Etica & Politica / Ethics & Politics
XVII, 2015, 2

www.units.it/etica

Edizioni Università di Trieste

ISSN 1825-5167
<table>
<thead>
<tr>
<th>MONOGRAPHICA I</th>
<th>Current Issues in Virtue Ethics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alessio Vaccari</td>
<td>Guest Editor’s Preface</td>
</tr>
<tr>
<td>Michael Slote</td>
<td>Saucers of Mud: Why Sympathy and Altruism Require Empathy</td>
</tr>
<tr>
<td>Julia Driver</td>
<td>Virtue and Moral Deference</td>
</tr>
<tr>
<td>Gopal Sreenivasan</td>
<td>A plea for moral deference</td>
</tr>
<tr>
<td>Lorraine Besser-Jones</td>
<td>Two Objections to Virtue Ethics</td>
</tr>
<tr>
<td>Sophie-Grace Chappell</td>
<td>Lists of the virtues</td>
</tr>
<tr>
<td>Julia Annas</td>
<td>Virtue, Skill and Vice</td>
</tr>
<tr>
<td>Caterina Botti</td>
<td>Feminine Virtues or Feminist Virtues? The Debate on Care Ethics Revisited</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MONOGRAPHICA II</th>
<th>The Legacy of Bernard William’s Shame and Necessity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alessandra Fussi</td>
<td>Guest Editor’s Preface</td>
</tr>
<tr>
<td>Alessandra Fussi</td>
<td>Williams’s Defense of Shame as a Moral Emotion</td>
</tr>
<tr>
<td>Alba Montes Sánchez</td>
<td>Shame and the Internalized Other</td>
</tr>
<tr>
<td>David Roochnik</td>
<td>Courage and Shame: Aristotle’s Nicomachean Ethics III.6-9</td>
</tr>
</tbody>
</table>
Laura Candiotto  
*Aporetic State and Extended Emotions: the Shameful Recognition of Contradictions in the Socratic Elenchus*

VARIA

Walter Block  
*The Trolley: a Libertarian Analysis*

Irene Dal Poz  
*The Concept of Ethos: Aristotle and the Contemporary Ethical Debate*

Pierpaolo Marrone  
*Condividere relazioni: riflessioni leibniziane a artire dall’esperienza dei social*

*Information on the Journal/Informazioni sulla rivista*
MONOGRAPHICA I

Current Issues in Virtue Ethics
Guest Editor’s Preface

Alessio Vaccari
Sapienza Università di Roma
Dipartimento di Filosofia
alessio.vaccari@uniroma1.it

Virtue ethics was originally presented by its proponents as a rather special normative conception that was radically different from that of what was known as the ethics of duty.¹ There were at least two important thematic lines behind this claim. The first concerned the way of delimiting the area of thought, behaviour and passions belonging to ethics², while the second concerned questions connected with the nature of practical deliberation.³

As regards the first aspect, those upholding virtue accused deontological ethics and utilitarianism of delimiting ethics to questions such as respect for other persons and their individual moral rights or the promotion of their happiness. In this description ethics is the combination of those institutions, rules and psychological dispositions that focus on how we should relate with others (with their freedoms, their goods, their desires) and does not directly concern the self, its needs and its development in its various dimensions. The proponents of virtue ethics have notably modified this way of delimiting the content of their area of enquiry. Ethics is that area of thought, behaviour and passions that concerns characters: that is to say, practical dispositions that involve passionsal and cognitive elements that arouse our approval and disapproval. A virtue is such, not only when it is the expression of our benevolent traits or those that inspire impartial behaviour, which are necessary for social cooperation, but also when it is the basis of those activities that mainly concern care of one’s self. The moral subject described by virtue ethics not only has benevolent traits or dispositions necessary for respecting the rules of group life, but also possesses other admirable qualities such as courage,

prudence, steadfastness in achieving one’s goals, and the vocation to seek and follow one’s deepest impulses.\(^4\)

As regards the second aspect, those upholding the virtues claimed that these structural features modified the nature of the thinking that was in play in deliberative processes. It was a transition that was generally described as a change in the way of understanding the kind of practical questions that are at the centre of ethics, which is marked by the shift from the question “what should I do?” to “what kind of person do I want to be?” In this reconstruction virtue ethics put on one side what had been regarded as the perspective of modern philosophy that limited itself to actions and invited us to consider more complex practical problems as to the kind of life we should live.

More recently, after the initial opposition between virtue ethics and deontological ethics and consequentialism, a new phase has begun, marked by a flowering of many different virtue ethics.\(^5\) Alongside new theories that offer to develop Aristotle’s moral philosophy we have seen various new currents that are working on theories of virtue elaborated in modern moral philosophy. In recent years, not only Aristotle, but also Hume, Kant, John Stuart Mill and Nietzsche have been the most frequent sources for conceptions of ethics that revolve around ideas of virtue and character. The interest in Kant’s *Doctrine of Virtue*, the development of a Kantian ethic of character and new positions on the role of the emotions in Aristotelian ethics have been the basis of recent attempts by Barbara Herman, Marcia Baron and Rosalind Hursthouse to show important areas of convergence between Kantian and Aristotelian ethics.\(^6\) Other perspectives, however, have shunned Aristotle’s rationalist and eudaimonistic ethics. Julia Driver has expressed the need to reconcile recognition of the centrality of virtue with a perspective that reduces the importance of moral knowledge and that reconstructs moral value starting from the consequences of actions.\(^7\) More recently, Julia Annas has formulated a virtue ethics that starts from the Aristotelian thesis

---


of virtue as a practical skill.\textsuperscript{8} A virtue ethics that draws on David Hume informs the work of Michael Slote\textsuperscript{9} and Lorraine Besser-Jones\textsuperscript{10}. It is a particularly promising conception as it combines a pluralist and sentimental perspective on the virtues with a conception of human nature centred on sympathy and on the self-conscious emotion of pride.\textsuperscript{11}

The essays in this collection are an expression of this second happy phase of contemporary thinking on the virtues. Each of them has a clear philosophical perspective behind it. Yet they are not presenting an abstract defence of their orientation, but are seeking, rather, to defend it by showing how it is best fitted to respond to important aspects of our moral experience or able to provide a better explanation of some constituent principles of human psychology.

Lorraine Besser-Jones examines how well eudaimonistic virtue ethics is holding up in the light of two recent objections: the \textit{self-effacing objection}, which claims that virtue ethics is problematic because it sets out to justify the virtues in a way that is not part of the motives of those acting in favour of virtue, and the \textit{self-centeredness objection}, which claims that virtue ethics, at least in its eudaimonistic version, expresses a selfish conception of ethics that does not take account of the other regarding passions that make up a central aspect of moral theory and of our shared ethical experience. Through an extensive discussion of the interdependent nature of the self, Besser-Joness shows how these objections are effective only in the false hypothesis that the self is egoistic. On the contrary, a virtue ethics that takes on board a relational conception of the self, supported, moreover, by influential research carried out by empirical psychology, can defend a version of eudemonics that is immune from these criticisms.

The essays by Gopal Sreenivasan and Julia Driver both discuss the theme of moral deference. More precisely, they examine the situations in which a moral agent defers to a so-called moral expert in ways that leads us to consider him as an agent that falls short of our ideal of moral agency. Using various arguments, these essays show that though moral deference is a less than perfect outcome from the point of view of agency, this does not constitute a proof that the agent is not virtuous or, more precisely, that she does not possess that specific virtue on which she is asking advice of the expert. As Driver observes, what matters in these cases

is that the agent is properly oriented in performing the action: that is, that the agent is responding to moral reasons although she may not be able to perceive them as such. The two essays, and particularly Sreenivasan’s, also show how virtue ethics, especially his version, which rejects the Aristotelian thesis of the unity of virtuous traits, has a greater chance of grasping this fundamental aspect of our everyday moral experience.

The theme of the limits of moral understanding, considered in the framework of feminist ethics, is again present in Caterina Botti’s essay. Using a complex and eclectic method of enquiry, which draws on psychology as well as philosophy, Botti underlines that, though the relations between human beings are fact of human life that cannot be by-passed, they are marked by a constituent opacity. In her contribution, Botti explores the effects of this epistemological thesis on how to describe the virtues at the centre of the ethics of care. On the one hand, care cannot be based merely on the spontaneous exercise of compassionate inclinations, but will presuppose the cultivation of our imaginative resources, our receptiveness and out attention, which bring better understanding of the characteristic, distinctive aspects of specific care situations. On the other, Botti’s perspective brings out how the effort to understand others meets a limit, a threshold beyond which the other is not knowable. According to Botti, our willingness to care must therefore be associated with the virtue of humility, a trait that expresses our awareness that our cognitive resources are limited and imperfect.

In continuity with his most recent works on care and moral sentimentalism, Michael Slote investigates the psychological causes of our altruistic inclinations. After distinguishing empathy – the psychological mechanism by which human beings and animals communicate their passions and opinions to each other – from sympathy, which is a term that identifies the active psychological principle that leads us to take care of others, Slote claims that there is a fundamental connection between these two psychological principles: empathy motivates sympathy. Slote’s enquiry offers to give a new and original explanation of this tie. Unlike psychologists such as Gregory Batson, Nancy Eisenberg and Martin Hoffman, Slote claims that the relation between empathy and sympathy/altruism is not an empirical, but a conceptual question. In this essay, Slote develops in particular Elizabeth Anscombe’s well-known thesis on the conditions of intelligibility of some desires. Following the structure of Anscombe’s argument, Slote claims that empathy constitutes a condition of intelligibility of our benevolent desires. Slote’s intention, with this explanation, is to provide an important argument in favour of a sentimentalist conception of the virtues. Slote shows that a perspective that appeals to David Hume’s teaching and that regards the mechanism of empathy as a constituent element of human nature can give a full account of that part of the virtues that coincides with the benevolent traits of character.

Julia Annas and Sophie-Grace Chappell examine the continuing relevance of Aristotelian ethics. Annas discusses the inexplicably neglected topic of the nature
of vice, while Chappell deals with the question of the method, or methods, for determining the list of the virtues.

In Annas’ contribution, an examination of our everyday moral experience reveals, in her view, the soundness of Aristotle’s analysis of this concept. Like virtue, vice too can be regarded as an internally unconflicted state of character. More precisely, just as the virtuous person, unlike the enkratic person, succeeds in doing the right thing without having to combat contrary motivation in order to do so, in the same way the vicious person is one who does not posses virtuous motives without feeling any regret for this lack. Secondly, Annas convincingly shows that vice, like virtue, is a psychological state that cannot be explained except as the result of a particular kind of upbringing. As a stable state of character, vice should be learnable in a way that virtue is learnable. According to Annas, the difference between these two states of character can be explained if we use the metaphor of a skill. Annas claims that every skill has its own intrinsic standard that concerns the acquisition of goods that are in some way intrinsic to that skill and that need to be pursued for themselves. Annas’ thesis is that though the vicious agent, unlike the virtuous one, can learn the skill, she will no longer be able to satisfy their standards. As she does not consider the goods internal to the skill to be pursued for their own sake, the vicious person merely acquires the skill because she considers it a means for pursuing other purposes.

In her fascinating contribution, Chappell reflects on the methods for identifying the virtues within a certain community. Chappell is impatient with Foot’s ethical naturalism and, more generally, with every form of foundationalism that claims to derive admirable character traits starting from a morally neutral description of the excellent exercise of human faculties. Following this approach, Chappell proposes a cautious and piecemeal methodology that brings together three different suggestions: a non-finalistic conception of Aristotelian eudaimonistic ethics, McIntyre’s argument about the virtuous traits necessary for the successful pursuit of human practices, and the aesthetic and emulative value of the living exemplars of a given virtue.

My aim was to assemble a cornucopia of varied current issues in virtue ethics. Yet, as is probably already evident from this preface, it is easy to identify some recurrent themes. For example, in their different ways the essays by Gopal Sreenivasan, Julia Driver and Caterina Botti all give attention to the issue of moral epistemology. Julia Annas and Sophie-Grace Chappell have a common interest in Aristotle’s legacy to contemporary virtue ethics. The notion of the social and empathic agent runs through the papers by Lorraine Besser-Jones and Michael Slote. And, taken as a whole, they show the endless fascination of virtue ethics.
Saucers of Mud: Why Sympathy and Altruism Require Empathy

Michael Slote
University of Miami
College of Arts & Sciences
mslote@miami.edu

ABSTRACT
Empathy and helping motives are more closely connected than philosophers and psychologists have realized. Empathy doesn’t just cause sympathetic concern for others, but is conceptually tied to it. When we empathize with someone’s distress at their pain, we ourselves are distressed by that pain and that in itself necessarily constitutes a motive to rid them of that pain. But helping motives like compassion or concern for others can be shown to be conceptually impossible in the absence of empathy. Compassion as a feeling and compassion as a motive are thus inseparable from one another, and this then lets the Chinese complementarity of yin and yang enter the picture. Yin can be viewed as a kind of receptivity, and compassion as a feeling instantiates such receptivity; but compassion as a motive instantiates yang conceived along somewhat traditional lines as a form of strong purposiveness. If moral sentimentalism is on the right track, then the motives and feelings it views as foundational to normative morality turn out to instantiate yin and yang conceived in traditional terms as an indissoluble complementarity. Moral sentimentalism properly pursued allows East to meet West in the field of ethics and possibly in other areas of philosophy as well.

KEYWORDS
Empathy, compassion, Yin and Yang, receptivity, strength, complementarity, sympathy, moral sentimentalism.

I

I have written a great deal in defense of a sentimentalist form of virtue ethics in recent years, and like David Hume, my work has placed a great emphasis on empathy (Hume didn’t have the word “empathy” but often spoke of what we mean by empathy using the term “sympathy”). I have argued that empathy not only motivates us to help others (and even to conform to deontological side-constraints and the moral demands of respect and justice), but actually gives us a criterion for distinguishing right from wrong across the complete spectrum of
possible cases.¹ But this is not the place to repeat those arguments. I have also defended the role of empathy in the making of moral judgments (though somewhat differently from the way Hume advocated such a role), but, again, this is not the place for me to try to recapitulate all that I have said in that direction.

So what am I going to do here? Well, as I said just above, my view subscribes to and depends on the motivating force of empathy, but until very recently I think I had a somewhat distorted view of how empathy motivates altruistic behavior or just plain sympathy with the plight of others; a distorted view, however, that I shared with some of the most significant psychologists who have written about empathy. Martin Hoffman, Nancy Eisenberg, C. D. Batson, and I myself (following their lead) have long believed that the relation between empathy and sympathy/altruism is an empirical issue, that human sympathy and altruism develop as a result of developing empathy and that this is an entirely contingent matter that we have to learn about from the science of psychology (or personal observation).² But I now think we have all been mistaken about this. I therefore propose, initially, to tell you why I think we have been confused on this subject, and this conclusion will prepare us for the main topic of the present essay, the question whether there can be such a thing as altruism/sympathy independently of empathy. If there can’t be such a thing, if such a thing turns out to be unintelligible, then the case for a sentimentalist account of morality will have been considerably strengthened.

² Let’s first talk about terminology. Most of us don’t find it very difficult nowadays to distinguish between empathy and sympathy. When Bill Clinton said “I feel your pain,” he was talking about what we now call empathy, and sympathy, by contrast, simply means a desire to see someone’s lot in life or present condition improved. And (though this is a point that hasn’t, I believe, been made in the philosophical or psychological literature) sympathy can be said to be a kind of minimum level of benevolence and of altruism/altruistic motivation more generally. A benevolent and altruistic person wants to help

another person, and sympathy may just be expressed in the desire to see the other person helped (by someone, not necessarily oneself). But I don’t think one can actually count as sympathetic in this latter sense or way unless there is something one would oneself do to help the other person. If one wouldn’t help the other person even if it was incredibly easy to do so, then any expression of sympathy would have to be considered hypocritical, “crocodile” sympathy. (I am reminded of Doctor Johnson’s complaint about people who just “pay you with feeling.”) But having made these points we still have a seeming conceptual divide: between empathy, on the one hand, and sympathy/compassion/benevolence/altruism, on the other. And I think that in the most important sense or way, this divide is actually illusory. Empathy’s connection with sympathy, benevolence, compassion, caring, etc., is conceptual, not empirical.

However, to make this point I first have to distinguish, as many philosophers and psychologists nowadays do, between two kinds of empathy. There is projective empathy, which involves putting oneself into the shoes or the head of another person (or animal), and then there is what is variously called associative, receptive, or emotional empathy, which occurs when we are invaded, so to speak, by the feelings or attitudes of another person. This is the kind of empathy Hume mainly spoke of using the term “sympathy,” and it is the kind that plays the most central role in my own approach to virtue ethics and that, as I now think, is conceptually tied to sympathy, etc. Empathy in some of its embodiments depends on a certain degree or amount of conceptual and cognitive sophistication. A child of four cannot empathize with the sufferings of the people in another country the way an adolescent or adult can, because they simply lack the requisite concepts. And when a father is infected by his daughter’s enthusiasm for stamp collecting, this too requires the father to know something about stamps, about collecting, and about his daughter. But the empathic infection occurs without the father consciously willing for that to happen, so we are talking here, not of the projective kind of empathy, but of the associative emotional kind.

And notice one thing. The father who is infected or, to switch metaphors, who takes in his daughter’s enthusiasm by a kind of empathic osmosis, doesn’t merely become enthusiastic in an unspecific or vague way. The enthusiasm has the same intentional object as his daughter’s, namely, stamp collecting. (There is some ambiguity or leeway here as to whether he starts wanting to help her collect stamps or starts wanting to collect for himself or both.) In other words, receptive or associative empathy takes in an attitude, motive, or feeling with its intentional object, and this is something that Hume seems to have recognized
when he pointed out in the *Treatise of Human Nature* that we humans have a strong tendency to take in the attitudes of those around us. Thus if my parents love Winston Churchill (mine did), one can and will, without knowing it, take in that attitude, but the attitude taken in isn’t just some generalized or vague form of positive feeling: it is positive feeling directed toward Churchill as its intention object.

These observations give us all we need to show that there is a conceptual, not a mere empirical, connection between the emotional or associative kind of empathy, on the one hand, and sympathetic, benevolent, altruistic, and/or compassionate motivation, on the other. If someone feels pain and is distressed about it, then they automatically, *ex vi termini*, count as motivated to alleviate that pain. That’s just what distress *means*. But then consider someone who empathizes with, who empathically takes in, the other person’s distress at their pain. This means feeling distressed oneself about their pain, and, again *ex vi termini*, this constitutes motivation to alleviate that person’s pain. Which is what we mean by altruistic or benevolent motivation. So on strictly conceptual grounds empathy involves sympathy with and motivation to help another person. (Of course, the motivation may not issue in action if stronger contrary motives are also in play in the given situation.)

However, those who accept the above argument often have an interesting way of (in effect) resisting its force and implications. They say (at talks I have given) that even if empathy entails sympathy on the grounds I mention, there still might be such a thing as sympathy *without* empathy. And till very recently I haven’t known how to answer them. I have tended to grant that there might in principle be such a thing as sympathy and benevolence without empathy in (extraterrestrial) species other than our own and have usually just insisted that in the human case, there is no such thing as sympathy or benevolence without a developed capacity for empathy. But now I think the case can be made stronger, that sympathy is conceptually impossible in the absence of empathy, and the very title of the present essay gestures allusively in the direction of the kind of argument I am now prepared to offer for that conclusion.

In her famous book *Intention*, published in 1957, Elizabeth Anscombe made a conceptual point that very much bears remembering.3 She argued (roughly) that

---

certain desires don’t in fact make any sense, that if someone claimed to desire a saucer of mud, we couldn’t attach any sense to what they were saying unless they went on to suggest some intelligible reason why they wanted this: e. g., someone had told them they would give them a lot of money if they gave them a saucer of mud, or they needed the mud for a beautifying facial, or some such. The point is well taken; it makes no sense to suppose that someone just wants a saucer of mud and has no further reason for this desire beyond the simple desire itself. But how exactly does this bear on the topic of the present essay?

Well, I think it bears immediately and, I hope, decisively. Those who questioned whether sympathy needs to be underlain by empathy were saying that one might just be sympathetic or benevolent without there being any further basis for that motivation, that such motivation might, for example, be simply instinctual. But I now think this is a mistake, a conceptual mistake, and my reason for thinking so partly comes from Anscombe’s earlier example. Though the matter is far from as obvious as what we (with benefit of hindsight) can say about the desire for a saucer of mud, I think, I now think, we in fact can’t make sense of there being benevolence or sympathy lacking any further basis or reason for the benevolence or sympathy.

However, in order to show you this, I think I need to begin by making use of a maneuver that Bishop Butler used to very good effect in his Sermons from the Rolls Chapel. Butler wanted to show that people are capable of altruism, but he recognized a certain difficulty standing between himself and that goal: the fact that people know that helping others can be a very rewarding and even pleasurable experience and are inclined to conclude that we sympathetically help others for egoistic, rather than altruistic, reasons. (There is some evidence that even Kant was taken in by this kind of thinking.) So Butler, rather ingeniously, changed the subject from benevolence and altruism to malice and revenge. He pointed out that the person who feels malice toward another may get pleasure from hurting them, but also typically risks his or her own happiness and comfort in their effort to do dirt to the person they hate. Hurting the person is their goal, and the only pleasure they get in this connection is from the fact, when it is a fact, that that goal is (thought to be) achieved. And this is something it is fairly easy for us to recognize. But once we see the case for regarding malice and revenge as non-egoistic motives, it is easy or easier to see that an analogous case can be made for regarding benevolent action and motivation as non-egoistic. And I propose to use a similar maneuver to deal with the issue of whether sympathy and benevolence can ever be sheer and basic. To help us with that issue, it will be useful to focus on the opposite of these feelings,
on malice and the desire for revenge, the very feelings or motives that proved so useful to Butler’s purposes.

Now no one could think that (the desire for) revenge could exist all on its own: revenge is always based on some offense or injury, imagined or otherwise. But this then contrasts, or seems to contrast, with malice (or malevolence). The Shakespearean critic A. C. Bradley once described Iago as having felt “motiveless malignity” toward Othello, and though this may not be entirely accurate to the play (Othello had passed over Iago for promotion at the time the play *Othello* begins), the idea of motiveless malignity doesn’t seem a contradiction in terms the way the idea of motiveless revenge does. However, I still think we should be suspicious about the notion of motiveless malignity. When we think of the malice that actually exists among or in humans, there always seems to be an element of revenge or some other deeper motivation for the malice. Iago bore malice toward Othello, but also had a motive of revenge against Othello, and just think about it. When you and I feel hatred toward someone, don’t we always think we have a basis or justification for feeling the way we do? Even the paranoid schizophrenic who deliberately injures others imagines that the others are out to get him or have done him dirt in the past, and so their malice and hatred toward others doesn’t, in psychological terms, stand on its own.

But perhaps the psychopath raises a problem here. Some (but I don’t think all) psychopaths have hereditary or congenital brain lesions or abnormalities. And many such psychopaths seem to want to hurt or harm others even if this in no way advances or promotes their own well-being. May such psychopaths not demonstrate the sheer malice that I am saying is conceptually impossible? Well, I am not sure. But I don’t think we should rush too quickly toward classifying such people as having motiveless malignity. Perhaps, for example, their brain/neurological deficiencies make them paranoid like some schizophrenics, which would undercut the claim that there is nothing motivating or behind their malice toward so many others. Alternatively, their brain abnormalities or earlier psychological/sexual abuse may make it harder for them to control or moderate their anger when unpleasant things occur or are done to them, and this may make them angry at the world in a way most of us aren’t. But once again, such anger and the malice that embodies it are not unmotivated or lacking in intelligible psychological grounding. So at the very least I am inclined to say the following.

Aside from bizarre cases like psychopathy that we don’t yet perhaps know enough about to characterize properly in moral-psychological terms, the idea of
motiveless malice doesn’t make a lot of sense. To say of someone “he just hates people, and there is absolutely no reason why he does” seems to me to be saying something quite difficult to make sense of. And in that respect I think the case is similar to Anscombe’s example. The simple desire for a saucer of mud is not something we can readily understand, and I say the same about the simple desire to hurt people, sheer malice. And if one can always wonder whether the neurologically damaged psychopath might not have a basic and unaccountable-for hatred of others, one can wonder too whether such a psychopath or someone else with a brain abnormality might not just have an unaccountable-for desire for mud. If we can’t rule the former out, how can we rule out the latter? But, turning the tables, might we not rather conclude that it is as difficult to make sense of the idea that brain malfunctioning might make us hate people for no psychologically operative reason as it is to make sense of the idea that such malfunctioning might make someone desire saucers of mud for no other reason than that desire itself. Every case of malice we are actually aware of seems to have some psychological basis other than the malice itself, and I can see no reason to think that any possible or conceivable malice could really be otherwise.

But then this argument transposes to benevolence, sympathy, and altruism. If the idea of sheer raw malice makes no genuine sense, why should the idea of sheer raw benevolence make any more sense? Of course, we are aware of our sympathy for others more vividly and/or self-consciously than we are of the (potential) empathy or empathic transmission of feeling that I say underlies sympathy. But this is an epistemic matter, not a causal or ontological one. Even if we typically know sympathy before we know empathy, it doesn’t follow that the former can exist without the latter. We may also know our own desire for revenge more vividly and immediately than we know the cause of that desire, but the former still depends for its existence on the latter. And similarly for lots and lots of other cases where the ordo essendi and the ordo cognoscendi proceed in opposite directions. We run into difficulties when we prescind from issues of knowability and just try to imagine malice with no psychological cause, and the case of benevolence/sympathy seems analogous. Imagine someone who is in trouble and someone who has sympathy for their plight. Doesn’t there have to be something that gets them to be or makes them sympathetic with the other person’s trouble? After all, a psychopath can recognize that someone is in trouble and feel no sympathy whatever for them, so the sheer recognition of another’s trouble doesn’t automatically arouse the sympathy of a bystander or onlooker. And in parallel doesn’t there have to be something in, something
about, the onlooker that leads them to feel sympathetic? The idea that someone might just automatically want another person’s suffering isn’t a very clear one, and the idea that someone might just automatically want another person to escape a difficult situation isn’t a very clear one either.

Now in the case of the psychopath, there is something missing, something which, if present, would allow and account for their sympathetic motivation toward someone who is suffering or in trouble. And the missing element or ingredient is, of course, empathy. As we have seen, empathy is a mechanism that converts distress on one person’s part to similarly-directed distress on the part of another person, and this yields or constitutes altruistic motivation and sympathy on the part of the person to whom the distress is conveyed. (This can happen with some non-human animals as well.) In such a case, we can understand, and understand very well, how altruism and sympathy can arise. But if we don’t posit empathy, then the motivation behind or for altruism seems difficult or impossible to fathom. And that is just what I am saying.

Now the reader may want to reply at this point that I am forgetting how easy it sometimes is to feel sympathetic concern in the absence of empathy. If (to take a famous example from Confucian thought) a child is about to fall into a well, can’t one feel concern and act on that concern even if the child doesn’t see their own danger and there is therefore no distressed state of the child to empathically latch onto? Yes, all of this is possible, but it only constitutes an objection to what I am saying about sympathy and altruism, if empathy exclusively takes in actual psychological states, and that assumption is far too limiting. It is possible to empathize with the distress, and suffering one knows someone will have if one does or doesn’t do something, and in the case of the child about to fall into a well, the observer can have a quite vivid sense of what the child will feel and suffer once they have fallen into the well. Sympathetic, caring, altruistic adults are capable of feeling empathy with what can or will happen to or in another person (the psychology literature on empathy discusses this possibility), and so in the kind of normal case of helping motivation the reader may have worried about, both empathy and sympathy are present.

In the end, therefore, I think sympathy and altruism without empathy are very much like a basic desire for a saucer of mud: something we really can’t understand. And at this point this comparison shouldn’t perhaps be so surprising. If desire has intelligibility conditions and cannot attach to some intentional object independently of those conditions, then sympathy and altruism, which involve specific (and positive) forms of intentionally directed
desire, may have their own intelligibility conditions; and can one think of a better candidate for such a condition than empathy?

In the light of the above, we might want now to consider the more general question whether empathy is necessary to any and every sort of moral helpfulness toward others. Sympathy and benevolence as I have been describing them don’t rest on any specifically moral or ethical thinking (this is another aspect of their naturalness in Hume’s sense of the term). When I benevolently help another person, I needn’t be thinking that this is my moral duty or that it would be a virtuous thing for me to do—I may simply be impelled by my sense of what the other is suffering and my empathy for their state of mind. But in describing things in these terms, I have remained pretty much entirely within the sentimentalist moral tradition that derives from Hutcheson and Hume. Rationalists have other ways of justifying and explaining the motivation behind actions that seek to promote the well-being of others, and it would be interesting to consider whether any of them allows coherently for beneficent actions based on something other than benevolence, compassion, empathy, or (psychological) altruism.

Kant seems, for example, to have thought that our rationality (or rational freedom) as such can lead us to help others. According to Kant, reason grounds the Categorical Imperative, and one of the duties that follow out of the Categorical Imperative is the (imperfect) duty to promote the welfare of others. But, in addition, Kant thinks pure reason not only grounds this obligation but makes it have a certain motivational force with us. However, rather than try to tease out Kant’s reasons for saying all this, I would rather talk about some recent ethical rationalists who I believe make the case for Kant’s conclusions easier to understand than Kant himself does. Let’s see if some more recent rationalist approach allows for helping motivation independently of an appeal to empathy.

I think John McDowell has made the overall best case, in rationalist or cognitivist terms, for the idea that empathy isn’t necessary to helping motivation. In his “Virtue and Reason” he argues (roughly) that if someone isn’t motivated to help someone in dire need or distress, that can only be because
they don’t fully appreciate what the other person is going through. A thoroughgoing apprehension of relevant facts can be automatically motivating, and in such cases, McDowell argues, the cognitive and the motivational are inextricably bound together—they can’t even be conceptually prized apart. Now McDowell’s view and others like it have been criticized as “queer” for tying the cognitive and the motivational so tightly together. But I am not going to object to this aspect of what McDowell is committed to. I can think of many cases where it seems plausible to suppose that cognition and motivation are inseparable and have discussed the matter at great length in my book *A Sentimentalist Theory of the Mind*. My objection to McDowell’s cognitivism/rationalism will come at his views from a somewhat different direction. We need to go back to the case of psychopaths.

What prevents a psychopath from fully appreciating how bad or painful it is or would be for one of his (potential) victims and being motivated to help them (or not hurt them in the first place)? Psychopaths are famous for being able to “get into the heads” of their potential victims (this presumably is projective empathy), so how can the rationalist like (the early) McDowell say that their lack of appropriate motivation is due to their lack of appreciation of relevant facts, to their not seeing certain facts as salient in the way a moral person would? Well, let me suggest that empathy may make the difference here and may be (part of) the only possible explanation of the difference of motivation between a psychopath and a moral person. Psychopaths may be able to get into the heads of other people, but they characteristically lack the ability or tendency to empathically feel what others feel. So if the psychopath fails to appreciate certain facts about another’s need or suffering, fails to see those facts as salient in a way that would motivate them to help, that may be precisely

---

4 McDowell, “Virtue and Reason” in R. Crisp and M. Slote eds., *Virtue Ethics*, Oxford: Oxford University Press, 1997. Incidentally, if someone argues against the analogy between malevolence/malice and benevolence that I have been making use of on the grounds that benevolence seeks to bring about something impartially good, the flourishing of other human beings, and malice seeks to bring about something bad, their suffering or faring poorly, they have actually given the game away. If benevolence is based in the thought of the goodness of what it seeks to bring about, then it isn’t the sheer desire for the welfare of another, but rather anchors itself in a conception of what is good in itself. So this sort of objection does nothing to show that sheer benevolence, the sheer desire for the welfare or happiness/non-suffering of another, makes sense.

because they lack (associative or emotional) empathy. It would then turn out that the lack of empathy makes it impossible for the psychopath to fully apprehend the suffering of others and to be motivated on that basis to help rather than hurt them. So McDowell’s purely rationalist/cognitivist route toward motivating helping behavior seems blocked, and once again we need empathy in order for such helping morally good behavior to occur.

Is there any other possible way for the rationalist to argue that empathy isn’t necessary to the kind of helping motives that morality (at least in part) depends on? Well, let me mention one other possibility. In What We Owe to Each Other, T. M. Scanlon argues that people have reasons for action and belief, that they can recognize such reasons, and that the reasons are capable all on their own of motivating actions or beliefs. For Scanlon, the notion of a reason is not reducible to any naturalistic notion, but can in any event be seen as equivalent to the idea that a given consideration favors a certain belief or action. Scanlon holds that many of us think we have reason to help others, and on his view the fact that it appears to one that one has a reason to help someone can on its own motivate one to help that person. But such a rationalistic view of moral or altruistic motivation (Scanlon is not talking about a “natural virtue” of benevolence here) makes no mention of empathy or of any need for empathy in order for the appropriate helping motivation to occur and eventuate in actual helping, and it is worth considering whether in fact Scanlon’s view can coherently avoid any appeal to empathy in the way it seems committed to doing.

I think that, as with McDowell’s views, the psychopath represents a stumbling block for Scanlon’s rationalistic views. One thing seems clear: even if there appear to some of us to be reasons to help others, such appearances don’t occur to the psychopath. It doesn’t seem to him or her that he or she has reason to help rather than hurt the person they want to victimize. (I am leaving aside cases where the psychopath has an egoistic and ulterior motive for wanting to help some other person.) And the best and most obvious explanation of why the psychopath doesn’t seem to see any reason to help is that they lack the kind of empathy with others that involves feeling what others feel. In that case, empathy and the capacity for empathy seem to make the difference between the psychopath and those of us to whom there appear to be reasons to help others, and so again, as with McDowell, the rationalist account that Scanlon gives of helping motivation seems essentially incomplete. The sentimental factor of

---

empathy has to be brought into any full picture of the phenomena here, and so I am inclined to conclude that any non-egoistic motivation to help others depends on empathy. Not only do the so-called natural virtues of sympathy, benevolence, and compassion require empathy, but empathy has to be brought into any possible account of non-egoistic morally mandated or desirable helping motivation.

But let me now mention one final way in which someone might want to claim that the latter sort of motivation might be explained without bringing in empathy. It could be said (this is the sort of thing Kant says) that the conscientious desire to do one’s duty can motivate someone who thinks it is their duty to promote the welfare of others to actually help other people, and it is surely far from obvious that the desire to do one’s duty and the recognition or belief that one has a duty to help others require empathy. There are in fact two ways one might go about answering this possible objection, one more critical, one more positive. The critical route will tell us to focus once again on the psychopath and ask why such people lack the desire to do their duty (and are incapable of guilt for moral failures). Surely, the absence of empathy will be part of the most plausible answer, and in that case one cannot invoke the conscientious desire to do one’s duty as motivating helping behavior without implicitly assuming that the person with such a desire is capable of empathy.7 But there is also a more positive way of answering the present objection to what I have been saying here. The objection effectively assumes that we can have moral concepts without having empathy, but many of us hold that psychopaths are like the congenitally blind. The latter lack full color concepts even if they eventually can tell you that grass is green and blood red. And by the same token one might say that psychopaths don’t really understand what terms like “right” and “wrong” mean even if they learn (for adaptive social reasons) to be able to tell people that stealing and killing are wrong.

But even if there is some initial tendency nowadays to think that psychopaths lack full moral concepts, can this view of them be supported in some more definite or positive way? Well, I think it can be, but that is a very long story. In my book *Moral Sentimentalism*, I argue—to some extent following Hume—that our second-order empathic experience of being warmed by the first order empathic warmth some agent displays toward some third party (e. g.,

---

7 Terminologically, the desire to fulfill one’s duty to promote the welfare of others can be regarded as a form of altruism even if it isn’t thought to be based on empathy, but I am using the term “altruism” more narrowly as a catch-all for all and only natural (in Hume’s sense) helping motives.
their friend or sibling) serves to fix the reference of the term “morally good” for us; and I argue that the way cold-heartedness empathically chills us helps fix the reference of “morally wrong” in a similar fashion. This is not the place to repeat the earlier arguments. But if the ideas just mentioned are on the right track, they serve to explain why psychopaths are like blind people with respect to the relevant concepts. Their lack of associative/emotional empathy undercuts their ability to fix the reference of moral terms in the way I believe is essential to the meaning, to a proper semantics, of moral language.

All in all, then, I don’t think we can have beneficent or helping moral motivation in the absence of empathy. This goes well beyond anything I or anyone else has previously said about the moral role of empathy, and it moves us toward the conclusion that empathy is probably the most significant factor in the moral life.

Conceptual/philosophical speculation is a very risky business, and I have been doing quite a lot of that here. But in this concluding section, I would like to speculate further and in a new direction. Let’s say that the above discussion supports the idea of an inextricable connection between empathy, on the one hand, and sympathy, benevolence, etc., on the other. One can’t have either one of them without the other. But consider how this relates to the fact that we think of sympathy, benevolence, compassion, and the like both as feelings and as motives. The above argument shows or seeks to show that the feeling side of, say, compassion is irrecusably tied up with the motivational side, and this in and of itself seems to me to be an interesting result. It can seem strange (it always has to me) that compassion is considered to be both a feeling and a motive, and the argument I have given helps explain how that can make sense. If the empathy/feeling side of compassion and the motivational side of compassion cannot be separated, then one can see how it makes sense to hold, as common sense does, that compassion is a feeling and also a motive. And exactly the same points can be made about benevolence, caring, and sympathy.

But now I want to make what will seem to most of you like an incredible leap of topic. I think what I have just been saying offers a philosophical foothold for the ancient Chinese complementarity of yin and yang, and if that is the case, moral sentimentalism illustrates some themes that go beyond Western culture. Now yin and yang are nowadays not thought to be serious topics for
philosophical thinking—even by the Chinese. Like us Westerners they are accustomed to various popularizations of yin and yang—as with macrobiotic diets; and they are aware, as most of us Westerners are not, of how ancient yin/yang explanations of physical phenomena (e.g., of how sunlight differentially affects the two sides of a hill) have had to yield to more quantitative and mathematical explanations of such phenomena of the sort that are the mainstay of (elementary) modern physics. But despite these problems or limitations, I think suitably updated notions of yin and yang can be useful for present-day philosophical purposes, and I am going to try to persuade you of that here and now.

What do I mean by updated versions or notions of yin and yang? In a recent article, I have argued that we can make the most ethical sense of yin and yang via the Western notions of receptivity and active/rational control.\(^8\) Yin is often equated with passivity and often with pliancy or pliability, but it is also often equated with receptivity (there is no term in Chinese for “receptivity” and “yin” may be the closest that language comes to our notion of receptivity). And I think that, unlike passivity and pliability, receptivity is a positive and broadly valued quality that, equated with yin, can be counterbalanced with or against the quality I am proposing to equate with yang, the quality (and notion) of active/rational control.

I have argued elsewhere that Western philosophy has tended to emphasize active/rational control at the expense of the value and virtue of receptivity, but the point then is that we need and need to value both active/rational control and receptivity in our lives and thought. And I think these two qualities can be viewed as necessary complements in the moral or ethical life. Again, I have made the arguments for this conclusion elsewhere.\(^9\) But for present purposes and given what was argued earlier, something very interesting (I think) follows if we conceive yin and yang in this updated philosophical way. When we empathize with the distress of someone who is in pain, we are receptive to them in a way the psychopath never is with anyone. And when we ipso facto are then motivated to help (remember, though, that this doesn’t mean we actually will help—other motivational factors may override our compassion), we are motivated to actively do something effective as a means to alleviating the pain of the other person; and this motivation to help shows us as active and

---


interested in exerting control over what will happen to the other person. (Rationality comes in because if one doesn’t, cognitively, seek to find or learn about the best means to help the other, the fact of compassion is criterially challenged or undercut; one doesn’t count as compassionate if one is slapdash about finding proper means to helping the other person.)

So I am saying that compassion, benevolence, and the like have the yin quality of receptive feeling and the yang quality of desiring actively to help in a specific way—they have both of these at the same time and, as I have been arguing, each aspect is inseparable from the other. And this gives yin and yang a deeper, further foothold in our discussion because that complementarity is traditionally viewed involving just such an inextricable or irrecusable relationship. The traditional symbol of yin and yang depicts yin with a small circle of yang in it and yang with a small circle of yin in it, and this is one way to symbolize the ancient view that yin and yang is a necessary complementarity, that yin and yang are really yin/yang.

But the present discussion gives these ancient and philosophically somewhat vague (and suspect) notions a particular and definite embodiment. If you can’t have compassion as feeling without compassion as motivation and vice versa, then you can’t have a certain sort of receptivity without also having a certain sort of control-seeking activeness and vice versa; and if one buys my updating of the notions of yin and yang, then in the sphere of moral sentiments you can’t have yin without yang or yang without yin and they are invariably instantiated together. The (valued or positive) moral sentiments thus all have a yin/yang character, and that is a philosophically significant fact both about the sentiments and about the ancient Chinese complementarity of yin/yang.

But if moral sentimentalism lends itself to an interpretation via the Chinese categories of yin and yang, we really shouldn’t be too surprised. What we call moral sentimentalism had its origins, in the modern West, in eighteenth-century Britain, but there is a strong element or aspect of sentimentalism in traditional Confucianism: in Mencius and in neo-Confucians like Cheng Hao and Wang Yangming who were strongly influenced by him. However, the specific idea that yin/yang applies to compassion and other particular moral sentiments doesn’t seem to have occurred to any Confucian or neo-Confucian (or later Chinese) philosopher, so what I have just been saying is intended as a contribution to the overall Confucian tradition at the same time that it represents, as I believe, a philosophical application of yin/yang to or within moral sentimentalism. I also think yin/yang has applications outside of ethics, but that is a long story to be told on another occasion.
Virtue and Moral Deference*

Julia Driver
Washington University in St. Louis
The Department of Philosophy
jdriver@artsci.wustl.edu

ABSTRACT
Moral deference has come under attack for undercutting moral understanding, and thus, for undermining virtue. This essay defends the view that moral deference and virtue are not incompatible. Indeed, responsible moral deference is a kind of coping virtue, a trait that has value in counteracting either and agent's own deficiencies, or conditions that undermine the application of moral understanding in a particular practical situation (such as when the agent has insufficient time to deliberate).

KEYWORDS
Virtue, moral deference, moral understanding, moral worth.

A consensus has been reached that moral knowledge can be transmitted via testimony (Jones 1999; Driver 2006; Hills 2009). That is, one’s belief, say, that “x is wrong” can be true, and can be justified solely on the basis of the testimony of a reliable and trustworthy authority. Yet, even granting this, one might still hold that a person ought not to rely solely on the testimony of someone else in deciding what one ought to do, that such reliance exhibits moral failure. Alison Hills, for example, argues that deference to moral testimony exhibits a lack of moral understanding, and it is moral understanding and not moral knowledge that is the “centrally important concept in moral epistemology” (Hills 2009, 97). Moral understanding is quite distinct from moral knowledge. Attributing knowledge to an agent with respect to a particular

* An earlier version of this paper was presented at a workshop on Moral Testimony at the University of Zurich, at the 2014 meeting of the Scots Philosophical Association in St. Andrews, at a meeting of the Saint Louis Ethics Workshop in December 2014, at Duke University, at Hebrew University, and at the 22nd Annual Philosophy Graduate Student Conference at the University of Waterloo. I would like to thank the members of those audiences for their very helpful comments, and in particular Ann Margaret Baxley, Georg Brun, Michael Campbell, Timothy Chappell, David Enoch, Kathryn Lindeman, Walter Pedriali, Walter Sinnott-Armstrong, Paulina Sliwa, Gopal Sreenivasan, Hasko Von Kriegstein, and Eric Wiland.
proposition is too thin to capture what we care about in assessing actions and in assessing an agent as a *moral* agent. This is because there is both something morally defective in the action an agent performs on the basis of pure deference to testimony, and something morally defective in the agent herself who needs to rely on testimony. What is missing is moral understanding, which is unpacked in terms of the agent possessing a full appreciation of the reasons that justify the action. When an agent possesses moral understanding, and acts in light of that understanding, her action possesses moral worth, and the possession of that understanding underlies moral virtue in the agent. Without such understanding, the action lacks moral worth and the agent fails to exhibit virtue. In this paper I assume for the sake of argument that understanding is very different from knowledge (though this has been very effectively challenged by other writers, see Riaz 2015). However, I disagree that *all* actions performed purely on the basis of testimony lack moral worth, and that *all* agents who rely on testimony lack moral virtue.

*Deference*

What is justified deference?

D defers to A that *m*, iff: D’s judgment that *m* is adopted *solely* on the basis of A’s judgment that *m*; D’s deference is justified given that A is deemed by D to be the appropriate relative authority on the basis of [an appropriate level of] evidence that A is a reliably good judger in moral matters [within the appropriate domain of expertise], and must also responsibly judge A to be trustworthy regarding the testimony provided.

For the deference to be justified the deferrer must be justified in regarding the person providing the testimony as trustworthy and the relevant sort of relative authority. By ‘relative authority’ it is understood that A is an authority on *m* relative to D. In many cases, A will be an expert relative to D, though this need not be the case. For example, A might simply be more competent relative to D in some way that is time or situation specific. D might be drunk, for example, and A sober.

The word ‘solely’ is important. All are in agreement that moral knowledge can be responsibly and appropriately acquired via testimony *without* deference when the authority makes clear to the agent what the reasons are and how they are to be weighed, and the agent takes that on board; that is, she sees and appreciates what the reasons are and how they are appropriately weighed. The problem cases are those in
which the agent must defer because she either doesn’t see the reasons at all, or has no understanding of their appropriate relative weight.

**Moral Understanding**

Hills develops an account of moral understanding that is very demanding.

If you understand why X is morally right or wrong, you must have some appreciation of the reasons why it is wrong. Appreciating the reasons why it is wrong is not the same as simply believing that they are the reasons why it is wrong, or even knowing that they are the reasons why it is wrong. Moral understanding involves a grasp of the relation between a moral proposition and the reasons why it is true. (Hills 2009, 101)

For Alice to understand why stealing is wrong, she must grasp the reasons why it is wrong via an appreciation of those reasons, and that goes beyond simply knowing that it is wrong. Further, moral understanding does not hold for isolated facts. Moral understanding requires a “systematic grasp of morality.” It would be odd for Mary to realize that she should be nice to Sandra, to avoid causing Sandra pain, and yet not be able to grasp that this counts as a reason to be nice to others as well, and have some understanding that this is because pain is bad. Understanding requires the following abilities:

If you understand why \( p \) (and \( q \) is why \( p \)), then in the right sort of circumstances, you can successfully:

(i) follow an explanation of why \( p \) given by someone else;

(ii) explain why \( p \) in your own words;

(iii) draw the conclusion that \( p \) (or that probably \( p \)) from the information that \( q \);

(iv) draw the conclusion that \( p' \) (or that probably \( p' \)) from the information that \( q' \) (where \( p' \) and \( q' \) are similar to but not identical to \( p \) and \( q \));

(v) given the information that \( p \), give the right explanation, \( q \);

(vi) given the information that \( p' \), give the right explanation, \( q' \).

(Hills 2009, 103)
Because understanding is itself so demanding, one can know that \( p \) without any understanding of \( p \). Acting on knowledge without understanding both undermines the moral worth of the action and exhibits the lack of virtue in the agent.\(^1\)

Friends of moral deference hold that moral deference is not only appropriate, but also sometimes required, especially in cases where the harm in failing to defer is quite large.\(^2\) Suppose that Mary believes, with very good reason, that she herself tends to become flustered when required to make decisions about her elderly mother’s nursing home care, and that her sister, Donna, is much better able to make decisions that protect their mother’s interests. It would seem that, for the sake of her mother, she ought to defer to Donna’s judgment on how their mother’s care should proceed, and follow Donna’s advice about what instructions to provide the nursing facility. Here, Mary has enough self-awareness to realize that there is a feature of her temperament that interferes with her practical deliberations on a specific topic.

Do Mary’s actions lack moral worth? Given how Hills characterizes moral worth, they must lack it. The notion of ‘moral worth’ has a long history, and accounts of moral worth attempt to capture Kant’s idea that some actions are deserving of a special moral esteem in virtue of being properly motivated, or performed for the right reasons. Hills unpacks it the following way: “Your actions is morally worthy only if it is a right action performed for the right reasons…” (Hills 2009, 113) Kant’s honest shopkeeper case is supposed to distinguish right action from morally worthy action (Kant [1785] 2002, 397). The action of an honest shopkeeper, who correctly charges his customer, and who is motivated by self-interest, is right, but lacks moral worth. This is because the honest action is done from a motive of duty. The honest action that is motivated duty, on the other hand, possesses moral worth. For an action to be morally worthy the person performing the actions needs to act for the reasons that make the action right. What makes charging one’s customers the correct amount for their purchases is that duty requires it. Self-interest is not what properly justifies honesty. Of course, Hills is not at all committed to the way Kant happens to spell out the proper justification for moral actions. She is simply making use of the idea that there is, intuitively, a distinction between an action’s rightness, and something else, call it “moral worth”, which picks out a special sort of esteem for actions performed with the proper motivations, however we spell out those motivations.

There is an ambiguity here. Nomy Arpaly points notes that there is a difference between acting for reasons known or believed by the agent to be moral reasons, and

\(^1\) Hills discusses other ways in which moral understanding is important, but the focus on this paper will just be the issue of how understanding is allegedly undermined by deference, and how this relates to moral virtue.

\(^2\) David Enoch focuses on cases in which the risk of significant harm is quite large. In these cases it seems clear one ought to defer. See his “A Defense of Moral Deference” (2014).
acting in response to moral reasons, even if the agent does not perceive them as such.\(^3\) On her view an action has moral worth if it is done for the right reasons, and “done for the right reasons” is understood in terms of the agent simply being responsive to the right reasons – she need not herself think of the reasons as the right ones. The case of Huckleberry Finn is an illustration: Huckleberry has, out of genuine feelings of sympathy, helped a friend of his, Jim, escape from slavery. However, Huckleberry views what he has done as wrong since he doesn’t question slavery itself. Does Huckleberry’s act of helping Jim have moral worth? Arpaly correctly notes that it depends on how we reconstruct the case. If we understand Huckleberry as someone who is not accidentally, or whimsically, doing the right thing, as someone who is rather “…racist in conscious opinion but viscerally more egalitarian…” (Arpaly 2002, 229) then we might see Huckleberry as someone who, though lacking in moral insight still sees Jim as a person, and thus someone whose action has moral worth. I make a similar point about Huckleberry Finn in Uneasy Virtue. I note there that Huckleberry, “…though lacking a correct conception of the good, was still acting in accordance with the correct conception of the good….In order to be virtuous…one need not know that what one is doing is good or right.” (Driver 2001, 52) Huckleberry’s actions were virtuous, deserving of praise, and thus morally worthy, though he lacked systematic moral insight.

What seems crucial for moral worth is that the agent be “properly oriented” in performing the action. One way to spell this out is to hold that one’s actions are only properly oriented if they are done for the reasons that justify them. But this is not the only way to spell this out, and there are alternatives that would make deference to another’s testimony display a proper orientation.

One of Hills’ examples of an action that lacks moral worth is the following, which is a modified version of a case discussed by Arpaly:

**The Knowledgeable Extremist** – Ron is an extremist, believing that killing a person is not generally immoral but that killing a fellow Jew is a grave sin. Ron would like to kill Tamara, but he refrains from doing so because he wants to do the right thing, and he knows (on the basis of his rabbi’s testimony) that the right thing to do is to refrain from killing her. (Arpaly 2002, 115)

When Ron refrains from killing Tamara, he does the right thing. Hills believes, though his action clearly lacks moral worth. Though he is motivated by a de dicto desire to do what is right, he does not see that the fact that Tamara is a person gives him sufficient reason not to kill her. This is certainly a failure to be properly oriented. That much seems clearly right to me. Ron’s action lacks moral worth. However, I

---

don’t believe that the lack of proper orientation that Ron displays is best unpacked as a mismatch between his reasons for acting and the justifying reasons for the action. Though, I also believe that Arpaly may be mistaken in her analysis of Huckleberry Finn. This might be a good example, though, of a case in which virtue and moral worth part ways. Huckleberry is described as someone who does the right thing, even though he thinks it is the wrong thing to do, so it is odd to view him as having a desire to do the right thing – he only has the desire to do the thing that, in fact, is right, but not under the description of it being right. So Huckleberry does not act out of a de dicto desire to do the right thing. If this is necessary for the moral worth of an action, but not sufficient, his action lacks moral worth.

Consider the case of Matilda:

Matilda wants very much to do the right thing, and has been brought up to believe that lying is wrong. However, she has a friend, Marilyn, who she justifiably takes to be an expert on morally appropriate social interactions, and Marilyn assures her that sometimes lying is not wrong, it just depends on the circumstances. Indeed, sometimes lying can be morally required. This puzzles Matilda. One day both she and Marilyn are attending the same baby shower. Marilyn, understanding Matilda’s predilection for scrupulous honesty, tells her “You ought not to say the baby is ugly, or that the baby doesn’t look like his parents at all – that would be wrong because it would hurt the parents’ feelings to no good end.” Again, this puzzles Matilda. It’s not that she fails utterly to see the moral significance of hurting people’s feelings, but she simply has no feel for the trade-offs Marilyn’s judgment requires. However, she defers to Marilyn and though she does in fact think that the baby is ugly, and though she is asked her opinion, she lies and says the baby is quite lovely.

In certain structural respects, this case is similar to Ron’s. Matilda wants to do the right thing – she has a de dicto desire to do the right thing. There is a mismatch between her reasons for acting and the reasons that justify the action. She is acting because she wants to do the right thing, and she knows, from Marilyn’s testimony, that lying is the right thing to do in this circumstance. She lacks moral understanding because she clearly lacks the sort of systematic understanding of morality that is required on Hills’s view. However, unlike Ron, she seems properly oriented. Her motivations, more broadly considered, are basically good ones. She wants to do the right thing, she has some idea of what sorts of considerations go into moral justification, but lacks more precise and systematic grasp of the relations
between reasons and competing considerations.\textsuperscript{4} She is not \textit{utterly} unresponsive to the right sorts of reasons, unlike Ron, but she lacks the sensitivity to those reasons that would allow her to make the finer judgments and discriminations that Marilyn is capable of.\textsuperscript{5}

It might be useful here to make a distinction between the act of deferring itself and the action that is the result of the deferring. Hills might hold that in the case of Matilda, her decision to defer to Marilyn is done for the right reasons, and thus has moral worth, though when she tells the lie she does not act for the right reasons, and so \textit{that} action lacks moral worth. I don’t think that this is open to Hills, really, if we take her at her word on what is required of moral understanding. \textit{None} of Matilda’s actions possess moral worth, if what is required is a systematic understanding of morality. But maybe this can be weakened a bit. When Matilda decides to defer to Marilyn she is doing so out of a desire to do the right thing, and surely that, along with the belief that Marilyn knows best in this case, is what justifies the deference. But this does not seem sufficient for proper orientation. Otherwise, Ron’s decision to defer would have moral worth. Thus, there seems to be something more that is required for an action to have moral worth. Even in deciding to defer Ron displays a failure to appreciate, at any level, the sorts of reasons that count against something – the sorts of reasons that factor into an action’s rightness or wrongness.

Does deference undermine moral virtue? Consider the case Hills’ uses to motivate the intuition that it does:

\textit{The Incompetent Judge} – Claire has just been appointed as a judge and is very anxious to sentence people justly. But she finds it exceptionally difficult to work out the just punishment for various offenses, though she listens to the evidence presented carefully and tries her best to get the right answer. Luckily she has a mentor, a more experienced judge, Judith, who has excellent judgment. Claire always consults with Judith and gives her decision in accordance with Judith’s

\textsuperscript{4} One might hold both Ron and Matilda to be morally fetishistic, since they act from a \textit{de dicto} desire to do what is right. However, that someone is morally fetishistic is not really problematic. See Jonas Olson, “Are Desires \textit{De Dicto} Fetishistic?” (2002). Further, Matilda does not seem to be pure case, since she has awareness of the moral reasons, and is just unclear on how to weigh them (though it is true that it would be hard to characterize her \textit{de re} desire to do the right thing).

\textsuperscript{5} Eric Wiland (2014) notes that the focus in the literature has been on what is wrong with those who need to defer on matters such as “Suffering is wrong” and deference to something like that is odd indeed. The more realistic cases are those that involve weighting different factors. Wiland discusses this issue in light of Ross’ distinction between prima facie duties and duties proper. We are clear on our duties, but weighing them is more complicated since an act in one way may be prima facie wrong, but in another way prima facie right.
guidelines, offering Judith’s explanation of why the sentence is just to the defendants.

Claire lacks moral understanding, and thus lacks virtue. This is another instance in which I agree with Hills’s verdict on the case itself, but disagree on the lessons to be drawn from it.

In rendering the just verdict Claire is not responding to considerations of justice (i.e. the right reasons). Rather, she is responding to Judith’s judgment – and on Hills’s view this is the wrong sort of reason, at least if we want to characterize the agent as virtuous. This is because she holds that the virtuous person must be a moral authority, an authority on what is right (Hills 2009, 111-112). Further, the virtuous person is someone whose “…thoughts, decisions, feelings, and emotions as well as her actions…” are “…structured by her sensitivity to morality” (Hills 2009, 112). Claire, like Matilda, is someone who is motivated to do what is right. What she lacks is good judicial judgment, and this is what undermines her virtue, on Hills’s view. But we can agree with Hills that Claire lacks virtue. The critic only needs to find cases where deference is responsible, and even virtuous, the critic need not hold that all cases of deference that lead to the right decision are indeed virtuous.

Human beings are subject to limitations, both epistemic or cognitive, and temperamental. We often need to seek out advice to make responsible decisions. Of course, it would better if we had no limitations, and we did not need to seek out advice. But this shouldn’t detract from our capacity for virtue. A responsible moral agent will want to do the right thing more than she will want to exhibit the sort of moral understanding Hills regards as necessary for virtue. She should care about that more than her own possession of virtue. In cases where she is unsure of which reasons obtain, or have greater weight, as in the case of Matilda, the desire to do the right thing should lead her to get moral advice. But it may be that in spite the advisor’s best efforts she still doesn’t grasp how it is that one reason has greater weight than another. She knows what the right thing to do is. She must defer. Hills has defined virtue so that moral understanding is a part of it. It looks as though being concerned with the moral worth of one’s actions, and with virtue, are obstacles to doing what is right since doing what is right in situations where deference is required for doing what is right requires that one act in such a way that one’s action lacks moral worth, and one is failing to exhibit virtue. On my view, responsible moral agency requires deference in such cases, and given the very plausible assumption that those who are responsible moral agents possess some virtue, we have an argument that responsible moral deference is required of the virtuous person in some circumstances. Responsible moral deference is not incompatible with virtue.
When we rely on advice we can do so responsibly or irresponsibly. In order to responsibly rely on advice we need to consider factors such as the reliability of the putative advisor, along at least two parameters: does the advisor reliably come to correct judgments on the matter at hand, and is the advisor trustworthy – someone who can be relied on to honestly communicate those judgments?

A person who seeks to make important decisions in isolation, without seeking out advice, does not represent the relevant sort of ideal for us. Again, certainly, it would be better if we all knew everything we needed to know to engage in effective practical deliberation on all issues of moral significance, and if we possessed a fully systematic understanding of morality. It would be better if we had capacities that allowed us to use that knowledge more efficiently – if we could think more quickly, for example. But again, given our limitations, seeking out advice from reliable sources is not just in keeping with virtue, it is required of it. And a similar argument can be made – by extension – to deference. The conditions for responsible deference are the same as those for responsibly trusting advice. The difference is that when one is relying on advice as opposed to deferring, one comes to appreciate for oneself the moral reasons at play in the decision or judgment and their relative weight and significance. Full deference involves lack of moral understanding. A person who seeks out and relies on advice either possesses moral understanding already or comes to possess that understanding through the advice. But if we are operating with the very demanding notion of moral understanding articulated by Hills, it will be hard to motivate the distinction between pure deference and relying on advice. If the person possesses moral understanding to begin with, the only advice that person will need will be on purely empirical matters. For example, I may need advice on which charity is the most efficient in making a decision about where to send a contribution. No moral advice is required. However, if a person does need moral advice, it seems incredible to me to suppose that any single provision of that advice will result in moral understanding – given that this requires a systematic understanding of morality. Thus in practice, as a practical matter, on Hills’s view the distinction between moral deference and advice will be difficult to make.

Further, the view that a person who relies on deference is acting incompatibly with virtue relies on a very narrow understanding of what a virtue – or excellence of character – is. One way to approach virtue is to ask what sorts of character traits a morally conscientious person would want to have – including, what a morally conscientious person who is not perfect – would want to have in light of those imperfections. A similar issue arises in the debate between possibilists and actualists: should I do the best that I think I can do, or should my decisions about what to do be guided by my views about what I will do, given various features of my character? The classic example used to illustrate the distinction involves a procrastinator who
needs to decide whether or not to agree to write a book review (Jackson and Pargetter 1986). The best option would be to agree to write the review and then write it in a timely manner. If the review is not written in a timely manner, however, it would be better not to agree to write it in the first place. The procrastinator is aptly named – she knows that she has a tendency to procrastinate and that it is unlikely that she will finish the review in a timely manner. What should she do? The actualist holds that she should not agree to write the review – and, after all, isn’t that the advice a friend would give her? The possibilist, however, believes that she should agree to write the review, and then write it in a timely manner. Taking into consideration what she will likely do, given her character flaw, is letting herself off the hook. The connection to the question we are considering here has to do with how we should approach the question of what we should be like, given that we have flaws and limitations? This is taking an actualist perspective to the question of virtue. However, the question we are considering isn’t wholly analogous to the question of what we should do given our limitations – in the actualist/possibilist debate the actualist concedes that it is possible for the agent to do what is best, it is just very unlikely. Here, though, we also consider cases where it isn’t possible for the agent to make the right decision on her own, without deferring. In this case, even someone who leans in favor of the possibilist would hold that the deference is warranted, and given that the agent has certain limitations, the deference is good, not bad. Thus, the opponent of deference when the agent cannot make the right decision on her own, the opponent who thinks that such an agent is lacking in virtue, is in the odd position of holding that in deferring she is doing everything she ought to do, for good reasons, too, and nevertheless is exhibiting vice in so deferring.

Can we make a stronger case for deference counting as virtuous? Some virtues seem to be ‘coping’ virtues. They are virtues that we need or that benefit us precisely due to certain character deficits. Philippa Foot had the view that all virtues were correctives: “...each one standing at a point at which there is some temptation to be resisted or deficiency of motivation to be made good” (Foot 1978, 8). Her claim is both broader and narrower than the one I would like to make. It is broader in that she views all virtues, not just a subset, as corrective. It is narrower in that she restricts what needs to be corrected for, or what needs to be managed in some way, to those deficits regarding our abilities to withstand temptation and be properly

---

motivated. I agree that these are two prominent ways in which a sort of deficit impedes practical deliberation, but there are other ways – some relating to features of the person herself, and some relating to that person’s circumstances. One person may benefit from a strong will precisely because she is tempted. We can call these ‘coping’ virtues. Deferring to greater authority on important moral matters may be such a virtue. It is only because Matilda lacks a systematic grasp of the reasons that she is required to defer. And there will be other cases where the agent has difficulty trying to figure out which reasons among a range of possible reasons are the right ones. So, for example, someone might know lots of reasons that are not the right ones, but be unsure about which ones are the right ones. Consider Melissa, who has taken a class on contemporary moral problems and comes to believe that abortion is morally permissible. She knows that there are good arguments for not viewing the human fetus as a person; she also knows that there are good arguments to the effect that even if the fetus is a person, abortion is still permissible. She doesn’t know what to think about fetal personhood at all, but she does know that either way, whichever is the right way to go, abortion is permissible. Here she isn’t able to identify which line of justification is the correct line of justification; just that permissibility ‘dominates’. Does this qualify as appreciating the reasons that justify the permissibility of abortion? Or does the agent need to get things precisely right? It seems that given Hill’s criteria such an agent lacks moral understanding. This seems highly counterintuitive to me, but perhaps this shows that we simply need to acknowledge different ways in which someone can possess moral understanding. One way is through a generalizable appreciation of the reasons that morally justify actions in such a way as to be able to apply them to particular cases; another way is by appreciating reasons more generally, without, in some cases, being able to tell which ones apply and/or how they apply, as in Mary’s case. In the latter sense, then, someone who completely doesn’t see that causing pain is a bad thing, and a wrong making feature of an action, is lacking moral understanding. Melissa, however, is not like this, though she lacks the sort of appreciation that allows her to properly identify the operative reasons.

There are coping virtues that help people counteract deficits they have, that may or may not be typical for the average human being. Some virtues, such as graciousness, involve responding appropriately to one’s own failure. This may include things like sincerely apologizing, not allowing one’s disappointment to interfere with showing the correct respectful behavior towards others. This raises the issue of diachronic versus synchronic coping. The latter example of graciousness

Foot seems mainly worried about deficits that afflict persons in general – e.g. so that a virtue is still a corrective even in cases where a person is not tempted to do the wrong thing. However, my reading of “Virtues and Vices” holds her as allowing that virtues also correct for individual deficits.
could be handled by the Aristotelian by noting that the gracious person has become virtuous – they had lacked it when they behaved badly – but the recognition of their own bad behavior and what the appropriate response is to said behavior shows that they have grown into virtue – thus, the coping is diachronic – not at the same time as the lapse – and this is not at all incompatible with the Aristotelian view. Synchronous coping, on the other hand, would be – because it involves coping with a lapse as one is experiencing or displaying the lapse. This doesn’t affect my overall point, since synchronous coping can still be virtuous – and responsible deference is an example of synchronous coping.

There are a variety of deficits someone might need to cope with, and how one copes says a lot about one’s level of virtue. There are epistemic deficits – some may involve features of the person, such as an inability to focus on solving a problem, confusion over what sorts of reasons apply in a given case, or lack of the relevant sort of experience; some may involve features of a situation a person might find herself in, such as lack of time to carefully think through the options, or the presence of other environmental factors that impede practical deliberation. Other deficits might be temperamental, such as a tendency to get angry too quickly. Yet other deficits may have to do with attention, or salience (perhaps these also qualify as epistemic). Perhaps Sally has a tendency to get lost in her own world as she walks home from work, and thus fails to notice the suffering of others around her, or perhaps she is too easily distracted by kitten pictures on social media, and fails to attend to more serious issues.

These deficits, in one way or another, undermine one’s ability to make the right judgment, not necessarily one’s virtue. In any of these sorts of situations, it may make perfect sense to defer to a trustworthy expert. It is actually required of the virtuous agent.

The coping virtues fall into the category of what Robert Adams calls ‘structural’ virtues (Adams 2009, 33ff.). He contrasts these virtues with what he calls ‘motivational’ virtues. Motivational virtues are defined by their good motives: so, benevolence must involve “desiring or willing” what is good for others. Structural virtues are not defined by their aims. Examples are courage, self-control, and patience. These may or may not have good ends in any particular case. All by themselves, then, they don’t make a person good. This is certainly true of deference. Responsible deference, on the other hand, deference for a good end, just like courage, strength of will, and patience, are good. As Adams notes, without such virtues one cannot be “excellently for the good”. A disposition to defer in standard ‘coping’ situations seems to be a structural virtue. One’s ends may or may not be good; but the deference is necessary to really do good in these situations, where the agent is motivated to act well.
To test this let’s look at some of the other structural virtues. Consider self-control. If one has a view of virtue in which it consists of having one’s desires in conformity with what is right, as many neo-Aristotelians do, then self-control is not a virtue either since one only needs it when one is working towards virtue. It is a kind of crutch for those who lack virtue – but even if we accept this picture of virtue, which is highly contentious – this view of how self-control functions as a crutch is too one-dimensional. I can learn virtue through the exercise of self-control – this could be part of the natural developmental picture for virtue. Rather than undermining virtue, it makes it easier for people to become virtuous. After all, what really is the alternative? And the same can be said for responsible deference. Even if one accepted the neo-Aristotelian view of virtue, responsible deference is a developmental aid as plausibly as it is an undermining crutch.9

But responsible deference doesn’t simply help someone develop out of limitations. We will not be able to avoid all limitations, and, we will need deference as part of our decision-making tool kit, a quality that if responsibly deployed is part of an excellent human character. Thus, ‘the virtue of deference’ is not an oxymoron.

Bibliography


---

9 Wiland (2014) notes that trusting in an advisor is one way to acquire moral understanding.
A Plea for Moral Deference

Gopal Sreenivasan
Crown Professor of Ethics
Duke University
Philosophy, Arts & Sciences
gopal.sreenivasan@duke.edu

ABSTRACT
It seems to be a commonplace of the philosophical literature that there is no such thing as moral expertise. Or perhaps, more narrowly, that there is no such thing as justified deference to moral expertise, when there is moral expertise. On the other hand, a warrant for moral deference seems to have a secure place in everyday moral experience. It is illustrated, for example, by the ubiquitous phenomenon of taking moral advice (this includes a role for exemplars of individual moral virtues, but is not limited to exemplars of virtue). In this paper, I shall defend moral deference against overblown philosophical skepticism. I hope to contribute to rehabilitating the notion for some role in moral theory.

KEYWORDS
Moral deference, moral expertise, moral virtue, ethics, virtue ethics.

Imagine that you face a practical situation in which you do not know what the morally right thing to do is. Suppose, moreover, that you also happen to know someone who does know what the right thing for you to do is (and she is in a position to advise you). Should you accept her moral advice? In due course we shall encounter and discuss some examples that fit this bill. It will also be useful to refine the parameters of our question in a little more detail along the way. But we can make a reasonable start in fairly general terms.

In the recent philosophical literature around this question, two facts stand out as fixed points of the discussion, with one of them having greater salience than the other. The first fact is that the question has some importance for moral theory, especially perhaps for theories of virtue in particular. The more salient fact is that, sociologically, there is a wide consensus among participating philosophers that, in some important sense, the correct answer to the question is ‘no, you should not accept the other person’s moral advice.’ Naturally, my two qualifications call for some elaboration.

A narrower way of construing the relevant consensus would be to express the point of agreement as the proposition that accepting the moral advice is somehow
notably objectionable: it carries some kind of toxic stain.\footnote{For example, it is ‘unacceptable’ (Hopkins 2007; Hills 2009), ‘off-putting’ (McGrath 2011), or ‘fishy’ (Enoch 2014).} While some parties to this agreement hold that it is nevertheless sometimes permissible to accept the advice—that, under some conditions, the objection may be overcome—all parties agree that there is something to be overcome. Indeed, a common project in this literature, undertaken even by ‘defenders’ of moral deference or testimony,\footnote{See, e.g., Sliwa (2012) and Enoch (2014).} is to explain the stain that attaches to accepting moral advice, i.e. to identify or diagnose it properly. The presence of the stain accounts for the ‘important sense’ in which you should \textit{not} accept the advice, even if, all things considered, accepting it remains permissible.

My own view is that no stain attaches to your accepting the advice of someone who knows (when you do not) what the right thing for you to do is. You may accept such advice, and often you should accept it. More significantly, in the basic and most instructive version of the case, there is no good objection to your accepting it. In arguing for this conclusion, I not only decline to join the recent philosophical consensus, but I reject it. (Of course, I do not deny that the consensus itself exists—hence, ‘sociologically’). Unlike many of its other defenders, then, the plea I shall be making for moral deference will be a plea without excuses.

Now different contributors to this debate enter it along different terminological pathways. The relevant terrain is defined by a nexus of inter-relations among the terms, ‘moral expertise,’ ‘moral testimony,’ and ‘moral deference,’ where (roughly) non-experts defer to experts in relation to their testimony. It is possible to distinguish sharply between any pair of these terms, thereby severing one term from the nexus. The motivation for so doing is usually to establish its innocence by disassociation. But, whatever their motivation, contributors often leave at least one of these terms outside the scope of their enquiry altogether. I shall do the same.

In what follows, I pay no particular attention to moral ‘testimony.’ Insofar as there is any reason for this, it is because I shall also be ignoring a subsidiary debate that commonly arises here concerning an alleged asymmetry between the moral and non-moral cases.\footnote{See, e.g., Driver (2006), McGrath (2011), and Howell (2014).} Everyone (or almost everyone) admits that there is nothing objectionable about deferring to non-moral testimony. Given the background consensus that something is wrong with deferring to moral testimony, a further question therefore arises of how to explain the resultant asymmetry. Among other things, framing this asymmetry in terms of ‘testimony’ facilitates comparisons with an established epistemological literature on (non-moral) testimony. Clearly, I reject
the presupposition of the comparison. To some extent, however, I am also simply opting for a narrower scope.

While I ultimately wish to focus on evaluations of moral deference, I shall begin with moral expertise. For the most basic objection to moral deference is that there is simply no such thing as moral expertise, and hence nobody to whom one might defer morally. Even though I believe that this objection is both mistaken and confused, it remains well worth discussing. I shall then introduce an example and develop my defence of moral deference in relation to it. Throughout I shall argue in terms of morality quite generally. Yet I am also interested in the ramifications for deference to exemplars of the moral virtues, specifically. Accordingly, I shall close with a coda on virtue. In the coda, I shall argue that, in several respects, the case for deference within the province of virtue is easier to make than the generic case for moral deference.

1. Moral expertise

Let us begin by considering an ‘in principle’ version of the objection that there are no moral experts. So construed, the objection lends itself to either a negative or a positive formulation. In its negative formulation, the objection derives from a denial that there can be any in principle moral epistemic elite—a denial, in other words, that there is any sub-class of individuals who know (or are even able to know) moral truths that the rest of us cannot know. In its positive formulation, it derives from an affirmation that, in principle, moral truths are all fully accessible to everyone, i.e. to every ordinary person. Although the historical origins of this view are commonly—in the philosophical literature, anyhow—attributed to Kant, they really belong to a cultural legacy to which Kant himself is much more heir than testator, namely, the legacy of the Protestant Reformation. Fundamentally, this version of the objection originated in an anti-clerical critique: specifically, in the idea that there are no human gate-keepers to salvation.4

In any case, as a matter of substance rather than history, the basic idea inspiring the objection is plainly very appealing, since it amounts to the democratization of morality itself. Equivalently, it applies something like a principle of equality of opportunity to the achievement of moral knowledge, and thereby to the achievement of a morally good life. Unfortunately, however, just because an idea is very nice—noble, even—that does not make it true. On inspection, moreover, this particular idea turns out to be philosophically suspect. Let us examine the matter under its positive formulation. What might account for the fact, if it is one, that the truths of morality are all fully accessible to every actual human being?

4 See, e.g., the helpful account in Taylor (1989, ch. 13).
At least when human epistemic capacities are taken as a given, the explanation for the universal accessibility of moral truths arguably has to rest either on a giant coincidence or on the operation of some prior constraint on the moral truths themselves. That is to say, in the latter case, that this accessibility results from morality’s having been ‘bent to fit’ our epistemic capacities. Which particular epistemic capacities served as the target for this bending depends on whether every actual human being is equally capable (in principle) of working out the requirements of morality. In the more difficult—but presumably, much more plausible—case in which actual human beings (even ‘normal’ ones) are not equally capable in this respect, the prior constraint would have the effect of bending morality to fit the lowest common denominator among (normal) human epistemic capacities.

For simplicity, let me reject the giant coincidence out of hand. What remains is the possibility of explaining the universal accessibility of moral truths on the ‘bent to fit’ model. But this explanation seems to commit us to some kind of constructivism about morality. As far as secular accounts of morality are concerned, then, the explanatory basis of the present objection to moral expertise appears to be inconsistent with the objectivity of morality. Whether or not that counts as a

---

5 How demanding a coincidence is required depends, inter alia, on whether the content of morality is fixed in advance of the comparison or not. On one model—probably more appropriate to the historical view, but much less philosophically robust—we assume a particular content for morality (conventional Protestantism, say) and then compare the epistemic capacities required to grasp that content to our actual (lowest common denominator) capacities. The claim on offer here is that the terms of this comparison happen to coincide. On another model, moral truths are held to be universally accessible, whatever the content morality turns out (correctly) to have. While this is philosophically more robust, the coincidence it requires is also correspondingly more demanding: for here the claim is that, of all the manifold contents morality might have, those contents that exceed the reach of our actual (lowest common denominator) epistemic capacities just happen, all of them, to be false. Of course, the staggering extent of this coincidence can always be avoided by falling back on the first, less demanding model. But the consequent reduction in coincidence will be encumbered by the presupposition that the particular content nominated for morality is both correct and complete. Either way, the position seems unsatisfactory.

6 On theological premises, by contrast, it is still possible to reconcile a version of constructivism about morality with its objectivity. Moreover, if the bending is to be God’s work, it seems that the desired congruence between our actual epistemic capacities and the capacities required to grasp morality could in principle be achieved by bending in either direction. In other words, a further option of bending our capacities (to bring morality within their reach) then comes into view. While a secular account can, of course, also allow human epistemic capacities to change, I am assuming that it has no way to make sense of the idea of bending them to fit morality.
decisive strike against the objection, it is certainly philosophically problematic. In any case, that is my ground for classifying the objection as a mistake.

I need not insist on this point, however, because the objection is also confused; and its confusion alone is adequate for my purposes. To see this, let us stipulate that every normal adult human being is equally capable (in principle) of working out the requirements of morality. Moral truths are therefore all fully accessible in principle to every ordinary person. We can regiment this landscape by saying that there are no moral ‘experts’ (in it). All the same, it simply does not follow that there is nobody to whom one might defer about what morality requires one to do.

To generate that further conclusion, even for a given point in time, we would need to be shown that every ordinary person had developed his or her in principle equal epistemic capacities equally, i.e. to the same extent as everyone else (at that time). Alternatively, differential moral learning among some group of individuals is both consistent with their underlying epistemic capacities all being equal in principle and yet inconsistent with there being nobody in the group to whom any one of them might defer on moral questions.

Those who have learned more are natural candidates for the role of someone to whom others who have learned less might defer. It remains a separate question, of course, whether there is anything objectionable about moral deference (and we shall join that question below). However, the present issue is not whether moral deference is objectionable, but rather whether it is even possible, in the minimal sense of there being anyone available to whom others might defer on moral questions. On the face of it, one person’s having some moral knowledge that another person lacks – or, perhaps better, her reliably having such knowledge – is sufficient to put the first person in the role of someone to whom the second person might defer morally. To occupy this role, the first person need not be a moral ‘expert’ in any stronger sense than that, and certainly need not be an expert in the sense of knowing some moral truths to which the second person lacks epistemic access in principle.

---

7 Here as elsewhere, we may be reminded of the familiar symmetry between *modus ponens* and *modus tollens*. McGrath (2011), e.g., drives something like this argument in the other direction, taking the (asymmetrically) ‘off-putting’ character of moral deference as a ‘datum,’ and deriving a challenge for moral realism from it.

8 We can refine the relevant conditions more fully in the next section.

9 I have no stake in the label: it makes no substantive difference whether ‘expert’ is in any sense an appropriate label for the ‘object of deference’ role. If it is, then the confusion in the objection we are discussing can be diagnosed as equivocation between weaker and stronger senses of ‘moral expert.’ If ‘expert’ is a wholly inappropriate label, then the objection’s confusion is to have (wrongly) assumed that the existence of moral experts is a condition of the possibility of moral deference.
Logically, then, the universal accessibility of moral truths (even when granted for free) is hopeless as a basis for pre-empting the possibility of moral deference. Nevertheless, it should also be acknowledged that ‘differential moral learning’ is not merely a logically coherent spanner that happens to lend itself to insertion in the anti-deference crusader’s works. On the contrary, I take it that differential moral learning is a plain fact of ordinary moral experience, and indeed a massively common one. All of us, presumably, are acquainted with people who are better than we are, or more reliably knowledgeable anyhow, in some or other department of morality (if not many such departments). There is no shame in admitting this. Nor (I’m guessing) does it require a lot of reflection, but only a little honesty or humility, to recognize oneself in that portrait.

One could always try to factor differential moral learning into its causal constituents. Besides the omnipresent possibility of differential underlying epistemic capacity (itself a complex of factors, clearly), there is certainly differential experience as well, and no doubt more. I shall not pursue this sort of analysis, since the raw fact of differential moral learning is adequate for our purposes. Still, in case some resist acknowledging the raw fact in the first place, it may be worth briefly going one more round.

Differential experience can be factored into differential net quality and differential quantity. Opening an explicit place for quantity reminds us that the ‘wise people’ who populate legend and folklore are invariably wise old people. That is to say, brute relative age—certainly, a generation gap—makes a prime contribution to differential moral learning, which is relevant here because real people are always distributed across a generational spectrum, unlike the weightless contemporaries of abstract analysis. On reflection, this is itself enough to yield our conclusion.

Consider, e.g., the platitude that it is appropriate (compulsory, really) for children to defer morally to their parents or elders. What makes a child’s moral deference to his elders possible is their differential moral learning. However, as long as this differential persists into the child’s adulthood, the upshot will precisely be differential learning between adults (the grown child and his elders still), i.e. our raw fact. To resist this conclusion, one would have to deny that this differential typically does persist for a significant interval, which requires one to embrace one of the following nettles: Either there is some early point in adult life at which moral learning effectively ceases or else the bare attainment of majority obliterates any

---

10 Its contribution is consistent, of course, with some cases of precocious moral wisdom, as well as with the age difference between two people’s declining in epistemic significance over time. I insert the qualification ‘typically’ into my claim about the persistence of differential moral learning, in the next paragraph, to cover just these points.
remaining gaps in the former child’s moral knowledge (relative to his elders). While the latter proposition is scarcely credible, the former is not at all plausible either.

2.  *Moral deference*

So let us take it that moral deference is possible and return to the question of whether it is objectionable. To focus our discussion, consider the following everyday sort of example:

*Country cousin.* Suppose that my poor cousin from the country is coming to town for a few days to interview for a job. He has asked me to put him up during his visit, since a hotel bill would be a real hardship for him. But my apartment is very small, with hardly enough room for my immediate family, which includes a baby. Since there is no spare couch, I would have to let him sleep on the floor. That would work, albeit with some discomfort for all concerned, especially him. The alternatives seem to be footing his bill at a modest hotel nearby or finding some friend with a larger place to help me help my cousin. Each of the available options—floor, footing, or friend—has its disadvantages.¹¹ I have no hesitation, let us say, in agreeing to my cousin’s request. But I also have no idea which of these ways of putting him up is best.

There are various ways in which I might consider how to respond to my cousin’s request. I might simply wonder which of the options is best. Or I might wonder, of some particular option, whether it is required (or somehow, best). Or I might wonder something yet more specific. With this case, it seems natural to wonder whether *generosity*, e.g., requires me to take the footing option. To wonder, that is, whether offering my cousin the floor or my friend’s couch would constitute a criticisable lack of generosity on my part.¹² Let me begin by following this particular thread.¹³

By hypothesis, I myself have no idea whether or not generosity requires me to take the footing option in *Country cousin.* (While I may have a nagging suspicion to that effect, I have no idea how to evaluate that suspicion). How then might I proceed, given my ignorance? One possibility, evidently, is to ask somebody who

---

¹¹ The hotel bill would be a strain for me, though I can certainly manage it better than my cousin. On the other hand, he might feel rather awkward or offended being sloughed off on someone he does not know.

¹² We need not concern ourselves with the issue of what the force of this criticism would be, exactly. We may simply suppose that I innocently aspire not to merit the criticism.

¹³ In the coda on virtue, I shall return to comment on the relation between this narrow thread and the widest question, what is the right thing for me to do (in *Country cousin*)? I shall claim, though, that it makes no real difference to the status of moral deference.
does know what generosity requires (or, more generally, someone in a better epistemic position). However, in order for this quite general possibility to be practically available to me, two further conditions have to hold.

To begin with, trivially, the situation must leave me enough latitude to seek counsel (sometimes immediate action is required). Furthermore, someone who knows what generosity requires has not merely to exist, but to be known (and available) to me. There is some question about how this second condition gets to be satisfied, but for the moment I shall simply stipulate that it holds.¹⁴ Let us say that I have a friend available, Gina, who is known to me as a model of generosity (i.e., I know that she is reliably knowledgeable about what generosity requires). For convenience, I shall refer to this version of the case, which builds in the two further conditions, as Country cousin (plus).

Now suppose that I ask my friend Gina for advice and that she says there is no particular reason to foot my cousin’s bill.¹⁵ “You could,” she says, “but it would be going over the top. Offering him the friend’s couch is perfectly good. You can offer him your own floor as well, to make it clear that you are not trying to slough him off. He can choose.” From here, we could proceed straight into an examination of whether there is anything objectionable about my accepting Gina’s moral advice. But this may trample over a distinction on which some critics of moral deference have wished to insist, namely, a distinction between accepting moral advice and moral deference proper.¹⁶

As applied to Country cousin (plus), this distinction turns, roughly, on whether I accept Gina’s judgement simply on the basis of her known reliability as a model of generosity (deference) or whether Gina’s role extends instead to bringing me somehow to grasp myself why generosity does not require me to foot my cousin’s bill (advice). Thus, to clarify that our case offers an example precisely of ‘deference,’ so understood, let us specify that I accept Gina’s judgement here simply on the basis of

---

¹⁴ This stipulation is dialectically legitimate because (following on the argument in the previous section) the controversy has now shifted, in effect, to the question of whether moral deference is objectionable, given that it is possible. My stipulation merely serves to tighten the scope of the italicized concession. Nevertheless, a profitable question does remain to be discussed here, and I shall return to it (too) in the coda on virtue.

¹⁵ To forestall irrelevant difficulties that arise when the objective and subjective dimensions of moral evaluation come apart, let me also stipulate that Gina is right about this.

¹⁶ See, e.g., Hills (2009) and McGrath (2011). For criticism of this distinction, see Sliwa (2012). I am very sympathetic to Sliwa’s criticism, as I am to her argument generally. But it is not necessary to fight that battle here.
A plea for moral deference

her reliability.\(^1\) While her judgment is indeed correct, it remains obscure to me why. I shall refer to this final version of the case as *Country cousin* (*). So I defer to Gina in the matter of what generosity requires. What is wrong with that? In general terms, a very common idea is that something morally important is lost when we defer to others. It somehow belongs to the ideal of an adult moral agent, we might say, that she works the answers to moral questions out for herself. Living up to this ideal therefore requires more of me than mere acquisition of a valid warrant to affirm some answer to my moral question [even, the correct answer], which I presumably do acquire in *Country cousin* (*). Taking a page from Hills (2009), a central example of what I lose (or rather, fail to gain) is moral understanding of why generosity permits me to take the friend or floor options.

Somewhat surprisingly perhaps, I fully agree that, in some sense, it is better not to defer morally to others. Alternatively, I agree that some valuable things are lost or forgone when we defer, and even that these plausibly include moral understanding. However, I deny that this entails any good objection to moral deference. The basic point is a structural one. But we should warm up to it by noticing that the relevant structure can be framed without appealing to any especially moral values, let alone fancy theoretical ones.

Situations in which one moral agent defers to another are situations in which the first agent falls short of our ideal of moral agency. The valuable things that stand to be lost here can be analyzed as dimensions of this ideal, i.e. as respects in which someone’s moral agency may prove sub-optimal. Moral understanding is one such dimension; and other contributors to the ‘explain the stain’ game, if I may call it that, propose various additional candidates. Yet the mundane advantages of self-reliance also fit the minimal bill perfectly well. That is to say, moral agents who can answer moral questions for themselves will have access to these answers even in situations where no other reliable advisor is available, whereas the same cannot be said of agents who have to rely on others to answer moral questions for them. This suffices to yield a fairly straightforward sense in which deference is sub-optimal or in which it is better not to have to defer.

However, from the mere fact that moral deference is sub-optimal, it simply does not follow that deference by some agent is at all objectionable or somehow stained or to be avoided. This basic point is demonstrable without regard to the particular

---

\(^1\) We nestle up against a separate issue here, concerning how articulate (intellectually or philosophically) a reliable moral judge has to be. I deny that Gina’s reliability in judgment entails that her ability to articulate the reasons for her judgments about generosity is equally reliable. But we need not engage that dispute here. Those who disagree may imagine either that (while she was capable of explaining her reasons to me) Gina was otherwise prevented from doing so or else that she did explain her reasons, but I still failed to grasp them.
dimension in which the relevant agent’s moral agency proves to be sub-optimal. It will nevertheless be dialectically useful to frame the argument specifically in relation to moral understanding (rather than self-reliance, e.g.). To sharpen the point, let us also assume that moral understanding is much more important than self-reliance (as Hills 2009 seems to hold) or than any of the other candidate dimensions.

So consider moral understanding. Its ability to license objections to moral deference entirely depends on whether understanding is actually on offer in the agent’s practical situation. Suppose, e.g., that I cannot understand why generosity does not require me to take the footing option in Country cousin (*). In that case, the valuable thing that ‘stands to be lost’ by my deferring to Gina is already lost. There is nothing I can do to change that.\(^{18}\) But then no objection to my deferring can arise (not from moral understanding, anyhow).\(^{19}\)

Notice, crucially, that this conclusion does not result from any trade-off. More specifically, it is not the case that while moral understanding gives me some reason not to defer to Gina (cf. Crisp 2014, pp. 132 and 134), this reason is defeated by my reason to learn what generosity requires [which, in Country cousin (*), favours deferring to Gina]. For in the present version of the case, my practical options are only to defer to Gina or to guess at the right answer or to do nothing.\(^{20}\) None of these options will gain me any moral understanding. While it may be obvious that moral understanding therefore gives me no reason to guess or to defer, it is important to see that it does not give me any reason to do nothing either. In comparison to the first two options, doing nothing amounts to a ‘dog in the manger’ option, as it will certainly cost me something else of some value (and for no purpose). But moral understanding gives me no reason to play dog in the manger. Under the circumstances, then, it gives me no reasons at all.\(^{21}\) Rather, moral understanding is practically inert.

This further observation explains why it is literally correct that no objection to my deferring to Gina is licensed (by moral understanding). Moreover, this conclusion is preserved under substitutions of other candidate dimensions of our ideal of moral

---

\(^{18}\) Enoch (2014, p. 247) makes a similar observation, but does not draw the conclusion to follow in the text.

\(^{19}\) As I read her, Hills actually concedes this point (2009, pp. 122-4). But it seems to me that this gives the game away. It no longer follows, e.g., that ‘we have strong reasons neither to trust moral testimony nor to defer to moral experts’ (p. 98). Hills also argues that actions performed without moral understanding (of why they are right) fail to be morally worthy. However, this is a separate point. While I reject it (too), I do not have the space to engage this further claim properly here. For some criticism of this aspect of Hills’ position, see Jones and Schroeter (2012) and Driver (2015).

\(^{20}\) Some may think that I always have a fourth option, of trying to understand and failing. Those who do should suppose that I have done that already.

\(^{21}\) Except perhaps to regret that my situation has the features it does.
agency, as long as any such candidate (self-reliance, e.g.) is not on offer in the agent’s practical situation either. Hence, my conclusion—that no objection to my deferring to Gina is licensed—holds independently of taking ‘moral understanding’ as the particular valuable thing that stands to be lost.

Of course, in other situations, the valuable thing might actually be on offer. In other variants of *Country cousin (*)*, e.g., I might well be able to understand the basis of Gina’s moral advice. We shall come to these variants presently. But we should first recognise that the variant in which I cannot understand—or, more generally, in which the valuable thing in question is *not on offer* in the agent’s practical situation—is the basic and most instructive version of the case. For that is the version in which our evaluation of moral deference is focused starkly and exclusively on comparing deference to guessing or doing nothing, i.e. to the available alternative specifications of ‘not understanding’ (or, more generally, of ‘not the valuable thing in question’). What emerges from this comparison is the realisation that moral understanding is indifferent as between these alternatives. Its indifference is what makes it practically inert under the circumstances.

Once we see this point in the basic version of the case, however, we become better able to appreciate that it extends to other versions, too, albeit more subtly. Suppose now that I might actually manage, in *Country cousin (*)*, to understand why generosity does not require me to take the footing option. To simplify, let us say that my deliberative options are (i) to defer to Gina, (ii) to guess at the right answer, and (iii) to work that answer out for myself. While moral understanding plainly favours option (iii), it does not follow that option (iii) is in fact my best option. Whether that conclusion is warranted depends on how various considerations besides moral understanding balance out, including the likelihood of actually reaching the right answer for each deliberative alternative and the importance of reaching the right answer in the case at hand. No doubt philosophers disagree about these questions. But to sharpen the point once more, let us simply accept that, in this variant of the case, option (iii) is my best option. (In the usual course of affairs, this conclusion would have to be earned by argument. I do not mean to be conceding that it will typically be an easy conclusion to earn.)

Against that background, consider my choice between (i) deferring to Gina and (iii) working the answer out for myself. Imagine that I decide to defer to Gina. In the specified context, my decision is gratuitous and unjustified. However, even treating options (i) and (iii) as inconsistent alternatives, my decision still has two halves, the half in which I decide against (iii) and the half in which I decide in favour of (i). Their

---

22 For an argument that option (iii) can never be my best option, provided that I am more likely to reach the right answer by deferring, see Enoch (2014, pp. 248-50).
independence is secured by the existence of my option (ii) to guess at the right answer. Strictly speaking, then, the objection to my decision licensed by our case stipulations only applies to its first half, where I decide against(iii). That is what is gratuitous and unjustified.

Of course, one might respond that, under the circumstances, ‘having decided against (iii)’ is a presupposition of deciding in favour of (i). In that sense, the objection remains attached to the second half of my decision, insofar as that half inevitably carries an objectionable presupposition. While this cannot be denied, it actually helps to explain why, all the same, no objection attaches specifically to my moral deference. For the objectionable presupposition attaches equally to anything I decide here, once I have decided against (iii) working the answer out for myself. In particular, and by way of illustration, it would attach equally to my decision (ii) to guess at the right answer.23 But any objection that applies equally to moral deference and to guessing thereby fails to discriminate between them. Hence it has nothing to do with moral deference per se.

3. Coda on virtue

When I defer to Gina in Country Cousin(*), the conclusion I accept on her authority is that ‘generosity does not require me to take the footing option.’ It does not follow from this, of course, that morality does not require me to take the footing option. As described, then, Gina’s moral advice is partial or incomplete.

In my own view, this feature of the example is not any kind of problem. On the one hand, to put it both starkly and mildly, that is how things actually are. In life as we know it, that is, very few people are reliably knowledgeable about morality across the board. One sense, therefore, in which the example is not problematic is that it is realistic. On the other hand, despite the incompleteness of Gina’s advice, it remains the case that, in accepting it in the manner described, I fully exhibit the phenomenon of interest, namely, moral deference. Another sense, therefore, in which the example is not problematic is that it is still relevant to our enquiry.

Now there are various ways to bridge the gap between Gina’s incomplete moral advice and a conclusion about what, all things considered, the right thing for me to do is in Country Cousin(*). In the simplest case, morality’s all things considered verdict coincides with the verdict of generosity because no other moral considerations apply in the situation at hand. In that case, all that I require to complete Gina’s advice is

23 It is worth bearing in mind here that guessing is something an agent can do all by herself. Whatever the objection to the presupposition may be, then, it has nothing to do with the agent’s relations to others.
access to the proposition ‘no other moral considerations apply here’; as it were, access to that plank enables me to bridge the gap in question.

While the necessary plank might be procured from various sources, by far the simplest source is my own judgment (i.e., the agent’s own). For present purposes, it does not matter how reliable the agent’s own moral judgment happens to be. If mine is not very reliable, then my final conclusion about what to do in *Country Cousin(\*)* may suffer correspondingly. However, that is clearly a problem I have anyway. Our present concern is to evaluate the contribution(s) made by moral deference specifically. The fact that agents can always fall back on their own judgment to bridge any gap between the advice on offer and the conclusion they seek about the right thing to do, all things considered, means that incompleteness in the moral advice offered by an exemplar of virtue does not prevent such advice from being practically useful.

In the cases of interest to us, moreover, the exemplar’s judgment will, by hypothesis, be more reliable than the agent’s own, at least with respect to moral considerations specific to the virtue in question (e.g., to generosity). Hence deference to an exemplar’s moral advice, however incomplete, can only serve to improve the reliability of the agent’s final conclusion about what to do, even if that conclusion remains unreliable in certain other respects. Since improved reliability was always the primary reason in favour of deferring, we thus return to our original question about deference, namely, whether it also introduces countervailing moral difficulties. That is the question we have already asked and answered. But we are now in a position to confirm that, as I said, structuring our discussion around cases in which the advice and deference at issue are incomplete, as they are in *Country Cousin(\*)*, makes no difference to the status of moral deference.

Let us turn, finally, to consider my suggestion that pleading for moral deference is easier when it is confined to the province of virtue in particular, as compared to morality at large. For those who happen to be interested in virtue anyhow, this conclusion will be no more than a happy bonus. But since most of the debate about moral deference is prosecuted in more general terms, vindicating my suggestion may still add something in behalf of the defense. I shall concentrate on two main points, both of which concern respects in which it is easier for exemplars of virtue to earn the credential of ‘expert,’ i.e. someone to whom others are licensed to defer.

For present purposes, credentialing someone as an ‘expert’ is simply a matter of his or her reliability with respect to a certain class of moral judgments. As we have seen, the central basis on which one person is licensed to defer to another (with respect to a certain class of moral judgments) is that the second person is a more
reliable judge than the first (with respect to that class of judgments); and that the first person knows this. By an ‘expert,’ then, I mean no more than someone who satisfies these two conditions (cf. note 9).

My first point is that \textit{being a reliable judge} is easier with respect to the moral considerations specific to a given virtue (generosity, say, or compassion) than it is with respect to the right thing to do, all things considered. Up to a point, this is just a boring consequence of the fact that reliability is easier to achieve with respect to a narrower class of judgments than it is with respect to a wider class (especially when the wider class encompasses the narrower one). Beyond this, however, the extra width in the scope of judgments about all things considered rightness also covers at least two sorts of judgment differing in kind, and not merely in number, from judgments about the moral considerations specific to generosity, say.

On the one hand, this extra width extends to judgments about the various moral considerations specific to each of the other virtues. (This is plainly an umbrella category). On the other hand, it extends to the judgments required to resolve conflicts among any number of more specific moral considerations. Even if someone is a reliable judge of generosity, then, there is no particular reason to expect that the basis of her reliability—the particular abilities that constitute it—will already equip her to be reliable about either of the other sorts of judgment, let alone both of them. This is not at all to say that an exemplar of generosity could not develop all of the other relevant abilities. But it is undeniably easier for her not to (have to) develop them.

Of course, the distinctions invoked here between various kinds of moral judgments presuppose that someone can be an exemplar of generosity or compassion without having all of the other virtues. My distinctions presuppose, in other words, that there is no unity or reciprocity of the virtues. Since I have argued for that conclusion elsewhere (Sreenivasan 2009), that is fine with me. But it may be helpful to recast my first point in these familiar terms. Under the unity of the virtues, someone can have the virtue of generosity only if she also has all of the other virtues—either because the various relevant abilities are not actually distinct existences (unity) or because, while they are, having the abilities relevant to

\begin{footnotesize}
\begin{itemize}
\item[24] Distinctly more reliable, if you prefer.
\item[25] The position I shall describe is broadly similar to the one advocated by Jones and Schroeter (2012), although the basis on which it is secured is somewhat different.
\item[26] Naturally, there may be additional moral considerations that are neither specific to any virtue nor common among all virtues. If there are, then all things considered rightness will include a third sort of judgment likewise differing in kind (or more than two sorts, anyhow). For what it is worth, it seems clear to me that there are such considerations, i.e. that morality includes more than virtue(s). Indeed, my argument against the unity of the virtues begins from this very fact (Sreenivasan 2009). But the present point does not depend upon it.
\end{itemize}
\end{footnotesize}
A plea for moral deference

generosity, say, does not count as having a ‘virtue’ unless one also has all of the other abilities, too (reciprocity). By contrast, under disunity, someone can have the virtue of generosity all by itself. So construed, my first point is simply that it is (much) easier to have the ‘disunified virtue’ of generosity than it is to have its ‘unified’ counterpart; and likewise for any other individual virtue.

Eric Wiland (2015: 171-77) has recently argued that moral expertise is more plausibly or appropriately located at the level of ‘all things considered rightness’ than at the level of specific judgments of compassion or gratitude, say.27 On the face of it, his claim runs precisely counter to mine. Nevertheless, this appearance is misleading in one notable sense. In another sense, it is not misleading, but does correspond to an important point that Wiland has overlooked. It will be useful to clarify these matters briefly before proceeding to my second main point.

To begin with, Wiland actually agrees that judgements about all things considered rightness are much more difficult than specific judgements of compassion or gratitude.28 Indeed, that is his point of departure. For specific judgements of gratitude (say) are not only easier, he claims, but in fact ‘self-evident,’ so that only a ‘moral idiot’ would require help to make them properly. That is what is supposed to make deference to experts at this level inappropriate; and Wiland’s fundamental observation is that this ‘problem’ does not arise at the level of all things considered rightness. Nor do I think that the greater difficulty attending judgements of all things considered rightness makes claims of expertise or deference to experts at this level in any way ‘inappropriate.’ Rather, my contention—to foreshadow the lesson to follow below—is only that it makes experts about all things considered rightness very scarce (and hence, less available to help the rest of us). I also agree that some specific judgements of gratitude (or generosity or whatever) are reasonably regarded as self-evident; and consequently, that few normal adults will be in position to require advice about them (at least, in ‘our’ culture).

What Wiland overlooks, however, is that not all specific judgments of this or that virtue are self-evident.29 Hence, it is not necessary to ascend to the level of all things considered rightness to find moral terrain on which normal adults may, with perfect propriety, find themselves in need of moral advice. More specifically, within the class of judgments specific to a given virtue, we can distinguish ‘paradigmatic’ judgments

27 Wiland himself formulates the (lower) level of specific judgments in the language of Rossian ‘prima facie duties,’ rather than in the language of virtue. Yet, as he recognizes explicitly (2015: 171), nothing turns here on the details of the formulations.
28 Following Ross, he focuses on the dimension of conflicts among prima facie duties.
29 The same goes for prima facie duties. Wiland’s twin contrasts—‘prima facie duties’ versus ‘all things considered rightness’ and ‘self-evident’ versus ‘not self-evident’—therefore fail to align neatly in the way that his argument requires.
of that virtue from ‘non-paradigmatic’ judgments. *Paradigmatic* (or ‘stereotypical’) judgments of some virtue are judgments that anyone who is acquainted with that particular virtue concept can make; and they are typically also the judgments by reference to which the concept is taught. ‘Helping an old lady to cross the street,’ for example, is a paradigmatic act of kindness. Anyone who is familiar with the concept of kindness knows that it is ‘kind’ to help [suitably situated] old ladies to cross the street. Wiland is certainly right that no such person requires moral advice to make that judgement.30

By contrast, *non-paradigmatic* judgments of kindness are simply moral judgments that are plausibly specific to kindness and yet not paradigmatic of it. It may be controversial exactly where (and how) to draw the line between ‘non-paradigmatic judgments of kindness’ and moral judgments that not judgments ‘of kindness’ at all. But I take it, in any case, that paradigmatic judgments of kindness clearly do not exhaust the class of judgments specific to kindness. So it is not controversial that there are some non-paradigmatic judgments of kindness. Along with their counterparts for other individual virtues, these judgments satisfy both Wiland’s criterion that someone may stand in need of moral advice about them without qualifying himself as a ‘moral idiot’; and mine that reliability about them is easier to achieve than reliability about judgments of all things considered rightness.

Recall that an ‘expert’ about kindness, in the present sense, is someone who is both *reliable* about making judgments of kindness and *known to be reliable* on this score.31 My second main point is that this second condition is also easier to satisfy within the province of virtue. That is to say, not only is reliability in judgment itself easier to achieve, but so too is a *reputation* for reliability—or, more precisely, the accessibility of the evidence for someone’s reliability—easier to secure with respect to individual virtues. As I said earlier (see note 14), there is some question about how those of us in need of moral advice are supposed to go about identifying moral experts in the first place. Indeed, various skeptics about moral deference seem to feel that this represents some kind of insuperable stumbling block.

In general terms, the challenge is to find an evidence basis for someone’s judgmental reliability that satisfies all of the following desiderata: First, it really is

---

30 It does not follow that, when co-located with an old lady and a street, everyone familiar with the concept of kindness always knows whether to help her across or not (not even when no other moral considerations apply). For there is still a significant difference between making a moral judgment in a stylized ‘textbook’ context and making the ‘same’ judgment out in the world. See, e.g., the persuasive account in Blum (1994). Accordingly, even within the core of paradigmatic judgments for a given virtue, there remains room for differential reliability in practice between normal adults, and thus room for moral deference.

31 For simplicity, I elide the comparative dimension of expertise here.
evidence of reliability (so that trust on its basis is neither blind nor arbitrary); second, one does not have to be an ‘expert’ oneself to employ it; and third, it is not itself controversial. The argument (or rather, claim) in the literature concerning possible asymmetries between the moral and non-moral domains in relation to expertise has been that, in the moral domain, these desiderata either cannot be satisfied or are very difficult to satisfy jointly.

At least within the province of virtue, I claim that a straightforward basis for identifying moral experts is made available by the distinction between paradigmatic and non-paradigmatic judgements. Consider kindness again. My suggestion is that experts in kindness can be identified on the basis of their (greater) reliability in performing paradigmatically kind acts. These acts can be identified as acts of kindness by anyone who is familiar with the concept. Hence, access to this (potential) evidence does not depend upon one’s being an expert oneself. Moreover, the open accessibility of paradigmatically kind acts still leaves room for differentially reliable performance between normal adults (cf. note 30). The rest of us are therefore in position, more specifically, to recognize others—if or when we encounter some—who are more reliable about acting kindly than we are (i.e., others who have comparative expertise or expertise, proper). Finally, the status of paradigmatically kind acts as ‘kind’ is not controversial in any relevant sense. Accordingly, paradigmatically kind acts or judgments satisfy all three desiderata for an evidence basis on which to identify reliable judges of kindness.

The scope of the licence to defer to experts in kindness one acquires on this basis plainly includes paradigmatically kind judgments in all their manifold variety in practice. It would also be reasonable, however, to treat the scope of this licence as extending further, to non-paradigmatic judgments of kindness as well. Unlike judgments of other virtues (such as generosity or gratitude), non-paradigmatic judgments of kindness respond to the same kind of moral considerations as paradigmatic judgments of kindness. The particular abilities that constitute a reliable agent’s reliability with respect to the latter judgments are therefore, I submit, reasonably seen as enabling her to be reliable with respect to the former.

---

32 Whether there are (also) good bases for identifying experts in the wider moral domain I simply leave open. Thus, by ‘easier to identify experts in particular virtues,’ I mean that we already have a straightforward basis for doing so, as argued in the text, whereas it remains obscure what basis there might be for identifying experts within morality more generally.

33 Their status as kind is not controversial within the community of those who endorse the operative concept of kindness, which community includes both experts and those in need of moral advice. Of course, as with any moral notion, ‘kindness’ may be controversial in other, deeper senses (e.g., to moral skeptics and nihilists or between different moral cultures).
In addition, we should notice that when I defer to an expert’s non-paradigmatic judgments of kindness, what I am deferring to her about is, in effect, the question of how best to generalize the application of certain moral considerations we both recognize from central cases on which we agree to neighbouring cases where I am unclear. Thus, while it is still clearly moral deference, my deference here has a less radical flavour insofar as it is confined within the bounds of a shared evaluative perspective.

Let us take stock. My plea for moral deference has two parts. Its main part, recapitulated below, consists in explaining why there is nothing objectionable about deference, notwithstanding the consensus in the literature. This explanation holds equally in all provinces of morality. But its usefulness presupposes that those of us in need of moral advice stand a chance of actually finding someone to whom we are licensed to defer (about the matter at hand). The second part of my plea concerns the satisfaction of this presupposition; and that is the part it is easier to deliver within the province of virtue. A licence to defer to someone depends both on that person (who must be a reliable judge) and on the agent himself (who must have a secure basis on which to identify another’s reliability). In this coda, I have argued that each of these conditions is more easily established with respect to individual (disunified) virtues. Experts in these virtues are consequently more available to the rest of us than experts in all things considered rightness.

4. Conclusion

I have sought to defend moral deference against a prevalent skepticism. My main argument has been that the mere fact that the circumstances in which deference occurs are typically sub-optimal does not ground any good objection to the agent’s deference itself. Someone who defers to another person may not fully achieve our ideal of an adult moral agent. Fair enough. Still, this fact does not result from the agent’s deference nor does it prevent her deference from being a blemish-free response
to the agent’s circumstances as they actually are. Often, deference will also be her best response.*

References


* Earlier versions of this paper were presented, sometimes along with other material, at the annual meeting of the Canadian Philosophical Association in Victoria; at a philosophy colloquium at the CUNY Graduate Center; and at the Varieties of Virtue Ethics conference, held in Oxford by the Jubilee Centre for Character and Virtues. Thanks to the audiences on all three occasions for helpful discussion. For comments on material specific to this paper, I am especially grateful to Sarah Buss, Julia Driver, David Enoch, and Eric Wiland.
Two Objections to Virtue Ethics

Lorraine Besser-Jones  
Middlebury College  
Department of Philosophy  
lbesser@middlebury.edu

ABSTRACT  
This paper explores two objections to virtue ethics: the self-effacing objection, which holds that virtue ethics is problematic insofar as it presents a justification for the exercise of the virtues that cannot be appealed to as an agent’s motive for exercising them, and the self-centeredness objection, which holds that virtue ethics is egoistic and so fails to accommodate properly the sort of other-regarding concern that many take to be the distinctive aspect of a moral theory. I examine the relationship between these two objections as they apply to eudaimonistic virtue ethics. While defenders of eudaimonistic virtue ethics often appeal to self-effacement in order to deflect the self-centeredness objection, I argue that there is nothing in the structure of eudaimonistic virtue ethics that makes it problematically self-centered. Analysis of the self-centeredness objection shows that self-centeredness is problematic only on the assumption that the self is egoistic. Because eudaimonistic virtue ethics is predicated upon a non-egoistic understanding of human agency, it is not problematically self-centered. As a result, there is no reason for it to be self-effacing.

KEYWORDS  
Virtue ethics, self-centeredness, self-effacement.

Introduction

In discussions of the basic structure of virtue ethics, two objections often lurk. These are the self-effacing objection and the self-centeredness objection. While those critical of virtue ethics believe these objections pose serious problems for the enterprise of virtue ethics itself, defenders of virtue ethics tend to brush them off. They recognize their existence, as well as virtue ethics’ vulnerability to them, but seem to think the objections really are not bothersome. Something is amiss here: what critics of virtue ethics deem to be serious problems with virtue ethics, defenders of virtue ethics just aren’t that worried about.

The self-effacing objection holds that virtue ethics is problematic insofar as it presents a justification for the exercise of the virtues that cannot be appealed to as an agent’s motive for exercising them. The self-centeredness objection holds that virtue
ethics is egoistic insofar as its justification of the virtues fails to accommodate properly the sort of other-regarding concern that many take to be the distinctive aspect of a moral theory. While I think we can all appreciate the potential these objections have to create problems for a moral theory that is vulnerable to them, most defenders of virtue ethics have not tried to extinguish them. Where they address these objections, it often seems as if their goal is to show that while virtue ethics is vulnerable to them, this vulnerability does not really present the problem that the objectors think it does.\textsuperscript{1}

In this paper, I seek to develop a new, more satisfying, line of response to these objections, and one which sheds light on the basic enterprise of virtue ethics. As there are many varieties of virtue ethics, and the application of the two objections varies depending upon the kind of virtue ethics we are focusing on, my discussion will be limited to consideration of eudaimonistic virtue ethics (EVE), which many think is especially vulnerable to these objections. I take EVE to include any virtue ethical theory that motivates the virtues by appeal to their connection to the agent’s flourishing. While Aristotle’s virtue ethics and the contemporary theories that it has inspired are the most well known examples of EVE, my discussion will focus on EVE considered more generally. Because both objections target the framework of EVE, my question will be whether or not the framework of EVE necessarily is committed to features that make it self-centered and require self-effacement.

I’ll begin with consideration of the self-effacement objection. After considering why self-effacement is problematic, I’ll argue that whether or not EVE must be self-effacing turns on whether it is self-centered. I then move to consideration of the self-centeredness objection, I’ll argue that this objection misconceives the nature of the self that lies at the heart of, and indeed, drives EVE. I’ll conclude by showing that, because EVE is not vulnerable to the self-centeredness objection, it need not be self-effacing.

\textit{Self-Effacement}

A moral theory is \textit{self-effacing} if considerations that justify a particular act cannot be appealed to as motive to perform said act. In a well-known paper, Stocker (1976) charges both deontology and consequentialism with being self-effacing. According to Stocker, an agent who attempts to do the right thing for the reasons offered by either the deontologist (“because it is right/my duty/specifies the rules”) or the

\textsuperscript{1} Toner (2006) calls this move, as it is made with respect to the self-centeredness objection, the “complacency defense”.

61
consequentialist (“because it promotes the best state of affairs”) frequently fails to do the right thing, because she has incorporated the justification for her acts into her motives. The woman who visits a sick friend in the hospital because it is her duty fails to act well: she ought to be visiting her sick friend not because it is her duty, but because she cares about her friend. While Stocker’s original critique was of consequentialism and deontology, others have since pressed this charge against virtue ethics, arguing that virtue ethics is self-effacing in the same sense in which deontology and consequentialism are self-effacing.² Hurka even suggests that the sense in which virtue ethics is self-effacing is “more disturbing” than the sense in which other moral theories are self-effacing, for virtue ethics is non-contingently self-effacing (Hurka, 2001, p. 247). Whereas other theories (e.g., consequentialism, might be require self-efficacy based on contingent features of our psychologies, self-effacement, on Hurka’s understanding of EVE, is built into the very structure of EVE given EVE’s justification of the virtues: “to avoid encouraging self-indulgence, [EVE] must say that being motivated by its claims about the source of one’s reasons is in itself and necessarily objectionable” (Hurka, 2001, p. 247).

Self-effacement is seen to be objectionable for good reasons. Stocker (1976) originally argued that self-effacing moral theories were problematic insofar as they generated a schizophrenia between one’s reasons and one’s motives, thereby making impossible a state of psychological harmony. He writes:

One mark of a good life is a harmony between one's motives and one's reasons, values, justifications. Not to be moved by what one values-what one believes good, nice, right, beautiful, and so on- bespeaks a malady of the spirit. Not to value what moves one also bespeaks a malady of the spirit. Such a malady, or such maladies, can properly be called moral schizophrenia between one's motives and one's reasons (Stocker, 1976, pp. 453–454).

There is something very plausible about this line of thought: when we think about a well-functioning agent, and especially about a flourishing agent, we think about a person who knows why she acts and identifies with those reasons, and incorporates them into her motivational outlook. A theory that presents as a flourishing agent one who (necessarily) sets up a clear divide between her reasons and motives seems flawed.

More recently, Stocker (1996) and others (e.g., Pettigrove (2011) have worried about a self-effacing theory’s potential to offer normative guidance. The concern is that reflection upon her reasons for action—upon why it is important to exercise the virtues—plays an important practical component within an agent’s deliberative process. Oftentimes, part of figuring out what to do involves thinking about why we

² See Keller (2007) for an overview.
ought to do it. But, if reflection on the reasons why we ought to do something interferes with our success in exercising virtue, virtue ethics must be self-effacing and so burdened with the problems associated with self-effacement.

Let us now turn to the question of whether EVE must be self-effacing. Keller (2007) argues that any virtue ethics is subject to the self-effacing objection to the extent that its theory of right action refers to a conception of what the fully virtuous person would do, a consideration that in some instances cannot serve as an effective motive to act virtuously (where an “effective motive” is one that enables the agent to successfully exhibit the virtues). Keller argues, for example, that the woman who helps her friend because it is what a virtuous person would do fails to be fully generous. According to Keller, we must conclude that virtue ethics—in general—is self-effacing, and that the virtue ethicist must “say that what makes an act right is its being what the fully virtuous person would do, but add that having the governing motive of acting like the fully virtuous person precludes the possibility of being like the fully virtuous person—so it is often undesirable for people to take as their motives the considerations that provide reasons for action” (Keller, 2007, p. 227).

Is it fair to say that EVE is self-effacing insofar as it is committed to understanding right action in terms of what the virtuous person would do? While I think Keller is right to posit that any virtue ethics, including EVE, is committed to this understanding of right action, I worry that Keller’s formulation of the self-centeredness objection may be based in a mis-understanding of the role this theory of right action plays within EVE. In determining whether or not a moral theory is self-effacing, what counts are the reasons a theory appeals to in order to justify any particular act as right. For some moral theories, these reasons, and subsequent justification, are quite transparent. According to a simple consequentialism, the reason why any act is right is because it promotes the best state of affairs. The fact that an act promotes the best state of affairs also makes that act right and so serves as its justification. The same reason, then, both explains the rightness of the act and justifies it as right. When it comes to virtue ethics, however, things are less straightforward. The reasons that explain an act as right are not necessarily the same ones that justify an act as right.

Consider again what Keller takes to be the virtue ethicist’s justification for right actions: “the virtue ethicist says that the primary explanation of why right acts are right is that they are in accordance with the virtues, or would be performed by a fully virtuous person” (Keller, 2007, p. 224). Notice that in this quote Keller writes that the appeal to what a fully virtuous person would do, or to what is in accordance with the virtues, explains why right acts are right. Keller’s discussion presumes this explanation of rightness also serves as the justification of rightness and it is here that I think he errs.
Sometimes explanations can serve as justifications. As we’ve seen, for the simple consequentialist, that an act produces the best state of affairs both explains what the right act is and makes that act the right act. However, it is a mistake to read the virtue ethicist’s explanation of what acts are right as her justification of the rightness of the act, i.e., as an explanation of why right acts are right, and I worry that Keller may be making this mistake. For the proponent of EVE in particular (although the same probably holds for proponents of virtue ethics more generally), appeals to what a fully virtuous person would do have always been intended as an explanation of what agents ought to be doing. More specifically, the appeal to what a fully virtuous agent would do is meant to provide normative guidance to the person who is not fully virtuous, guidance which doesn’t explain why she should do the right thing, but rather, practical guidance, which explains what she should do. These are importantly different tasks. An explanation specifies what it is the person ought to be doing. Aristotle, for instance, would say that the person ought to act for the right reason, in the right manner, and at the right time. This is what the fully virtuous person should do. But it is not what justifies her actions and likewise should not be understood as providing her with justifications for acting. For Aristotle, and EVE more generally, the justification for developing and possessing the virtues lies in the virtue’s connection to flourishing. Their connection to flourishing makes them virtues and is what justifies their status as traits we ought to cultivate. While EVE holds that people ought to strive to act as the fully virtuous person, this is only because doing so enables them to develop a state of flourishing.

Because the justification of the virtues lies in their connection to flourishing, if EVE is self-effacing, it is so in a different manner than we find in Keller’s analysis. Whereas Keller gauges whether or not virtue ethics is self-effacing by whether or not a desire to do what a fully virtuous person would do can serve as an effective motive, the real challenge for EVE is whether EVE’s justification, which appeals to one’s own flourishing, can serve as an effective motive.

Hurka (2001) argues that EVE is self-effacing in precisely this sense. He argues that EVE must be self-effacing because it justifies the virtues by appeal to egoistic considerations of flourishing that are incompatible with the demonstration of genuine virtue:

---

3 Pettigrove (2011) notes that different kinds of persons need different kinds of reasons and so that justifications vary depending on the person towards which it is directed. I agree that reasons may vary depending on the person involved, but want to resist the idea that justifications vary. An appeal to what the virtuous person would do might serve as a reason for the child to do something, given the child’s lack of developed rationality to demand, expect, or appreciate more, but I do not think it would serve as a justification for the child’s actions. The justification ought to refer to why the virtuous person would do it.
A flourishing-based theory . . . says a person has reason to act rightly only or ultimately because doings so will contribute to her own flourishing. If she believes this theory and is motivated by its claims about the source of her reasons, her primary impetus for acting rightly will be a desire for her own flourishing. But this egoistic motivation is inconsistent with genuine virtue, which is not focused primarily on the self. (Hurka, 2001, p. 246)

Hurka is surely right in stipulating the inconsistency of egoistic motivation with genuine virtue. The person trying to act compassionately while driven by egoistic motives presents a classic illustration of someone whose specific motive is an ineffective one, which prevents her from succeeding in her actions and from developing genuine virtue.

In order to evaluate Hurka’s version of the self-effacing objection, however, we must first consider whether or not it is accurate to say, as Hurka does, that EVE is committed to an egoistic justification that gives rise to the egoistic motivation that proves incompatible with the exercise of genuine virtue. The self-effacing objection thus hinges on whether or not EVE is also subject to this, the self-centeredness objection. Let us now turn to consideration of this objection and then re-visit the question of whether it is self-effacing.

Self-Centeredness

A moral theory is self-centered if it takes, as its primary aim, promoting self-regarding concerns. This is problematic for those who think the function of a moral theory is to promote a concern for others, i.e., to inculcate genuine concern and care for those around us. Many think that EVE is self-centered and straightforwardly so: EVE justifies the virtues by appeal to the agent’s flourishing, thus offering, as Hurka highlights, what appears to be an egoistic justification for the virtues. While EVE maintains that part of developing the virtues is developing non-instrumental, irreducible other-regarding concerns, the worry is that the structure of EVE nonetheless is such that it inescapably places priority on self-regarding concerns (of personal flourishing). This is why Hurka thinks it must also be self-effacing.

One response defenders of EVE make against the charge of self-centeredness is to distinguish between “formal” self-centeredness and “content” self-centeredness. Annas (1993, p. 225) makes this move in her analysis of the self-centered nature of EVE. According to Annas, ancient conceptions of virtue ethics, including Aristotle’s eudaimonistic virtue ethics, are formally self-centered or egoistic insofar as they
maintain that an agent’s own good serves as her final end. But this doesn’t mean that their content is self-centered; rather, they direct an agent to develop other-regarding concerns. Where people worry about “egoism”, Annas argues, is with respect to content self-centeredness, not formal self-centeredness. Thus, she concludes, even though EVE is formally self-centered, it is not problematically so, for its formal self-centeredness does not affect the content of its normative prescriptions.

Notice, however, what has happened here: Annas’ defense against the self-centeredness objection is to say that the degree to which EVE is self-centered is not a problematic one, insofar as self-regarding concerns enter into the justification of EVE yet not the content. This is just to say that EVE is not self-centered in a problematic way because it is self-effacing—its egoistic justification does not factor into an agent’s deliberations about what she should do.

This reply to the self-centeredness objection is unsatisfying on two levels. First, as we’ve seen, it commits EVE to being self-effacing, and so subjects EVE to the problems that come with being self-effacing. Second, allowing that EVE is egoistic in its justification of the virtues overlooks the central insight of EVE, which is that human beings are not egoists. This is a point Annas hints at in a later article, where she seems to depart from her earlier position and argues “that aiming at flourishing is not egoistic” (Annas, 2008, p. 215). I agree with this basic sentiment and think it needs and ought to be fleshed out more concretely than it stands in Annas’s discussion. As I’ll now argue, the reason that aiming at flourishing is not egoistic, is that human nature is not egoistic. This, I believe, is the central insight of EVE; EVE is based upon the view that we are not egoists. Recognizing this provides a response to the self-centeredness objection that does not require self-effacement and so defends EVE, decisively, against the two objections in question. More importantly, however, it uncovers what I think is the real issue at stake in these debates. This has to do with the picture of the self to which EVE is committed, a picture that, we will see, is not, in any sense, egoistic. By bringing to light this understanding of human nature and the vision of the self that lies at its core, we can reach a better understanding of the basic framework of EVE and what distinguishes it from other normative moral theories.

To make my case, I begin with evaluation of the self-centeredness objection. I argue that self-centeredness is problematic only if the self that lies as the object of concern is construed egoistically. Because EVE construes the self in non-egoistic terms, it is not problematically self-centered, nor is it at all mysterious as to how the development of other-regarding concerns can be justified by appeal to flourishing.
A Non-Egoistic Self

While a theory will be self-centered if it grants priority to an agent’s self, I’d like now to suggest that self-centeredness is only problematic if the self prioritized by the theory is an egoistic one. Call an egoistic self one whose true interests can be described without making essential reference to the interests of others such that one person’s interest usually (although not necessarily) stand in a zero-sum relation to the other, whereby one person’s gain is the other’s loss. Call a non-egoistic self a self whose true interests can be described only through reference to the interests of others, such that one person’s interests usually (although not necessarily) stand in a positive-sum relation to the other, whereby one person’s gain adds to the gain of the other and no one wins at the expense of another.

The justification of the virtues upon which EVE rests presupposes the existence of a non-egoistic self. It presupposes that one cannot flourish unless one takes into account the needs and interests of others and understands that doing so is not sacrificial of one’s interests, for our interests are interconnected and their satisfaction dependent upon the other. This non-egoistic understanding of the self underwrites the very justification of the virtues, according to which developing and exercising virtue enables individuals to flourish; without the assumption of a non-egoistic self, this justification breaks down: virtue is not necessary for the egoist to flourish, and very well may stand in conflict with the flourishing of the egoist.

While my goal here is to illuminate the structure of EVE and not necessarily to defend the plausibility of this way of thinking about the self, it is worth taking a minute to explore some research suggesting it is both a viable and accurate way of conceiving of the self, lest we think this vision of a non-egoistic self is an ancient relic—an assumption that cannot be supported given our growing knowledge of human nature.

I’ll first consider some research on motivation that supports the thesis that we function best when we operate as non-egoistic self, thus affirming in part the connection EVE makes between the virtues and flourishing. While there is a host of psychological literature attesting to this basic idea, I’m going to focus on a line of research by Jennifer Crocker that explores the effects motivation by self-image goals (representative of the egoist self) or compassionate goals (representative of the non-egoist self) has for the agent. Her research explores interpersonal relationships and everyday goal pursuit in general, and not exclusively instances of helping behavior, but the extension is clear.

---

4 I develop a full defense of this connection in Besser-Jones (2014).
In two longitudinal studies, Crocker and Canevello (2008) explored the motivational orientation of first-year college students and their success in learning, self-regulation, well-being, and relationships. Student’s motivational orientations were assessed according to whether they were driven by compassionate goals (which reflect a non-egoist perspective) or self-image goals (which reflect the egoist perspective). The defining feature of compassionate goals is that they do not make reference to a benefit to oneself. They include:

- Be supportive of others
- Have compassion for other’s mistakes and weakness
- Avoiding doing anything that would be harmful to others
- Avoid being selfish or self-centered. (Crocker & Canevello, 2008, p. 560)

Self-image goals are specified in terms of making reference to a benefit for oneself. They include:

- Get others to recognize or acknowledge your positive qualities
- Avoid showing your weaknesses
- Avoid taking risks or making mistakes
- Convince others that you are right. (Crocker & Canevello, 2008, p. 560)

Students completed surveys on their goals and experiences both before and after their first semester, and every week in between.

What Crocker and Canevello (2008) found was a significant correlation between having compassionate goals and experiences positive affective states (feeling “clear and connected”) and between having self-image goals and experiencing negative affective states (feeling “afraid and confused”). Unsurprisingly, these negative affective states tracked higher levels of anxiety and depression (Crocker, Olivier, & Nuer, 2009). Student’s motivational orientations also tracked the progress they made in their academic and social goals. Their research showed that students driven by higher than average level of compassionate goals made higher than average progress towards their other goals; it moreover showed a correlation between a weekly increase in pursuit of compassionate goals and an increase in the weekly (non-compassionate) goal progress (Crocker et al., 2009, p. 261).

This research gives us good reason to expect that agents driven by compassionate goals will be more successful in whatever they pursue than will those driven by self-

5 In this particular study, goal progress was self-reported and not tracked by objective measures. The findings, though, are consistent with Moeller et al (2008), which identified the same correlation using objective standards of goal progress (e.g., vocabulary test scores).
image goals. And, this, in turn, supports the thesis that we operate at our best when we function as non-egoists. This is the spirit of Crocker’s own explanation of this phenomenon and I think it captures in a very basic way the picture of human nature lying at the foundation of EVE.

Crocker believes that the positive effects emerging from agents who are high in compassionate goals arise largely because in embracing compassionate goals, one transcends the self. “When people transcend the self,” she writes, “caring less about how others view them and more about the well-being of others, others are mostly likely to regard them highly and provide support, and relationship quality improves. Consequently, well-being improves” (Crocker, 2011, p. 142). Her thesis is that we improve our own well-being by improving the well-being of others, and that this happens most effectively when we are driven by compassionate goals, taking on a position she describes as an “ecosystem motivational perspective” (Crocker et al., 2009; Crocker, 2011). Individuals working from this motivational perspective take on, in Crocker’s words,

a perspective in which the self is part of a larger whole, a system of separate individuals whose actions nonetheless have consequences for others, with repercussions for the entire system, that ultimately affect the ability of everyone to satisfy their fundamental needs. Like a camera lens aimed at the self but zoomed out, people with an ecosystem motivational perspective see themselves and their own needs and desires as part of a larger system of interconnected people (and other living things), who also have needs and desires. (Crocker et al., 2009, p. 254)

What Crocker calls the “ecosystem perspective” is analogous to what I’ve been calling the non-egoist self; the individual working from this perspective does not sharply distinguish between her interests and others and she does not calculate the personal costs of helping others; she just does it, and benefits as a result.

In contrast, individuals driven by self-image goals work from an egosystem motivational perspective, which is analogous to what I’ve described as the egoist self:

Like a camera lens zooming in on the self, they focus on themselves and their own needs and desires. They view the relationship between the self and others as competitive or zero-sum—one person’s gain is another’s loss. They evaluate and judge people, including themselves, and they expect evaluation and judgment from others. They are concerned with the impressions others hold of them, leading to self-consciousness and social anxiety. They focus on proving

---

6 This, incidently, affirms the fundamental message of the paradox of egoism: living life as an egoist is self-defeating.
themselves, demonstrating their desired qualities, validating their worth, and establishing their deservingness. (Crocker et al., 2009, p. 252)

Key features of this egosystem motivational perspective are that the self is always in perspective and is invoked as the standard from which the agent evaluates her options as well as how she negotiates with others.

Crocker’s analysis puts into concrete form the two different ways of understanding basic human agency that lie at the root of much of the debate between defenders of EVE and those who think that EVE is self-centered. Defenders of EVE believe that in order to flourish, agents must transcend the self and operate from an ecosystem perspective, as non-egoists. Those who think EVE is problematically self-centered may contest that the only reason agents have to adopt this ecosystem perspective is that it makes sense to do so from an egosystem perspective, i.e., they will say that the reason why we have to transcend the self is because it benefits the self, conceived as egoistic. But this misses the point. We don’t adopt an ecosystem perspective because it makes sense to do so from an egosystem perspective. Rather, the ecosystem perspective enables us to flourish precisely because, at our core, we are not egoists. The best explanation of why we operate at our best when we operate as non-egoists is because we are, at our core, non-egoists.

A second range of research affirming this position draws on the deeply rooted needs we have for engaging with others in meaningful ways, research which we can see as both affirming and explaining why we operate best as non-egoistic selves. It has long been acknowledged that there is within human nature a need for relatedness. The need for relatedness shows itself earliest in the form of attachments between parents and infants. Infants need to develop attachments to an adult that make them feel safe and secure; this allows them to begin exploring new territory, all the while confident that they have a secure base to return to and to support them. The need for attachment transforms as we mature, but never disappears. We need to feel connected to others, to feel a sense of belongingness. Importantly, what we need as adults seems to be to develop interactions that exhibit mutual care and respect for both parties: it is not enough for others to be cared for; we need also to care for others—one-sided relationships do not typically fulfill our need for relatedness regardless of which side one is in. That individuals are driven by this need to engage well with others (and that their well-being is diminished when this need goes unsatisfied) affirms EVE’s assumption that we are non-egoists, and that we flourish

---

7 Many different research perspectives affirm this need, ranging from attachment theory (Bowlby, 1979), to evolutionary theory (Fowers, in draft), to research on motivation (Deci & Ryan, 2000) and self-esteem (Leary, Tambor, Terdal, & Downs, 1995).

8 I discuss the above research at length in Besser-Jones (2014).
when we recognize that we are essentially connected to others and that the interests of others do not stand in a zero-sum relation to one’s own interests.

Revisiting the objections

Recognizing that EVE is predicated upon a non-egoistic conception of the self allows us to put into perspective the concerns regarding both self-centeredness and self-effacingness. We can now see that self-centeredness per se is not a problematic feature of a moral theory; rather, it is only problematic insofar as it works in conjunction with an egoistic conception of the self. A moral theory committed to an egoistic conception of the self ought not to be self-centered. But EVE is not at all committed to an egoistic conception of the self. Its central insight, as we’ve seen, has always been that developing and exercising virtue is part of flourishing, that an agent must act well towards others in order to flourish herself, and that this is so because we are not egoists. The self-centeredness of EVE is thus not problematic.

I have argued that the self-centeredness intrinsic to the structure of EVE is not problematic because it works in conjunction with a non-egoistic conception of the self. Recognizing this allows us to respond to the self-centeredness objection in a way that helps us to better understand the enterprise of EVE itself. That EVE assumes a non-egoistic self is one of the things that set it apart from many other moral theories. EVE recognizes that the answer to the question “how ought I to live?” is to develop the virtues. It recognizes that this will involve caring about others as well as oneself and it recognizes that this will enable an agent to flourish. It does all of this because it recognizes that the self is fundamentally non-egoistic.

We are beings who are intertwined with others and for whom treating others well—exercising virtue—allows us to cultivate a state of flourishing. This is the fundamental insight of EVE and one that makes perfect sense when considered in conjunction with the picture of human agency revealed above. It is also one, I think, that ought to be recognized and reflected upon by the virtuous agent; this, I’ll now argue, both precludes and makes unnecessary self-effacement.

Self-effacement occurs when one cannot appeal to the justification of an act as also a motive to act. We’ve seen that Hurka believes EVE requires self-effacement because it offers an egoistic justification of the virtues that cannot be embraced as a motive, for egoistic motivation is incompatible with the development and exercise of genuine virtue. However, once we recognize that EVE offers a justification grounded in and dependent upon the thesis that we are not egoists, this picture changes significantly. EVE does indeed justify the virtues by appeal to flourishing, but the reason that this succeeds as a justification is because we are not egoists, and cannot flourish without caring about others and developing an irreducible concern for their
own well-being. Given this justification of the virtues, there is no need for self-effacement. Recognizing our interdependency, and the need we have for relatedness, can indeed serve as an important part of developing and embracing virtue; certainly, there is nothing inconsistent in reflecting on this justification and exercising virtue. Indeed, I’ve argued elsewhere that this kind of reflection helps individuals to identify and internalize the importance of acting well to others and so in itself plays an important motivational role (Besser-Jones, 2014).  

Conclusion

Recognizing that EVE assumes the existence of a non-egoistic self allows us to see that EVE is not problematically self-centered, and that EVE does not have to be self-effacing in order to avoid the self-centeredness objection. I’ve argued that thinking about how the virtues are justified—about the connection between oneself and others and of our mutual dependency—and being motivated by those thoughts, is perfectly compatible with exercising virtue. EVE need not be self-effacing, because it is not problematically self-centered.

References


9 Here I depart from Annas (2008), who, following Aristotle, argues that the mature virtuous person will not need to engage in thoughts about virtue, and so reflection on their justification, when she is exercising virtue. While Annas maintains those learning virtue need to engage in reflection about virtue, such thoughts about virtue simply drop out of the picture once the agent begins to master virtue, I think that, given the motivational power such reflection promises, and given the consistency of such reflection with the exercise of virtue, it is important that even the mature virtuous person engage in active reflection on virtue, as she is exercising virtue.


Lists of the Virtues

Sophie-Grace Chappell
The Open University
Philosophy
sophie-grace.chappell

ABSTRACT
Virtuous action is action according to the virtues. But which are the virtues? What might be our basis for a list of virtues? In this paper I consider some of the possible answers that have been offered, reviewing material from Plato and Aristotle and the New Testament, also from Philippa Foot, Rosalind Hursthouse, Alasdair MacIntyre, and Linda Zagzebski.

KEYWORDS
Virtue, Plato, Aristotle, naturalism, eudaimonism, exemplarism, Philippa Foot, Rosalind Hursthouse, Alasdair MacIntyre, Linda Zagzebski.

I
Virtue ethics tells us to act “in accordance with virtue, by which I mean contrary to no virtue.”

But which are the virtues? Broadly, we might agree with Aristotle Nicomachean Ethics 1106b36that nothing can be a moral virtue unless it is “a disposition of choice”, a character-trait that works to transmit our values into our actions; but consensus looks harder to come by when we try to get beyond this rather simple and obvious necessary condition. Or perhaps we will not even get this far: Hume notoriously defines a virtue as any trait the disinterested contemplation of which produces in us “the pleasing sentiment of approbation”. As is well-known, this seems to capture a much wider class of traits than Aristotle’s necessary condition does.

Beyond these points, one obvious problem for virtue ethics is the relativist worry where the virtue ethicist gets her list of virtues from: “what historical enquiry discloses is the situatedness of all enquiry, the extent to which what are taken to be the standards of truth and of rational justification in the contexts of

2 “[M]orality is determined by sentiment. It defines virtue to be whatever mental action or quality gives to a spectator the pleasing sentiment of approbation; and vice the contrary.” (Enquiry concerning the Principles of Morals, Appendix 1, p. 10) Query: is it morality that does this defining? Or sentiment? Probably the latter, but Hume’s “it” is ambiguous.
Lists of the virtues

practice vary from one time to another”. There is research that gives quantitative-analysis evidence of changing language about which are the virtues, and which ones matter most, even within one society, the US, during the twentieth century:

A study by Pelin Kesebir and Selin Kesebir found that general moral terms like “virtue,” “decency” and “conscience” were used less frequently over the course of the 20th century. Words associated with moral excellence, like “honesty,” “patience” and “compassion” were used much less frequently. The Kesebirs identified 50 words associated with moral virtue and found that 74 percent were used less frequently as the century progressed. Certain types of virtues were especially hard hit. Usage of courage words like “bravery” and “fortitude” fell by 66 percent. Usage of gratitude words like “thankfulness” and “appreciation” dropped by 49 percent. Usage of humility words like “modesty” and “humbleness” dropped by 52 percent. Usage of compassion words like “kindness” and “helpfulness” dropped by 56 percent. Meanwhile, usage of words associated with the ability to deliver, like “discipline” and “dependability” rose over the century, as did the usage of words associated with fairness. The Kesebirs point out that these sorts of virtues are most relevant to economic production and exchange.

Of course there is a sense in which this kind of finding is not news. Different societies have always had different lists of virtues. Classical Greece had the four cardinal virtues justice, temperance, courage, and wisdom - and in earlier lists, holiness/ piety as well. Christian Rome, and its successor civilization, has the three theological virtues faith, hope, and charity. (Alasdair MacIntyre famously wrote that “Aristotle would certainly not have admired Jesus Christ and he would have been horrified by St Paul”: AV p.184.) Some say that the Jewish tradition today recognizes six virtues: justice, truth, peace, loving-kindness, compassion, self-respect. Elsewhere in world history, Confucianism recognizes humanity, propriety, beneficence, reverence, practical wisdom, selflessness, and exemplariness. Buddhism sometimes gives us a list of three virtues (detachment, mindfulness, pity), sometimes a list of ten (good habituation, study, keeping good company, teach ability, helpfulness, truthfulness, industry, contentment,

---

3 A. MacIntyre, appealing to Robin Collingwood in the Prologue to After Virtue, 3rd edition, University of Notre Dame Press, 2007.
5 http://en.wikipedia.org/wiki/Jewish_ethics
6 http://philosophy.lander.edu/oriental/main.html
7 From personal conversations with Buddhist friends.
mindfulness, and insight). Islam, apparently, admits no fewer than thirty-six virtues. Back in the Western tradition, besides the three “theological virtues” of 1 Corinthians 13, St Paul’s epistles are awash with other lists of desirable characteristics for Christians to display: “the fruit of the Spirit is love, joy, peace, patience, kindness, goodness, faithfulness, humility, and self-control” (Galatians 5.22-23). (More on these lists later.) Much of the Secunda Pars of Aquinas’ *Summa Theologiae* is dedicated to making out and defending one long and intricate list of virtues and sub-virtues, graces and gifts and fruits, centered on, but not confined to, Aquinas’ fusion of the Classical Greek and Roman Christian lists into one list of seven virtues. Hymns, too, often contain catalogues of desired virtues, even if in some cases what the catalogue is apt to prompt today is the retort “Not desired by me”:

Let holy charity
Mine outward vesture be,
And lowliness become mine inner clothing;
True lowliness of heart,
Which takes the humbler part,
And o’er its own shortcomings weeps with loathing.
(Bianco da Siena, “Come down, O love divine”)

A list of virtues that looks very different from either Bianco da Siena’s or Aquinas’ is derivable from what is anyway an interesting exercise in street-level experimental philosophy - a survey of the Lonely Hearts columns. In this important contemporary list, alongside the perennial favourite GSOH, the most important virtues probably turn out to be kindness, romanticness, liking pets, and being a non-smoker.

There is this diversity even within the Western tradition which all of us here and now inhabit and from which most of us are culturally descended; never mind the further diversity outside that tradition to which we are nowadays equally exposed. So it looks hopeless for any virtue ethicist to simply accept a list of virtues wholesale from her society or tradition, and just construct a virtue ethics uncritically upon that unquestioned basis. This makes it all the more striking that the *Nicomachean Ethics* might seem to do exactly that. In NE 2.7, from 1106a34 onwards Aristotle just introduces one virtue after another for discussion, “taking them” (as he disarmingly says there) “from the diagram”. (We are presumably to imagine that he has a blackboard or the like next to him as he speaks these words.)

---

8 [http://www.buddhapadipa.org/meditation/10-virtues/](http://www.buddhapadipa.org/meditation/10-virtues/)
Aristotle’s Book 2 list of virtues includes only three of Plato’s four cardinal virtues, and adds in seven other virtues as well. It runs: *andreia, sophrosynê, eleutheria, megaloprepeia, megalopsychia, praotês/ philia, alêtheia, eutrapêlia, nèmesis, dikaiosynê*. (Wisdom, both theoretical and practical, he leaves to Book 6.) Aristotle makes no attempt to say (explicitly) why these and just these are the virtues that ought to be in his diagram. The list of virtues that he goes through in NE 2.7, where his main concern is clearly to establish the doctrine of the mean, he goes through again, in greater depth and generality, in NE 3.6-5.11. (But with two small modifications: (a) second time around he either forgets or deliberately leaves out *nèmesis*; (b) at 1126a20 he oddly says that the virtue that comes after *megalopsychia* in his list is nameless, though it seems to be the same as what the first list called *praotês* (1108a7) or *philia* (1108a28).) Second time around just as first time, Aristotle says nothing to justify his list of virtues: neither when he begins it at 1115a4, nor at any later point.

At the very least there is a sharp contrast here with Plato, whose entire career as an ethicist is devoted to the question which the virtues are, and why. If Aristotle seems to incline towards an uncritical traditionalism, Plato inclines, on the contrary, towards a hypercritical rationalism. For him, above all in the *Republic*, it is of the utmost importance to be able to give a complete theoretical justification of his list of the virtues: especially justice, of course, but the other cardinal virtues too, since justice cannot be defined with full clarity except relative to them. Famously, he sees the virtues as emerging one by one from his city-soul analogy, though perhaps not in the way that we might antecedently have expected. Courage, *andreia*, is the distinctive virtue of the warrior class in Callipolis, as it is of that part of the individual psyche that Plato calls the *thumos*; and wisdom, *sophia*, presumably is the distinctive virtue of the ruling guardians, as it is of the *nous* or intellect in the individual. But *sophrosynê*, temperance, is not the distinctive virtue of the lowest order of Callipolis, the wealth creators or business class, even though they are paralleled with the individual’s *epithumiai*, base desires. Of course that class, as much as any other, needs to have the virtues; but it has no distinctive virtue. Temperance is not its distinctive virtue; rather, temperance is an agreement in all parts of the city, or soul, about which part should rule, and thus comes to sound uncomfortably close to justice as Plato defines it (*Republic* 443b), which is the condition of the city, or soul, when each part within it knowingly and willingly performs its own proper function and no other part’s.

But perhaps, on second thoughts, something like Plato’s schema for generating the virtues is still present, albeit not explicitly spelled out, in the *Nicomachean Ethics*? There, it could be said, the cardinal virtues can be imagined to emerge in orderly sequence, in parallel with Aristotle’s review of human nature “upwards” from its lowest to its highest parts, in NE 1.13. As this picture has it, temperance regulates our desire for pleasure, and courage regulates our fear of pain; then
justice is there to give right order to our relations to our fellow-citizens. As for
wisdom, which had been the fourth and highest of Plato’s virtues, this maintains
its preeminence in Aristotle’s thought, but with a curious duplication in its nature
(which is also presaged in numerous places in Plato, beginning with *Meno* 97a-c’s
famous admission that true belief can be as good a guide as knowledge to human
affairs): Aristotle recognizes both practical wisdom, *phronēsis*, as a master-virtue
for human affairs, and also theoretical wisdom, *sophia*, as a virtue that takes us
beyond the human to the divine.

If this is, at least implicitly, the schema whereby Aristotle generates his list of
the virtues - or at any rate four of them - we might have almost as many doubts
about it as about Plato’s. The schema of the *Republic* is manifestly a creaky,
clunky, and contrived way to generate a list of virtues. The psychological schema
that I have just suggested might be attributed to Aristotle is a vast improvement
on Plato’s schema, but it is still extremely rough and ready. Of course, that might
actually be an advantage: given the extremely crude psychological science that
was available to him, Aristotle could hardly have done better by basing his schema
of the virtues on a preciser psychology.

Nonetheless, modern-day Aristotelian virtue ethicists do not typically go in
this psychological direction if they want to give a foundation for a list of the
virtues. They look instead to the notion of *eudaimonia*. Aristotle, as we have just
seen, simply presents us with a list of virtues; he never says explicitly that he is
going to generate a list of virtues from the contours of human psychology. No
more does he ever say explicitly that he is going to generate a list of virtues by
asking “What are the character-traits that humans need in order to live flourishing
lives?” That has not deterred a host of attentive and intelligent readers, with
Philippa Foot and Rosalind Hursthouse at the forefront of the host, from
concluding that this is what he is really up to in his ethics.

One attraction of this conclusion - we are told - is that it makes Aristotle into
what, in modern terms, is called a “naturalistic ethicist”; another, connected,
attraction is that it seems to make his ethics dovetail very neatly with biological
science. Now I am pretty sure that Aristotle would not only have found the
concept of “naturalistic ethics” unintelligible himself - he would have insisted too
that we don’t really understand what we mean by it either. This possibility has
not deterred Philippa Foot and her followers. One crucial advantage they claim is
that, on their reading, ethics can be given a descriptive or factual grounding. What
makes humans flourish or fail to flourish is, Foot liked to point out, a matter of
biological or zoological fact, just as it is a matter of biological fact what makes a
plant or a tree flourish. What the traits are that lead to this flourishing are also,
therefore, at least in principle factually establishable. Hence there is such a thing

11 For more on this see my “Aristotle’s naturalism”, in *A Companion to Greek and Roman Political
as the unique objectively correct list of the virtues, and the contents of that list can, at least partly, be read off the nature of human beings as a zoological species in nature.

As some species of animals need a lookout, or as herds of elephants need an old she-elephant to lead them to a watering-hole, so human societies need leaders, explorers, and artists. Failure to perform a special role can here be a defect in a man or a woman who is not ready to contribute what he or she alone - or best - can give. There is also something wrong with the rest of us if we do not support those of genius, or even of very special talent, in their work.

In spite of the diversity of human goods - the elements that can make up good human lives - it is therefore possible that the concept of a good human life plays the same part in determining goodness of human characteristics and operations that the concept of flourishing plays in the determination of goodness in plants and animals. So far the conceptual structure seems to be intact. Nor is there any reason to think that it could not be in place even in the evaluations that are nowadays spoken of as the special domain of morality…

Men and women need to be industrious and tenacious of purpose not only so as to be able to house, clothe, and feed themselves, but also to pursue human ends having to do with love and friendship. They need the ability to form family ties, friendships, and special relations with neighbours. They also need codes of conduct. And how could they have all these things without virtues such as loyalty, fairness, kindness, and in certain circumstances obedience?

Why then should there be surprise at the suggestion that the status of certain dispositions as virtues should be determined by quite general facts about human beings?12

One obvious objection to this philosophical programme – zoological naturalism, as we might call it - is that it is one thing for an ethics to dovetail with Aristotle’s biological science, and quite another for it to dovetail with our biological science.13 There is no such thing in modern evolutionary zoology as the notion of flourishing. For evolution, the only thing that counts is surviving long enough to pass on your genes, and for that it is simply immaterial whether you are flourishing or not. In many species, the may-fly for example, breeding is something that happens very late in the life-cycle, when the organism is already literally falling apart.14 Or consider the praying mantis: when the female praying mantis eats the male after they have mated, is the male flourishing?

12 P. Foot, *Natural Goodness*: 44-5.
But it did not, if you ask me, need the emergence of modern evolutionary zoology to tell us that there is something wrong with flourishing, understood in the biologically-based way that Foot and her followers understand it, as a basis for a virtue ethics, or any ethics. The notion fails to fit Foot’s own requirements, in at least two ways.

First, Foot and her school are resolutely anti-consequentialist (and hurrah for that). But zoological-naturalist virtue ethics itself is, or is very naturally understood as, a consequentialist view. Specifically, it is a consequentialism of the dispositions: it tells us to have the dispositions that will most promote flourishing. So zoological-naturalist virtue ethics fits the charge that Derek Parfit and Brad Hooker think applicable to virtue ethics in general, the charge of collapsing into at least indirect consequentialism.15

Secondly, if we are going to resist what Hursthouse likes to mock as “high-mindedness” (she takes John McDowell to be a key exemplar of this vice), and insist that flourishing for humans really is significantly like flourishing for wolves, then it seems impossible to avoid the objection that no such “low-minded” conception of flourishing can possibly be relied on to generate e.g. justice, charity, and temperance as virtues rather than, say, ferocity, cunning, and stealth. In conjunction with other materials, a zoological-naturalist conception of flourishing might produce an intuitively plausible list of virtues. But (first) there again, it might not. And secondly, when things do turn out right, it seems to be the other materials in the argument that have this happy effect - in particular, the account of human reasoning and rationality that the zoological naturalist offers - and not the zoological-naturalist’s distinctive account of flourishing.

Foot and her followers are of course not unaware of these difficulties, and have spent much energy on attempting to address them. Hursthouse, for one, is particularly insistent that her account of human flourishing is not offered from a neutral scientific view: “Everyone who is taking the Aristotelian naturalist line takes it as obvious that they are not pretending to derive ethical evaluations of human beings from an ethically neutral human biology, but are already thinking and… egg deposition. Winged existence may last only a few hours, although Hexagenia males may live long enough to engage in mating flights on two successive days, and female imagos that retain their eggs may live long enough to mate on either of two successive days... Mating is completed on the wing. After her release by the male, the female deposits her eggs and dies. A few species are ovoviviparous—i.e., eggs hatch within the body of the female generally as she floats, dying, on the surface of a stream or pond.”

15Derek Parfit has frequently expressed this view in correspondence with me, as a reason why he does not need to engage with virtue ethics in, for instance, the project of conciliating the different moral theories that he undertakes in his On What Matters, Oxford – New York, Oxford University Press, 2011. Brad Hooker argues the case that Judith Jarvis Thomson’ sand Rosalind Hursthouse’s virtue ethics are both really forms of indirect consequentialism in his “The Collapse of Virtue Ethics”, Utilitas14.1 (2002): 22-40; Hursthouse responds to Hooker (though not mainly to this charge) in the same issue, pp.41-53.
of human beings in an ethically structured way”. But first, one suspects a little hyperbole here: it is less than clear that either Philippa Foot or Michael Thompson, for instance, do in fact take this to be obvious. And secondly, if Aristotelian naturalists are indeed not “pretending to derive ethical evaluations of human beings from an ethically neutral human biology”, then there is a serious question what their often fairly detailed claims about (human and other) nature are actually for - what work those claims do to shape their moral theory if not, as I have been suggesting, to ground it in the factual and descriptive matter of “Aristotelian categoricals”. Thus for Hursthouse in Chapters 9-10 of her On Virtue Ethics - the most plausible response to these problems that I have seen - the naturalistic foundation of her virtue ethics is reduced to the four points that humans are social, that they seek enjoyment, that the continuance of the species is a priority for them, and they have close and particular ties with particular others, especially their families. All of which seems obviously right, and to furnish us, as Hursthouse says, with some important constraints on what ethics can be for creatures like us. Yet none of these claims seem necessarily dependent on anything like the kind of zoological naturalism that Foot lays out in Natural Goodness, or that Michael Thompson lays out in “The representation of life”.

What goes wrong in zoological naturalism, I think, is at least partly the philosopher’s characteristic mistake of over-ambition. Like Plato, the zoological naturalists seek a single uniquely complete and correct account of how to generate the virtues, from the ground up; like Plato, the picture they end up with is unconvincing.

II

It is also, I suggest, deeply un-Aristotelian. It isn’t Aristotle’s project in the Nicomachean Ethics to derive a list of virtues solely and exclusively from an account of flourishing. And this is not because his project is, rather, to derive a list of virtues exclusively from an account of human psychology. As already pointed out, it isn’t his project to derive a list of virtues at all. A fortiori, he isn’t trying to derive a list of virtues from anyone source exclusively.

---

17 P. Foot, Natural Goodness, chapter 2.
20 Hursthouse often remarks that she rejects foundationalism; but the relation of this rejection to her endorsement of the idea that the nature of morality, for us, depends on the natural facts about us, is not clear.
So what is he doing? Well, he’s doing what he says he’s doing in the Ethics itself (7.1, 1145b3-8):

We must, as in all other cases, set the observed facts before us and, after first discussing the difficulties, go on to prove, if possible, the truth of all the common opinions ... or, failing this, of the greater number and the most authoritative; for if we both refute the objections and leave the common opinions undisturbed, we shall have proved the case sufficiently.

Aristotle’s thoughts about the virtues, and his presuppositions about which is the right list of the virtues, begin in and with his own tradition. As he explains here, that does not mean that his ethics is necessarily wide open to the relativist accusation that he simply takes a list of virtues for granted. Tithenai ta phainomena is sometimes translated as “accept the appearances”; but “posit them as a starting point” would be better. What Aristotle undertakes to do in these lines – which happen to be prefaced to his discussion of akrasia, but could stand as his statement of method at almost any point in his writings – is to start from the endoxa (common opinions) of tradition; but not to end there. The common opinions, the traditional views that we’ve inherited, are to be exposed to each other, to check their internal consistency. But they are also to be exposed to whatever “difficulties” (aporiai) they may seem intuitively to face. And this - I suggest, though I admit the point can’t be proved - is not just a test of their internal coherence, but also of their correspondence to the way things are in the world beyond them.

The method that I think we can draw out of Aristotle’s words here is not just a method for arriving at a list of the virtues; it is a method of quite general usefulness in philosophy. As John McDowell puts it:

It is a deep truth that all thinking, just as such, is anchored in traditions. Reflection has nothing to go on, anywhere, but a putative grasp of the that, which (at least to begin with) is merely inherited.\(^{21}\)

The method is not the Cartesian one of getting rid of everything we already think, and trying to start somewhere else, somewhere suspended in the vacuous abyss of “pure inquiry”. The method, rather, is to start from everything we already think, and subject it, not to one single all-purpose philosophical test (such as, for instance, Cartesian doubt), but to a variety of different tests and questions that we can use, not usually just to abandon our initial view, but to refine it.

This way we can have the great advantage of traditionalism or conservatism, that it starts off with our pre-philosophical opinions. (As Aristotle says in NE 1173a1-3, there isn’t really anywhere else to start.) Yet we avoid its great disadvantage, that of being insufficiently critical about those starting-points for

thought. The critical element is supplied by the questions and tests that we apply to our initial presumptions. And our balance, our common sense, and our resistance to fanciful skepticisms are all preserved by the fact that it is questions, plural, and tests, plural. *Pace* writers like Peter Singer,²² there isn’t one super-methodology that does away with pre-philosophical common sense. Rather, there is a variety of methods for adjusting it, with none of those methods being preeminent and exclusively correct, and the correct application of them being, as always, a matter of judgment.

This way we can also have the advantages of systematic and constructive philosophy, that it will actually be possible for us to present an interesting positive and structured view as our philosophical position, rather than courting the familiar accusation that we are “quietists” or “conservatives” or “just being negative” (or opposed to what we might call edifactiousness²³). Yet we avoid the great disadvantage of system-building philosophy, that we make it look as if our theoretical structure were the only possible one, or as if our philosophical approach were all-or-nothing. We are freed of the supposed duty to derive our list of the virtues “from the ground up”, so that our derivation is either the only possible one and completely inexpugnable, or else completely impossible and a complete failure.

This cautious and piecemeal methodology, balancing a default presumption in favour of received opinion with a willingness to revise it for sufficiently good philosophical reasons, and balancing an openness to the sheer variety of what “good philosophical reasons” might be with a healthy skepticism about the philosophical fanaticism or monocularism that again and again becomes fixated with just one sort of philosophical reason - this, I want to suggest, is the truly Aristotelian method in philosophy, and the method that Aristotle himself tries to apply: with resounding success at times, and resounding failure at other times. In philosophy in general, there are explanations for most things, but there is no one explanation which is the explanation for even most, let alone all, things. In ethics in particular, there are reasons that ground most cases of rightness and wrongness, goodness and badness, but there is no one foundational reason, or kind of reasons, that grounds everything, or anywhere near everything. If we liked we might call this an intuitionist method, in the methodological sense of “intuitionism” that Bernard Williams endorses in “What does intuitionism imply?”.²⁴ Or we might prefer to keep things simple, and just call the method common sense.

---


I have abjured the search for one test or criterion that, all on its own, will give us a complete and definitive list of the virtues; I prefer talk of tests and criteria - plural. So what is in this plurality of tests, and how do the various individual tests interact to give us the right results? And how can we be sure they won’t interact to give us the wrong results? (An obvious piece of arithmetic shows that the number of possible interactions of the tests is a much larger number than the number of the tests: so this multiplicity might seem really worrying.)

The answer to the how-can-we-be-sure question is, of course, that we can’t be sure (in advance): that is what I meant by talking, a moment ago, about judgement. What we can be sure of is that intelligible argument about what the virtues are will always be conducted by appeal to these tests. But there is no guarantee that all intelligible argument about the list of the virtues will be plausible argument about the list of the virtues. Working out what is wrong with an implausible argument, or an implausible list, will often be a subtle and delicate matter, and will usually be a case-by-case one too.

We should enter another reservation too at this point, and before we go any further. This is about the supposed definitiveness or objectivity of the list of the virtues that we might hope to end up with by deploying the multi-criterial and intuitionist method that I have sketched. Here too there is a contrast with the over-ambition, as I see it, of zoological naturalism, which as we saw above, at least aims to give us a completely objective and definitive list of virtues, in principle applicable to all societies, but as far as I can see, specific to none.25 These claim strikes me as implausibly over-ambitious; and as unnecessary, even for those who, like me and the zoological naturalists, agree in endorsing moral realism.

Notice here the philosophical advantages of moral realism over moral irrealism of whatever stripe. If moral irrealism is the thesis that no moral propositions are objectively true or false, moral realism is simply the negation of that thesis. Moral irrealism, therefore, is refuted by a single example of a moral proposition that is objectively true, or false; moral irrealism is intolerant of objectivity.26 Moral realism, by contrast, need not be at all intolerant of subjectivity: since it is the thesis that some moral propositions are objectively true or false - and of course, preferably the central and important ones - there is room for subjectivity within moral realism, in a way that there cannot be room for objectivity within moral irrealism.

So in the present case, while the moral irrealist cannot say that there is - really, ultimately - any such thing as the correct list of virtues, and must say that there is

25 “Morality which is no particular society's morality is to be found nowhere”: MacIntyre, After Virtue, Second (corrected) edition (with Postscript), Norfolk, Duckworth, pp.265-266.
no such thing, the moral realist is free to say either that there is or isn’t such a correct list. Provided the moral realist takes something (and, of course, something important) to be objective, there isn’t even an apparent threat to his moral realism if he denies that lists of virtues are objective, or are fully objective.

Well, I am a moral realist; and although I don’t actually want to go as far as outright denial that there is such a thing as the correct list of the virtues, I do want to say that there is often something rather beside the point about anxious queries whether this or that list of the virtues is or is not “objectively correct”. To take it that lists of the virtues are actually meant, by those who propound them, to be - and to be nothing but - flatly descriptive of an antecedently given reality seems to me a rather naïve kind of literalism. Very often the intentions of those who offer us such lists are, and are patently, prescriptive rather than descriptive. They are not so much attempts to describe a reality that is already there, as to summon a reality into being by sketching an ideal and exhorting one’s hearers to live up to it.

This hortatory function of lists of virtues has a number of explanatory applications. For one thing, it makes it perfectly explicable, for the moral realist, why there should be so many differences between different societies’ lists of virtues. Some of those differences no doubt do reflect substantive philosophical disagreements: consider, for instance, the differences noted by MacIntyre between Jesus’ and Aristotle’s conceptions of the virtues, or again the differences between the virtues of a Socrates and those of a Thrasy machus.27 But many other differences are merely a matter of emphasis or of the division of topics. Between somebody who (say) takes gentleness to be a primary virtue and humility to be an offshoot or subcategory of gentleness, and someone who sees things the other way round, there need be no more than a difference in the order of exposition.

The hortatory function of lists of virtues is particularly obvious, I want to suggest, in the writings of a figure who - along with Plato, Aristotle, and Aquinas - is one of the three or four most influential propounders of lists of the virtues that Western ethics has ever seen, namely Paul of Tarsus. One list of his we have cited already:

The fruit of the Spirit is love, joy, peace, patience, kindness, goodness, faithfulness, humility, and self-control.

The δὲ καρπὸς τοῦ πνεύματος ἐστιν ἀγάπη, χαρά, εἰρήνη, μακροθυμία, χρηστότης, ἀγαθωσύνη, πίστις, πραΰτης, ἐγκράτεια: κατὰ τῶν τοιούτων οὐκ ἔστιν νόμος. (Galatians 5.22-23)

There are plenty of others:

So, as those who have been chosen of God, holy and beloved, put on a heart of compassion, kindness, humility, gentleness and patience; bearing with one

another, and forgiving each other, whoever has a complaint against anyone; just as the Lord forgave you, so also should you forgive each other. Beyond all these things put on love, which is the perfect bond of unity. Let the peace of Christ rule in your hearts, to which indeed you were called in one body; and be thankful. Let the word of Christ richly dwell within you, with all wisdom teaching and admonishing one another with psalms and hymns and spiritual songs, singing with thankfulness in your hearts to God. Whatever you do in word or deed, do all in the name of the Lord Jesus, giving thanks through Him to God the Father. (Colossians 3.12-17)

Finally, brethren, whatever is true, whatever is honourable, whatever is right, whatever is pure, whatever is lovely, whatever is of good repute, if there is any excellence and if anything worthy of praise, think on these things. (Philippians 4.8)

Let all bitterness and wrath and anger and clamour and slander be put away from you, along with malice. Be kind to one another, tender-hearted, forgiving each other, just as God in Christ also has forgiven you. (Ephesians 4.31-32)

It isn’t that there are no standards of correctness for such lists; if St Paul had recommended that the Ephesian Christians seek not only to be kind, forgiving, and tender-hearted, but also louche and cynical, there would have been something wrong with that. Or if he had recommended that the church at Philippi think not only on the true and the lovely, but also on the crafty, that too would have justified a readerly double-take. There are standards of correctness for these lists, but the standards are standards for exhortatory, not for descriptive, catalogues. Contrast the zoological naturalists and their would-be straight-descriptive lists of dispositions that, as a matter of fact, promote flourishing.

Though come to think of it, perhaps the zoological naturalists are also wrong about the sort of list that their account of the virtues implies: perhaps a list of the characteristics that, say, a wolf needs in order to live a life that is long, healthy, and largely undisturbed except by predation-opportunities is also rather more open-ended than they like to admit. (One obvious thought that points in this direction: what those characteristics are is determined in large part by the wolf’s environment; and environments change.) Still, as I say, their list is (at least

---

28 Is this a list of virtues at all? Isn’t it rather a list of things to meditate on?” It is certainly the latter; its being the latter does not preclude its also being the former, since these subjects of meditation are apparently classes of actions or of the dispositions that produce them. Though even if it isn’t a list of virtues, the relevance to a virtue ethics of such a list of reflective practices, or topics for reflective practice, is obvious; it is a key part of virtue—Paul is telling us—to reflect on things like these.

29 And of course behind all St Paul’s lists of virtues stands Jesus’ list, the Beatitudes: Matthew 5.1-12.
Lists of the virtues

intended as) a *straight-descriptive* list: it is meant to capture the facts about what characteristics actually achieve or tend to achieve a supposedly given end, namely biological or quasi-biological flourishing. If such a descriptive catalogue of the virtues is, in fact, open-ended, it will be so for quite different reasons from the reasons that make for open-endedness with St Paul’s lists of virtues and ideals.

We are close here to the large and interesting question, what a list of the virtues is supposed to be for anyway; exhortation has been suggested as one possible function, but maybe there are others. At any rate it would be bizarre to suppose that the list of the virtues is meant to be deployed directly in agents’ deliberation, as their main means of thinking about what to do. The zoological naturalists do not suppose this: that is why Hursthouse often invokes the notion of the “v-thoughts”, the thoughts that motivate the virtuous person, which can be but are not necessarily coincidental with the thoughts about the virtues that, according to her, give the criterion of rightness for all action.

It has often been argued that this manoeuvre rescues Hursthouse’s virtue ethics from an implausible picture of how agents deliberate, at the price of making the very distinction between deliberative procedure (DP) and criterion of rightness (CR) that makes so much trouble for utilitarianism.30 Without getting too deeply into this debate here, I will say simply that the DP/CR distinction in itself seems to me entirely unproblematic. The problem comes only when a theory is forced to suppose that an agent is bound to entertain thoughts on the one side of the distinction that subvert or contradict thoughts that she is bound to entertain on the other side. Utilitarianism of every variety known to me is clearly refuted by its failure to avoid this problem about internal consistency; it is less clear that every version of virtue ethics is.

Anyway, St Paul’s practice gives us further evidence that is that there are other things that lists of virtues can be for, besides straightforward description of moral reality, and the equipping of the agent with materials for (more or less direct) deliberation, and for devising and applying a criterion of rightness. A list of virtues can also be for _meditative attention_: it can be an ideal, or constellation of ideals, that we reflect on in order to internalize. It is clear that St Paul thinks that such meditative attention is a powerful form of moral discipline, and a powerful source of moral transformation. And on that, of course, modern thinkers such as Iris Murdoch will agree with him.

---

So though there are certainly (according to realists like me) antecedent realities that a list of the virtues needs to fit, presenting a list of virtues is not merely a matter of describing antecedent realities. As the case of St Paul makes clear, it is also, at least as much, a matter of prescribing: of exhorting one’s hearers to see things in a certain light, to dispose themselves in particular ways, to attempt to approximate some ideal or ideals, and to “think on” those ideals: to reflectively internalise them.

We can keep this thought in mind as we turn to the question that section III left outstanding, the question of what the tests are that should determine our list of the virtues. After the remarks I have just made about the exhortatory value of such lists, it would be odd to attempt to offer a complete and definitive answer to this. But I will offer three suggestions. The first is about *eudaimonia*; the second is about the notion of a *technê* or practice; and the third brings us back to the idea of reflecting on and imitating ideals or exemplars.

The first suggestion, then, is just the idea that the notion of human flourishing or *eudaimonia* can give us some help in formulating a list of virtues. Perhaps it looks as if I have already excluded all reference to *eudaimonia*; but actually what I have rejected is eudaimonism, by which in this context I mean the view that the point of the virtues is to promote *eudaimonia*, and so that our list of the virtues can be derived *solely* from thinking about what promotes human flourishing. I have denied this in particular where the notion of flourishing that we are working with has the misleadingly scientific look of zoological-naturalistic flourishing. But to deny these claims is not to deny that there is *anything* to the notion of human flourishing. Of course there is, but the notion is a “folk” notion, not a scientific one: what we count as flourishing is not part of our science but of our form of life.31

My second suggestion arises from MacIntyre’s discussion of the nature of the virtues in *After Virtue*. MacIntyre’s approach rephrases the question “Which character-traits are the virtues?” as the question “What character-traits do we need for successful pursuit of the practices?” His account of what a practice is - the notion has obvious affinities to Aristotle’s and Plato’s notion of a *technê* - is this:

Any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and the goods involved, are systematically extended.32

---

31 A point which, as we have seen, the zoological naturalists of course attempt, at least sometimes, to accommodate. In my view, not successfully; but that is an argument for another time.

Lists of the virtues

Examples of practices that MacIntyre gives include farming, chess, music, science, history, novel-writing, and (a very Aristotelian example, this) politics. As MacIntyre brings out, the mark of all these practices is that in order to practise them, one needs honesty, humility, preparedness to learn from others, and responsiveness to the legitimate demands of others: so one needs something very like justice. One also needs persistence and self-discipline – so something very like temperance – and the optimism to keep going when the practice seems impossibly difficult or overwhelmingly complicated: so something very like courage. And what one learns through the practices is how human understanding and expertise is articulated in particular cases and contexts, and how to move from those particular contexts and apply their lessons to the overall context of living our lives: so something like wisdom too.

The flourishing of the virtues requires and in turn sustains a certain kind of community, necessarily a small-scale community, within which the goods of various practices are ordered, so that, as far as possible, regard for each finds its due place within the lives of each individual, or each household, and in the life of the community at large. Because, implicitly or explicitly, it is always by reference to some conception of the overall and final human good that other goods are ordered, the life of every individual, household or community by its orderings gives expression, wittingly or unwittingly, to some conception of the human good. And it is when goods are ordered in terms of an adequate conception of human good that the virtues genuinely flourish. “Politics” is the Aristotelian name for the set of activities through which goods are ordered in the life of the community.33

We do not learn the virtues only through the practices; but that is one very obvious place where we do learn them. When we look at the dispositions of character that are required for expertise in some particular practice, what we arrive at very quickly comes to look pretty much like a list of the virtues that we need in any practice, and in the living of our lives overall.

Here then is a second way of arriving at some ideas about what the virtues are. It is a way that Iris Murdoch at least gestures towards as well:

If I am learning Russian, I am confronted by an authoritative structure which commands my respect. The task is difficult and the goal is distant and perhaps never entirely attainable. My work is a progressive revelation of something which exists independently of me. Attention is rewarded by a knowledge of reality. Love of Russian leads me away from myself towards something alien to me, something which my consciousness cannot take over, swallow up, deny

---

33MacIntyre, Preface to the Polish edition of After Virtue.
or make unreal. The honesty and humility required of the student - not to pretend to know what one does not know - is the preparation for the honesty and humility of the scholar who does not even feel tempted to suppress the fact which damns his theory... Studying is normally an exercise of virtue as well as of talent, and shows us a fundamental way in which virtue is related to the real world.34

My third and final suggestion about the tests that we might use to derive a credible list of the virtues is that at least some of the content of any plausible list of virtues is generated by thinking about examples of good people, and especially of remarkably good people - saints, or ideals, or exemplars of the virtues.

Exemplarism, the thesis that we can derive our conception of the virtues from thinking about exemplars of the virtues, is a venerable and important part of the virtue tradition. One way to develop it I deployed myself, in a paper that I toured for a couple of years in the early nineties without ever managing to get it published. The idea was to work from thin descriptions to thick: the key question was “If you know of Jane only that she is a good person, what further descriptions may you reasonably expect to be true of Jane?” I thought then, and twenty-five years later I still think, that there can be an interesting answer to this that gives us a good deal of information about the nature of the virtues. To be sure, we get that information in the form of a long, loose and defeasible disjunction; and to be sure, the information we thus get is filtered via my or our reasonable expectations, which no doubt are both cognitively imperfect and culture-relative. But these points are philosophical commonplaces. They constitute serious obstacles only to an attempt to use this “key question” as the unique method that definitively gets us to everything we need to know about the correct list of the virtues. As should be obvious by now, I am not in that game at all; I doubt I was, really, even in 1991.

In recent philosophy the exemplarist approach to virtue ethics has been argued for by other strategies than the one I tried out in that old draft of mine: by Linda Zagzebski in one way 35, and by me in another36. For both Zagzebski and myself, Aristotle is the source and authority for this thesis whom we quote the most, though we could also have quoted St Paul, who repeatedly tells the readers of his letters to imitate Jesus (e.g. Galatians 3.27, Philippians 2.5, Ephesians 4.14). (Paul also tells them to imitate himself at least twice: 1 Corinthians 11.1, Philippians 3.17.). The Christian tradition has made rich use of exemplarism: that, for instance, is one reason why the church has the lives of saints to meditate on.

Lists of the virtues

Part of the point of exemplarism is, as Zagzebskistresses, that a living exemplar of a given virtue gives us far more detailed information about what the virtue actually involves than any abstract description or definition of that virtue could. We might say, as she does, that what we learn from the exemplar is basically knowledge by direct ostension – “The virtue is like that” – rather than knowledge by definition. Or we might make what seems to be a closely-related point, that what the exemplar gives us is not, or not only, explicit and propositional knowledge of the virtue, but also tacit and non-propositional knowledge of what it is like for someone to have that virtue; or perhaps practical knowledge – knowledge how to exercise the virtue. Or again, we could observe with Iris Murdoch that attention to exemplars of virtue – those found in novels and plays, to give two obvious examples – feeds our imagination and our moral vision in a richer and psychologically deeper way than if-and-only-if equations do. Or we might make all of these points, as in fact I think we should; they seem perfectly consistent with each other. And what they add up to, in combination, is a striking picture of how much more there can be to moral knowledge than straightforwardly propositional knowledge.

How, though, do exemplars get established as exemplars? Very often, I suggest, it is because some person or some deed comes across to us as immediately admirable. In print or in person, we come across some Gandhi or St Francis or Martin Luther King or Sophie Scholl, and that person strikes us – directly and primitively – as awesome, as having done something noble or wonderful, perhaps even beautiful. Two well-known cases of this sort of experience are given by Rai Gaita early on in Good and Evil: the cases of the nun in the hospital working selflessly and unendingly to relieve the sufferings of her patients, and Primo Levi’s description of the sufferings of Ladmaker in Auschwitz: “Charles’s behaviour showed a goodness to marvel at”.37

What thinking about exemplarism gets us to see here, in fact, is something that we might find profoundly missing from approaches to virtue ethics such as zoological naturalism. It is that any plausible and attractive list of the virtues is going to depend at least as much on the notion of the morally fine or beautiful – in Greek, on to kalon – as on the notion of flourishing or the advantageous – in Greek, on to ophelimon. The virtues are not just the dispositions that tend to lead us to the desirable life. They are also, and perhaps even more fundamentally, the dispositions the exercise of which is admirable, and makes us admirable people.38 (But the point, of course, is to be admirable, not to be admired; it is to have in us what necessarily makes admiration apt, not what actually prompts admiration. Here talk of the beautiful is preferable to talk of the admirable, precisely because it lacks this misleading connotation.)

38 For the admirable/desirable contrast cp. Linda Zagzebski, op. cit.
We can put this as a point about the familiar old question “Why be moral?”.

So put, the point is that the answer to “Why be moral?” is quite often “Because that is the beautiful thing to do”. It’s not that the moral act is itself prudentially disastrous, but just happens to be, unfortunately enough, one of a class to the whole of which we are somehow committed, if we are committed to any part of it - as theories like rule-consequentialism and Gauthier’s contractualism too often tend to suggest. Nor is that the moral act is prudentially advantageous in some way - just a very obscure way, one which is consistent with the fact that the moral act is attended with terrible penalties like those that Hans and Sophie Scholl faced, or those described by Callicles in Plato’s *Gorgias* (486b). We need not think that there is any prudential advantage, in any sense, in the gravely-sacrificial moral act. (Notice here how close to the surface we find a false assumption that I have criticised elsewhere:40 that reasons for action have to be future-directed.) At least in some cases, advantage simply isn’t the point. It is rather that the moral act demands to be done even if it does involve a grave sacrifice - just because it is beautiful.

Perhaps this appeal to to kalon, The Beautiful, is the answer to the puzzlement expressed by the person who said of Sophie Scholl and those who suffered with her that “the fact that five little kids, in the mouth of the wolf where it really counted, had the tremendous courage to do what they did, is spectacular to me. I know that the world is better for them having been there, but I do not know why.”41 Perhaps it is also the best answer to the difficulty that Philippa Foot is struggling with in her rather convoluted discussion of the “Letter-Writers”, a group of victims of the Nazis who thought it was worthwhile to die rather than to give in to Hitler.42 Foot’s difficulty is, precisely, to square the Letter-Writers’ willingness to die, and their obvious virtue, with Foot’s own eudaimonism: the virtues are supposed to lead, or at least tend to lead, to flourishing, and here the virtues are, in all their glory, leading their possessors directly to death. (And not only leading but tending to lead, given the nature of Hitler’s Germany.) Despite repeated study of what Foot says about their case, I am not entirely sure how she thinks their case can be squared with the eudaimonist idea, which goes back at least to Socrates, that “the virtues benefit their possessor”. But I know how I want to respond to their case. On the grounds precisely of cases like the Letter-Writers’, I simply deny that the virtues do necessarily benefit, or even tend to benefit, their possessor. As the

---


40 In Chapter Three of *Knowing What To Do*.

41 Lillian Garrett-Groag, quoted in the Wikipedia article on Sophie Scholl. The remark is quoted—from the same source—and discussed by Eleonore Stump in *Wandering in Darkness*, Oxford, Oxford University Press, 2010:149 and 549. In the present context we are close, of course, to Gaita 1991’s remarks about “a goodness to marvel at”, cited in the last chapter; there is bound to be some overlap between the notions of glory and of nobility.

Letter-Writers show, there are cases where the exercise of the virtues can be utterly disastrous for their possessor - and yet, the virtues continue steadfastly to point us towards “the thing to do”. For the Letter Writers, what their virtues do is make their possessors, and their terrible submission to the suffering that confronted them, admirable/ fine/ beautiful/ kalos. But that, in the circumstances, was the very opposite of benefiting them.

Eudaimonia, the dispositions that we need for the practices, the power of examples: all of these are resources that we can appeal to when attempting to refine our list of the virtues, or assess whether we really think that this or that disposition of character is a virtue or not, and why. No doubt there are other resources too. My suggestion is not that any one of these resources would give us all we needed to derive a list of the virtues - even if deriving a list of the virtues from scratch, rather than refining the list(s) of the virtues that we have already inherited from our traditions, were really what we are engaged in doing. Rather, the three resources that I have looked at here suggest tests and diagnostic questions and criticisms that may be applied to already existing lists of the virtues. Such evaluation of lists of the virtues is therefore more like a matter of good judgment than of the application of a simple algorithm. But that, of course, is precisely what we would expect, and entirely as it should be.
Virtue, Skill and Vice

Julia Annas
University of Arizona
Department of Philosophy
jannas@u.arizona.edu

Abstract
It is widely accepted among philosophers and developmental psychologists that the acquisition of a virtue of an Aristotelian type can be illuminated by the acquisition of a practical skill. This model does not seem to fit the acquisition of a vice. The paper examines the issues involved and ways in which the skill model does and does not fit the acquisition of vice.

Keywords
Virtue, vice, skill, habituation, aspiration.

In recent years, as is familiar, there has been a flow of work on virtue in many fields, and notably in philosophy, mostly in the area of virtue ethics. Here I want to take up a point which is quite surprising, namely that this intense focus on virtue has not similarly focused on vice. Once you notice this, you scan complex works on virtue in vain for anything comparably developed on vice.1 It seems to be assumed that an account of virtue does not need also to provide an account of vice for it to be adequate as an account of virtue. I have been as guilty of this as anyone, and it may be that this assumption is not arbitrary. Aristotle, for example, tells us that we are studying virtue not for the sake of it, but to become better people2, and we can see why study of virtue alone would suffice for that, bringing in vice only insofar as we need to understand it to understand virtue.

However, vice is of interest, certainly to me, insofar as an account of virtue raises puzzles about the nature and structure of vice. This is especially so for those of us who defend an Aristotelian (or neo-Aristotelian) account of virtue, in which it is prominent that virtue is acquired in a way analogous to the way in which a practical skill is acquired. For such an Aristotelian account, it is central to the account of virtue that its structure, and so the way it is taught and learned, resembles that of a practical skill. Yet this does not carry over unproblematically

1 Theologians do not share this neglect to the same extent, but it is often difficult to make use of their work if you do not share the relevant tradition of understanding the idea of sin. Aquinas, for example, understands vice partly in terms of sin, which makes his work on vice less accessible to non-religious philosophers than his work on virtue.

2 Nicomachean Ethics II 2, 1103b26-30.
Virtue, Skill and Vice

to vice. This would not be a worry, of course, if we did not expect vice to be anything like virtue in its structure. But we do think of vice as being, in some respects, more like virtue than either is like other ethical states, and so we are faced by a problem: why do we think this if vice does not have the structure of a skill? What is the structure of vice? Some vices at least seem to display intelligence and application; how is this related to practical skill?

I hope to broach a topic which deserves more attention than it has had. I am not aiming to give anything like a complete account of vice, and there are many issues about vice and the vices which I must leave untouched. In order to get a handle on understanding vice I am focussing on the skill analogy for virtue, the question of its fit for vice and how we are to understand the intelligence often shown by the vicious.

I will begin with a brief account of an Aristotelian account of virtue, familiar though it is, in order to have a clear view of what it is that vices are being contrasted with. This account lays out the main features, and is obviously not a definition of any kind. A virtue is a disposition of a specific kind, to think, reason, feel and act in certain ways, namely the virtuous ways. This disposition has a specific history, namely of being built up through experience by selective and intelligent habituation, where this is guided by learning from those who, in the culture, already have the relevant virtues. (This can take the form of role models, books, movies, and so on.) What makes these dispositions virtues is that they aim at the good. When we are old enough to reflect on the virtues we have learned from our culture, we are able to revise our positions as to whether these virtues do, in fact, aim at the good, or are merely conventional, or even deeply misguided. There are various ways of interpreting the good that is aimed at, and in this paper I shall be assuming that we are dealing with a secular interpretation, where the good aimed at is that of a flourishing human life. The ‘skill analogy’ comes in because the habituation in virtue takes the form of the kind of education in acting that you get in making and acting when you learn a practical skill.

It’s important for this conception of virtue that virtue be a stable disposition, different from a mood or temporary commitment. It’s also important that habituation forms and educates our emotive and feeling, as well as our cognitive, aspects. Aristotle puts this in a striking way when he says that it is a mark of the virtuous person that she performs virtuous actions with pleasure rather than pain. What he envisages is the difference between, on the one hand, doing a generous act, but with a conscious effort, awareness of being pulled to do something else and with felt regret afterwards; and, on the other hand, just doing it readily and effortlessly. The virtuous person will just do it; her intelligent and educated disposition encounters no block from unwillingness, and so she experiences the

3 This is particularly important when focusing on vice, since non-secular views of flourishing often understand vice in terms of a non-secular concept like sin.
kind of pleasure in activity that Aristotle is concerned with. This is of course not
the kind of pleasure that we think of as a perceptible feeling, but the kind of
pleasure taken in effortless activity, whose hallmark is that it encounters no
obstacle from contrary motivations on the agent’s part. Virtue is *motivationally unconflicted* in a way that contrasts with the person who does the right thing but
unwillingly and so effort fully. Using Aristotle’s convenient term, we can call this
latter person the enkratic. The enkratic person succeeds in doing the right thing,
but has to combat contrary motivation in order to do so, and is not acting from an
unconflicted state.

We also take it, along with Aristotle, that there is a similar contrast between
the person who just does something cruel or mean without felt conflict or regret,
and the person who does it, but whose action registers with him as a lapse from the
conviction that this is the wrong thing to do, and who subsequently does feel
regret. Again using Aristotle’s convenient term, we can call the latter person the
akratic, and the former vicious.

Here we meet a point which needs to be considered in contemporary
discussions of vice. The above contrasts take over Aristotle’s contrasts between
virtue and vice on the one hand, and enkrateia and akrasia on the other – but do
we actually understand Aristotle’s conception of vice? In contemporary society we
are discouraged from being ‘judgmental’, and the idea of vice is often trivialized, so
that people will say that their vices are indulgences like food or buying shoes. This
might not matter much, as it does not much matter that we do not in ordinary
conversation use the term ‘virtue’ much: we still recognize virtues like bravery and
generosity, and we still recognize vices like cowardice and cruelty. So we might
retain the concept of vice, but in the form of recognizing vices rather than that of
trying to take over the common usage of the term ‘vice’.

An objection remains, however. We use the term ‘vicious’, but no longer use it
to cover all or even most of the vices. We talk of people as vicious when they are
cruel or aggressive; it sounds odd to us to say that cowardly or selfish people are
vicious. This might suggest that there is a deeper problem than our use of ‘vice’ in
finding unity among the vices. We could, it seems, recognize the vices without
finding any common structure in them. If so, then virtues would have a shared
structure, but vices would not, so that aiming to give a common account of the
vices would be misguided.

This is a serious objection, but there are two powerful points against it. One is
that much the same kind of objection could have been (and sometimes was)
brought against study of virtue as a unified entity before the rise of interest in
virtue and virtue ethics. In contemporary life ordinary discourse about virtue and
the virtues had been philosophically ignored for over a century, so it is no surprise

\footnote{4 Obstacles from outside the agent are another matter; here I am concerned with the lack of internal conflict in the virtuous.}
that clear distinctions, and clear outlines of virtue ethics, were not to be found on the surface. After forty years of philosophical and psychological exploration, we now do have a clear view of the many different options there are for virtue and for the role of virtue in ethical theory. (They are still, of course, very disputed.) Vice has not been the object of this kind of intensive and extended discussion, so the fact that our present discourse does not deliver clear results about vice may well not be philosophically significant; it may indicate no more than that ordinary discourse reflects an uncritical cultural tendency to associate using vice terms with being ‘judgmental’, which is acceptable only in cases of clear harm to others. This may be no more philosophically significant than the fact that before the rise of discussion of virtue, many people associated that term with prudery about sex, something now seen as obviously culture-bound and no longer relevant.

Rehearsing Aristotle’s contrasts makes us realize that, independently of ordinary discourse, we do recognize the difference between the person who does not help but later regrets it, and the person to whom it does not even occur to offer help, and we also realize that this lines up with the difference between the person who helps but has to overcome reluctance in order to do so, and the person to whom it does not even occur not to help. We recognize, that is, a state that corresponds to virtue in being unconflicted, but contrasts with virtue in being unconflicted about doing the wrong thing rather than the right thing. We are implicitly recognizing a concept of vice which so far corresponds to Aristotle’s: it is internally unconflicted by contrast with the states of the people who act in ways that express motivational conflict. This gives us something important to begin from: virtue and vice are more stable than the motivationally conflicted states.5

One obvious implication of the discussion so far is that vice will be a character trait. Will it be the same kind of character trait as virtue? The vicious equally with the virtuous act in accordance with their settled and motivationally internalized convictions as to what they should do. Virtue, however, is a state which is the result of a certain kind of history; it can’t be acquired by reading books or by sheer will-power, but requires a specific kind of education, namely habituation through experience acquired in an intelligent way by learning from the virtuous. This is what is illuminated by the model of practical skill; virtue involves more than skill (especially its orientation to the good), but we cannot do better than Aristotle’s famously mundane view that, since things that we learn to do have to be learned by doing them, learning to be just is like learning to be a builder: you learn to do what the experts do, and in so doing you learn to do it in the right way, the virtuous or expert building way. You acquire not a routine but the right training and grasp of the skill or virtue. If vice is a character trait like virtue, then, we

---

5 I am using ‘state’ here in a generalizing way to cover all four of the conditions I am concentrating on; this does not imply being static as opposed to being dynamic.
would expect it to result from a history of learning along the lines of a practical skill. However, this does not go smoothly.

Two aspects can be picked out of the skill analogy for virtue: I have called these the need to learn and the drive to aspire. The need to learn is the more obvious; I don’t learn to build by foolishly starting from scratch, but by learning from people who already know how to do it. With virtue I cannot choose to learn from scratch, since I begin learning when very young, before I am in a position to learn critically. I learn from various sources in the culture: role models, books, in large part my parents and local peers. Some people, whether in real life or other parts of the culture, are role models from whom I learn, in various ways, to try to be and to act like them.

Here at the start there is a disanalogy with vice. We don’t look up to cowards and dishonest people and try to be like them. We don’t make movies honouring famous cowards, with film stars competing for the role; cowards feature as butts in comic movies in which they are the objects of derision. We don’t use well-known cowards as sources of understanding to be conveyed to eager learners; people aim not to become well known as cowards.

The drive to aspire also does not seem to fit vice. The learner in virtue (ideally) comes to understand the point of what she learns to do, to come to acquire the skill in a way which is hers and not just mere imitating, and to keep improving at doing what she has learned. In all these ways she moves beyond merely doing what a teacher or role model does, and comes to grasp what the point of it is, so that she will respond to a new situation in a way which will probably be different from what she has learnt in being a different type of action, but also the same as what she has learnt in being an exercise of the same virtue that she learnt in the previous type of situation. All of these points lead to absurdity when we think of vice. We do not ask ourselves whether an action really responded to the situation in an appropriately cowardly way. We don’t ask if we were as cowardly as the situation demanded (indeed it sounds absurd to think of the demands of cowardice). We don’t ask, after the action is done, whether we were cowardly enough. We aren’t praised for the cowardice of our behaviour and encouraged to improve on this by making the next action an even better example of cowardice.

These points suggest that vice is not the same kind of character trait as virtue. Obviously people do learn to be dishonest, greedy and cowardly, but it can’t be in the kind of way that they learn to be honest, temperate and courageous. This is significant for the structure of virtue and of vice, since the issue of how we learn to be virtuous is not separable from the issue of what virtue is. A virtue is a disposition which has come about as a result of a certain kind of history, and its

---


7 These three aspects of the drive to aspire are explored further in *Intelligent Virtue* (n.6).
nature can’t be understood independently of that history. We expect the same to be true of vice; but it can’t be the same kind of history.

Yet vice is a character trait, a disposition, which, as we acquire it, ends up with our being motivationally unconflicted in a way comparable to that of the virtuous. Perhaps it is a mistake to apply the skill analogy to vice? Perhaps the history that leads to the character trait of vice is different from the history that leads to the character trait of virtue and does not involve anything to which the skill analogy is appropriate.

One answer which I will not explore here is that vice is a positive aspiration to evil in a way that mirrors virtue’s aspiration to good. This can be articulated in ways other than the one just rejected (aspiration to be cowardly, stingy and so on). I am leaving this aside in this paper because I am thinking of vice in terms of an account of virtue which is broadly Aristotelian in structure, in which, as we have seen, positive aspiration to evil does not fit. A full account of vice would explore this option.

A common and widely appealing way of thinking of vice is that is falling for temptation. The coward knows that she should not run away, but is unable to stand by that conviction because the thought of reaching her own safety tempts her to run, and she gives in the temptation. This account recommends itself for greed; the person is fully convinced that he should not have more to eat, but ‘gives in’ or ‘falls for’ the temptation to have that extra piece, and so acts greedily. Dishonesty is also readily seen on this pattern. The person knows that she should not swindle an easily fooled client, but ‘can’t help herself’ going for so easy a profit, and so falsifies the papers. This account, however, does not fit all vices equally well, such as cruelty or envy. In any case, independently of that it has a problem right at the start: it makes vice a matter of failing to stick to what you are convinced is the right thing to do, and this fails to distinguish adequately between vice and akrasia. There are cowards who just run away, gluttonous people who just take another piece of pie and dishonest people who just go ahead with the swindle, with no motivational discomfort, no idea of being tempted to go against what they think they should do. These look like people whose vice is not akratic; they seem to have a settled and unconflicted character trait of some kind.

In ancient ethical thinking there is a running temptation to think of the vicious person as the person who simply goes for what she most wants, neglecting overall priorities. This can be seen as a version of the ‘giving in to temptation’ interpretation of vice. The ancient model of reasoning is often the reasoning which takes in a life overall, as opposed to wants and desires which are focussed only on their own gratification, and this sometimes suggests the model of the vicious person as someone who goes only for what gives them pleasure at the time, without taking account of the priorities of their life as a whole. This picture, however, is inadequate for an account of vice as a whole, as opposed to the particular vice of
intemperance.\(^8\) Firstly, the objectives of the vicious need not be short-term gratifications: they might be as lengthy as Inspector Javert’s years-long vindictive pursuit of Jean Valjean. They need not have anything to do with indulgence; Javert is hardly a model of self-indulgence in his gruelling pursuit. The problem is not that the vicious are aiming at getting self-indulgent gratifications, but that they are going for the wrong things, whether these are what they most want at the time or not. Secondly, it is exaggerated to think that the greedy or dishonest person will have no, or weak, overall values and priorities. Being greedy for money is compatible with running your life overall in a perfectly competent manner, one focussed on getting money, rather than giving money its appropriate place in your life.

As we can see from the above examples, it is not giving in to temptation that is at the heart of vice, but going for the wrong things. Is getting what to pursue wrong what most characterizes vice? In one way this has to be correct, because the virtuous person is the person whose disposition is to get things right – to make the right practical decisions in all their aspects, evaluative and mundane. There is an initial appeal, then, to this idea of vice as failure to do what the virtuous succeed in doing. The coward fails to grasp which kinds of things are really to be feared and which are not; similarly the gluttonous failed to grasp limits to indulgence in food, and the dishonest fails to grasp which things are to be thought of as others’ and not your own. They are failures at being courageous, temperate and honest.

This idea also does not fit all vices equally well, but again there are deeper issues. Firstly, the general idea of failure will not do enough work here. Vicious people may turn out to have failed to learn something crucial, but vicious and bad people are not reasonably thought of as being blunderers, or dull and stupid. Becoming virtuous is not a course in which they get a failing grade. Some bad people are highly intelligent, and some vicious activities require intelligence, application and imagination. Vice is then not well characterized in terms simply of failure; it is a character trait, and a character trait can’t be made up just of failures.

The vicious are getting something important wrong which the virtuous get right. Vices, however, may display, and some may require, intelligence and foresight, in this being like the virtues. (The importance of this may differ among different vices. Dishonesty will not develop very far without considerable intelligence; laziness is clearly different.)\(^9\) The obvious way in which the vicious differ from the virtuous is that they get things wrong about value. The dishonest businesswoman thinks immediate gain is more important than running her

---

\(^8\) Aristotle’s remarks on vice are coloured by the fact that he focuses on intemperance as a paradigm vice. For this and other illuminating points see Terence Irwin’s interpretation of Aristotle on vice: ‘Vice and Reason,’ *Journal of Ethics* 5 (2001) 73-97.

\(^9\) Matthias Haase has suggested that he the same vices may come in an intelligent version and an unintelligent or stupid version, an idea worth further exploration.
business honestly. The coward thinks that immediate safety is more important than doing what will warn others. And so on. The vicious show, in a traditional formulation, ignorance of the good. They are mistaken in their priorities; they get wrong what matters in life, in ways that distinguish them from the virtuous.

An account of vice has to hold together, and do justice to, the following points about vice. It is a character trait, the state of the person who does the wrong thing and is characteristically (as a matter of character) unconflicted about it. It involves getting wrong matters of goodness and value which the virtuous get right. And (to varying degrees) it involves practical intelligence, something which also characterizes the virtuous. As a character trait, vice should be learnable in the way that virtue is learnable. But, as we have seen, the vicious do not have the same values as the virtuous, and so do not have the ethical aims that the virtuous aspire to as they learn to be virtuous. How then can the vicious learn to exercise the kind of practical intelligence that they often display? A cruel person may have as sensitive and practically intelligent a view of others’ psychology as a tactful person, but may use this to hurt rather than to support.

Since we do seem to learn to be cruel as well as learning to be tactful, let us return to the skill analogy for virtue despite the problems aired above. There has often been a temptation to think of vice in terms of skill in a different way. Perhaps, it is suggested, virtue and vice both have the structure of a skill, and differ merely in the ends that they have. The cruel and the tactful learn the same practical skill, but apply it to different ends, the cruel to hurting and the tactful to supporting. The cruel, after all, can adjust means to ends, be aware of the different kinds of issues salient in a situation and deliberate skillfully; these are aspects of skill that can be learned and exercised systematically. How does this work out in terms of the skill analogy?

Here are two examples, one from productive and one from performance skills. A painter becomes skilled in her art, and can produce great paintings. But since these do not sell well, she produces kitschy paintings which sell for large sums of money. A cricketer becomes a skilled batsman, but makes money from spot-fixing during matches. (Spot-fixing is taking money from betting syndicates for deliberately playing badly at a certain agreed point in a game; people bet on the results of parts of a game as well as the result of the whole.) These are clearly examples of having a skill and putting it to a mistaken end; they come from the point that ordinary skills, as has often been noticed, have the feature that the person who can paint well can also choose to paint badly, and the person who can play to win can also choose to play to lose. This use of skill to illuminate virtue and vice runs into problems with virtue, since it presents aspiring to be virtuous as detachable from the intelligent building-up of the virtuous disposition. But even

---

10 It is a form of cheating without necessarily affecting the result achieved by the whole team, and is thus hard to track. It clearly requires skill.
apart from that we can see that it is inadequate to explain the nature of vice, since nothing in it corresponds to the unconflicted nature of vice as opposed to akrasia. The painter and the cricketer might be perfectly happy with the use to which they are putting their skill, or they might be torn and internally conflicted. We have, then, not yet found the nature of vice.

However, if we return to the examples, they suggest a more subtle way in which we might find how we learn to be vicious. Instead of thinking of the painter and the cricketer as having the formed skill which they then apply to a wrong end, think of someone learning to do what they do – apply the skill instrumentally to achieving some other end. A painter learns to paint, but only to produce kitschy pictures which will sell well; a cricketer learns to play only in ways that focus on spot-fixing, and so learns to play badly as much as to play well. They haven’t learnt to do the same thing as the expert painter and cricketer, since their exercise of skill is limited, and they cannot yet (and perhaps never can) paint or play as well as the expert painter and player do. On the other hand, it’s clearly false that they haven’t learned to paint, or to play cricket; they may be making lots of money for painting, or playing, the way they do. What is going on?

In their case the acquisition of the skill involves a misdirection as it is learnt. The pupils learn to do what painters and players do, but not in the way that painters and players typically do – that is, aspiring to acquire the skill. A skill has standards intrinsic to it, which a pupil has to respect if she wishes to acquire the skill. This is most obvious when a pupil comes to see that she won’t succeed in reaching these because of lack of strength, aptitude and so on. Someone may, however, have the capacity to reach them, but instead of regarding doing this as a goal to be pursued for its own sake – for the sake of having the skill – regards reaching them merely as a means to something else, for example money. The painter will feel no need to practice beyond what she can do to produce money-making pictures; the cricketer will practice to lose as much as to win and so will not feel the need to develop the attitude of being part of a winning team. Their acquisition of the skill is limited, because it has been misdirected.

Making this distinction between reaching the standards of a skill, and learning the skill only for other purposes, does not imply that acquiring a skill should be ‘above’ thinking of uses to which it should be put. Of course a painter, and a cricketer, need to make money, and can reasonably hope that their work will be appreciated and famous. The important difference is between acquiring the skill, and then applying it to make money among other things, and misdirecting the acquisition of the skill towards those things as it is learnt. The standards of a skill both guide and constrain what the skilled person does with it. A skilled painter will aim to make money by exercising her skill, not by exercising her skill in a way that involves no respect for the standards of her art. A skilled cricketer, and other athletes, will aim to make money by exercising their skill to win, not by exercising it to lose. The pupils who have learnt the misdirected skill have in a sense not fully
acquired that skill, since they have no idea of what it is to exercise that skill in a non-instrumental way, a way in which the standards of the skill guide and constrain the way the skill is exercised. They have learned in a way that is indifferent to the standards of the skill, and to the extent that they are satisfied with this result they are mistaken about what it is to reach those standards. If, for example, the painter making her living from kitschy paintings feels thoroughly self-satisfied, she is mistaking what it is to be a painter, to strive to satisfy the standards of her skill. (This is not to say that we are in a position to blame her; skills have a different role in life than virtues do.) Similarly if the cricketer is thoroughly self-satisfied he has mistaken what it is to be a cricketer, as opposed to someone exploiting cricketing skills.

Does this illuminate the difference between learning to be virtuous and learning to be vicious? Take a young person whose father has run a business honestly for many years and been successful. The son learns the business from him, but at every point his emphasis is on how much money is made. Focussing on this he notices not only ways in which more money could be made in the honest ways assumed hitherto, but also dishonest ways in which money could be made, or ways that never occurred to his father. Going for getting the most money he proceeds to cheat unsuspecting clients. The difference between him and his father is not that his father was uninterested in money; it’s that for the father the pursuit of money was guided and constrained by honesty, whereas for the son it is not. The son thus does not aspire to be honest as his father did; he holds a value which honest people see to be mistaken. However, as far as the running of the business goes, the son has learned from his father the kinds of actions an honest person would have learned (by way of business expertise, marketing skills and so on). It was while he was learning them, not after, that the son learned to make money in ways not guided or restrained by honesty. In learning to run the business in ways focussed only on making money he is like the cricketer who learns to play in a way focussed on making money by spot fixing. Both of them are missing something important about the whole endeavor they are occupied in, but they can exercise cleverness in the ways they act.

Here as often we can learn from Aristotle, who contrasts practical intelligence with another ability, which he calls cleverness (deinotes). This is involved in practical wisdom, since practically wise people not only have the right aims but are intelligent in the ways they achieve them, but it can also achieve ends in a way detached from aspiring to the right ends. We would all, of course, prefer to be intelligent rather than to be slow or dull; there is nothing to be rejected about cleverness in itself. But someone who is merely clever can end up pursuing money,

---

11 I add this point to remove the idea that the son might be reacting against deficiencies in what his father has been doing.

12 Nicomachean Ethics, 1144a20-b1.
success, status or other such aims in ways that are detached from concern with virtue. This is most obvious with money, but it is easy to think of other examples: a writer who achieves success by writings which make hurtful revelations about other people, an academic who exploits the work of others without acknowledgement, and so on. There are of course people who do this sort of thing once, or occasionally, and feel conflicted about it. However, we are thinking of people who act like this because they have learnt to pursue their aims in ways that are detached from the ethical values which guide and constrain the way the virtuous pursue them. And this seems promising as an account of vice. These are people who are not conflicted – they act without internal struggle because they have learnt to act this way and have got used to feeling positively about it. They do not share the aims of the virtuous: they are not aiming to be honest, or compassionate. At the same time, they are not aiming to be dishonest or mean. They are not aiming to be any kind of person at all. They are just aiming at having money, status or whatever, and they have learned to do this in the same ways that the virtuous have, and may be good at it. In their case they have limited their lessons to ways of achieving things like money and status, ignoring the considerations which guide and constrain the virtuous as they achieve the same aims, considerations which enable the virtuous to aspire to become honest, brave and so on.

This account of vice meets the desiderata mentioned above, that vice is an unconflicted character trait, that the vicious get wrong about value what the virtuous get right, and that the vicious can nevertheless display a kind of practical intelligence in exercising the vices. The intelligence in practical matters which the vicious display may overlap with that of the virtuous, as with the cruel person who is as good as reading people’s reactions as the tactful person. But the virtuous have developed this practical intelligence while, and in the course of, learning to aspire to be a certain kind of person – tactful, honest and so on – while the vicious have learned it in a way detached from these aims. There is truth in the idea that the intelligence of the vicious is limited, and that because of this narrowness of scope and vision it may end up undermining, rather than furthering, the person’s goals. The ‘smart’ vicious person may trip herself up because of failing to comprehend what is beyond her limited range. However, this idea, which has often served to make virtuous people feel better about worldly failure, should not be pushed too hard; some of the vicious do die prosperous and successful.

The present account of vice also accounts for the point that the cowardly are not aspiring to be cowardly; vice turns out not to be an aspiration to be vicious, but to go for the same everyday things in life that the virtuous go for – money, success and so on – but in a way detached from what the virtuous take to be important. What is this ‘detachment’, and how does it relate to the traditional thought mentioned above, that the vicious are ignorant of the good? Here it is helpful to bear in mind that for vice to be a character trait which is relevantly like
Virtue, Skill and Vice

virtue in being unconflicted, it has to have been developed in ways which involve not only the building up of beliefs but the formation of our emotional and affective side. The virtuous person learns that it is right to do the fair thing, and learns through experience and teaching not to feel resentful about a fair distribution which disadvantages him. The vicious person is the person who does feel resentful, and who does not come to see the point of respecting fairness. She develops in a way that tries to evade fairness or to exploit occasions where fairness is demanded, feeling resentful when she is forced to be fair and pleased when she can get away with being unfair to her own advantage. We may say that she is ignorant or mistaken about fairness; she gets it wrong about what fairness is and what its value is, where the virtuous gets it right. Or we may say that she is indifferent to these things.\textsuperscript{13} With someone who is vicious (rather than akratic) these are not alternatives; both are right. The virtuous person comes to feel less and less impeded in virtuous action by contrary feelings or inclinations, as he comes to be more and more confident that the virtuous aim is worth aiming for. Similarly the vicious person comes to feel more and more indifferent to considerations of virtue as he becomes more and more confident that the virtuous aim is silly, or for losers, and the like. Ignorance of the good, and insensitivity to it, are mutually reinforcing, just as are true beliefs about the good, and increasing motivation to pursue it. In one case we may be struck by someone’s indifference to, say, suffering, and in another case it is the person’s misconception of what is worth doing that we notice. But indifference and ignorance of value go together; both are involved in a developed vice, for the same reason that practical intelligence and positive motivation are in the case of virtue.\textsuperscript{14}

This account of vice also goes some way towards accounting for another point. Since virtues all involve practical intelligence, which is not compartmentalized by the different areas it deals with, there is a tendency for virtues to be integrated or unified by the exercise of practical intelligence in them. This does not go over to vices, however. The practical intelligence in a vice is, as we have seen, cleverness, focussed on getting various ordinary aims in life, such as money and status, in indifference to considerations of virtue, which might restrain the cleverness in some directions, and guide it in others. The focus of cleverness on one such aim, such as money, has no tendency in itself to unify its pursuit with others in the person’s life, such as having a family, or being famous. This myopia of cleverness means that a vice can develop without integration from other character traits (except those whose exercise is instrumentally required to achieve the end of making money). Insofar as the person’s life is unified by the pursuit of a single end

\textsuperscript{13} For interesting remarks about indifference in the context of vice see Neera Badhwar in chapters 6 and 7 of her \textit{Well-Being}, Oxford – New York, Oxford University Press, 2014.

\textsuperscript{14} In this paper I do not go into the details of how the vicious’ characteristic way of getting it wrong about value involves ignorance and mistake; I hope to extend work on the vices in this direction.
like money, this pursuit produces a unification which is generally regarded, in the wider culture and not merely in philosophy, as artificial and forced. Lives of the vicious, in fact and fiction, are often represented as compartmentalized in this kind of way; the all-consuming pursuit of money in particular is generally recognized as producing a life which is unsatisfactory in its structure.

Further, the disunified way in which vices can develop in the person means that the life of the vicious can be internally conflicted in ways that have nothing to do with the internal conflicts of the enkratic and the akratic. Cowardice will tend to interfere with cruelty, arrogance with ambition, stinginess with greed. Moreover, vices are unlikely to develop on any scale without the assistance of two virtues which have been labelled ‘executive’ virtues, courage and temperance. If you are unable to stand up for your projects or to resist short-term gratifications, you are unlikely to achieve anything either virtuous or vicious on a large scale. To the extent that vices do develop on a large scale, then, there is likely to be an internal conflict with the recognition of the value of resisting danger and temptation.15

I have argued that we can give an account of vice which shows how it is similar to virtue in being an unconflicted character trait, which is learned in the same way as a virtue is learned, but in a way which is misdirected because indifferent to and mistaken about the kind of person to aspire to be as we learn to exercise our practical intelligence. I’ve also suggested that the account fits well with some aspects of vice. There is more to be said about vice, both within an account of virtue of an Aristotelian kind and beyond it to other types. I have stressed the skill analogy in this paper because it is so important to an account of virtue which is Aristotelian in structure, and because, at least at first, it appears to offer problems for a corresponding account of vice. I hope to have shown that the development of a practical skill is as useful for giving an account of vice as it is for giving an account of virtue.16

15 It may be objected that executive virtues are just traits that can be harnessed to any end, and so not in themselves virtues. However, insofar as courage and temperance import the value of being a brave and self-controlled person they do import the idea of a virtue trait as worth having in itself, not just instrumentally to achieving some end.

16 I am grateful to the members of my Virtue Ethics seminar in Spring 2015, and in particular to Phoebe Chan, Vincent Colainni and Greg Robson. I am also grateful to the audience for a very different first version of this paper at the Workshop on Virtue and Skill at the Centre for the Study of Mind in Nature, University of Oslo, June 1-2, 2015, and especially to Christel Fricke for organizing it with me as well as for comments, and to Matthias Haase, Matt Stichter, Peter Railton, Will Small, Darcia Narvaez and Sascha Settegast.
Feminine Virtues or Feminist Virtues?
The Debate on Care Ethics Revisited

Caterina Botti
Sapienza Università di Roma
Dipartimento di Filosofia
caterina.botti@uniroma1.it

ABSTRACT
In this paper, I would like to offer a reinterpretation of care ethics both as a feminist perspective on moral reflection and as an interesting remapping of the moral domain in itself. The feminist nature of care ethics can be understood in different terms. The leading idea of this paper is that the effort of distinguishing these terms may have important implications for a more structured philosophical understanding of our account of care ethics (and therefore of ethics). As I hope will become clear in what follows, this can be thought of in terms of distinguishing – at least metaphorically, if not technically – between considering care ethics as an ethics which puts at its centre (more traditional) “feminine virtues” or alternatively (some new) “feminist virtues”.

KEYWORDS
Care, feminism, sense of one’s own limits.

Introduction
Care ethics is nowadays considered one of the most thought provoking contributions of feminist thought to moral reflection and an interesting moral paradigm in itself. In the wake of Carol Gilligan’s first attempt to envisage an alternative – conceived in terms of responsible care for relationships – to the universalist, rationalistic, impartialist and individualistic moral paradigms (which characterise – to say it with Anscombe – “modern moral philosophy”), a significant literature has emerged. In fact Gilligan’s suggestions, but also those made in the same period by Sarah Ruddick and Nel Noddings, are considered insightful by many moral philosophers and have been further elaborated along different lines of development.

As is well known, a rich, ongoing debate among feminists and among philosophers who are, in their turn, critical of universalist and impartialist moral

---

1 I would like to thank Catherine Bearfield for her help with the nuances of English language and her thoughtful advice.
conceptions, has developed, focusing on the possibilities for understanding, refining and using those initial insights, giving them a more definite philosophical structure and considering the breadth of their implications for restructuring the field of moral philosophy.

Care ethics has in fact been developed as a distinctive moral paradigm (for instance by Joan Tronto and Virginia Held), but also many critics of universalist and impartialist moral conceptions have shown an interest in this proposal and in the consonances to be found with forms of moral sentimentalism and virtue ethics (see for instance Annette Baier’s and Michael Slote’s work), with moral particularism (as in the work of Lawrence Blum), or – more recently – with moral perfectionism and ordinary language ethics (as in Sandra Laugier’s writings). These encounters have produced fertile dialogues.

This wide debate notwithstanding, I think there is still room for offering another contribution on care ethics, focusing on some of its salient aspects, but also its limits. In this paper, therefore, I would like to sketch the main lines of such a contribution, aimed at offering a reinterpretation of care ethics both as a feminist perspective on moral reflection and as an interesting remapping of the moral domain in itself.

It should be made clear from the outset however, that the main object of this paper is not that of drawing a comparison between the above-mentioned different lines of research, or to argue in favour of one or the other of this vast array of philosophical positions. Rather, my attempt is to grasp more clearly, from a particular point of view, some elements which are relevant to a fuller understanding of care ethics, and therefore of this wider philosophical debate. This viewpoint will involve going back to Gilligan’s initial insights and considering them in the light of some more recent feminist considerations. Accordingly, while engaging in dialogue with certain well known (mostly sentimentalist) interpretations of care ethics and maintaining a rather superficial reference to the moral language of virtues, my main effort will be that of offering an illustration of some recent developments in feminist thought, which I find interesting both in themselves and as a contribution to a fuller understanding of care ethics and thus of ethics as such.

The feminist nature of care ethics, though often invoked, is not straightforward. It can in fact be understood in different terms, and the leading idea of this paper is that the effort of distinguishing these terms may have important implications for a more structured philosophical understanding of such an account of ethics (also in relation to other attempts at giving shape to non universalist and non impartialist accounts of ethics). As I hope will become clear in what follows, this can be thought of in terms of distinguishing – at least metaphorically, if not technically – between considering care ethics as an ethics which puts at its centre (more traditional) “feminine virtues” or alternatively (some new) “feminist virtues”.
It should be made clear that the idea of considering whether and to what extent care ethics and feminist thought are intertwined does not proceed from an ideological standpoint: there is no assumption that one should be feminist, nor any request of coherence for philosophers who declare themselves feminists, as many protagonists of the care ethics debate do. It proceeds instead from my opinion that feminist thought, in its development, offers some important considerations to be taken into account in the kind of reflection on ethics that care ethicists undertake (of value also in a wider philosophical debate on ethics). As will be argued, I believe that the evolution of feminist thought offers some important insights, not only with regard to the problem of women’s oppression but also with the need for a reconsideration of the human condition, of subjectivity and of morality, all of which are relevant to mapping the moral domain. A deeper analysis of what care can come to mean in the light of these feminist considerations (on subjectivity, humanity, morality or epistemology), may therefore be of some interest. The core issue of this paper is, thus, to assess whether care ethics, in the specific understanding I will be trying to carve out, is able to accommodate some of these considerations. At the same time, I will be arguing in favour of the value of these kinds of considerations in themselves.

Of course, the idea of focusing on this parallelism was inspired by the declared feminist nature of many reflections on care ethics, and by the ongoing debate that has developed in order to characterise this. In fact, as will be described in the following pages, there are at least two different ways to consider the feminist meaning of care ethics. One interpretation is that this approach to ethics is a way of doing justice to women: recognising in them a specific moral “voice”, based on the particularity of women’s experiences or on specific feminine endowments (in this sense we can consider care ethics as envisaging a specific form of “feminine ethics”, based on particular “feminine virtues”). The other interpretation is to consider care ethics as an account of ethics which is able to deal with the particularity, difference and concreteness of all human beings. This is one of the themes feminists elaborate by reflecting on women’s experience in order to give shape to a more adequate account of ethics for all. An account of ethics which, unlike universalist abstract accounts, is able to consider the importance, but also the difficulties, of caring for others in their differences. My contention is that clarification of the sense in which care ethics is feminist – and I will argue it is feminist in this second sense – is a way to clarify care ethics as such.

In what follows I would like not only to distinguish these two interpretations from each other but also to elaborate on the second, in an attempt to delineate what is required in order to consider a (feminist) care ethics not as the elaboration of a peculiar feminine endowment but rather as a form of discourse on ethics which is able to accommodate those feminist contributions in relation to all human beings, and, as such, as important contributions to moral reflection.
In the final section, I will therefore try to develop a few suggestions, but also to leave open certain questions which may hopefully be of interest both to those concerned primarily with the debate on care ethics and to others who are attempting to map the moral domain in ways that offer alternatives to universalist, rationalist, and impartialist approaches.

1. Circumscribing the problem: the core contents of care ethics and the feminist context

In order to characterize the nucleus of care ethics I will go back to Gilligan’s work, since I see her work as offering the raw material on which care ethics as a distinctive moral approach has been developed, but also as offering some specific clues (not always maintained in subsequent developments), as to what renders such an approach so interesting. I will try to show how Gilligan’s claims can be read as a specific kind of feminist claim, contextualising them within a (personal) reconstruction of the development of feminist thought.

It was 1982 when Carol Gilligan published *In a Different Voice*, the book containing the results of her work which, starting from empirical psychological studies on the development of moral judgement in adolescents of both sexes, launched the idea of a different voice in ethics and started to configure it theoretically.

As she says at the beginning of her book, it was in years of “listening to people talking about morality and about themselves”, that she came to hear a distinction between: “two ways of speaking about moral problems, two modes of describing the relationship between other and self”, and it was in trying to account for this second voice, that of girls and young women, or rather, in trying to solve what she considered the puzzle of female morality, that she came to think of “care” or, more precisely, of “responsible care” as a crucial notion for a different conception of morality.

As is well known, it was mainly with reference to the results of the work done and ideated by Lawrence Kohlberg, concerning the moral development of adolescents, that she started her research. This aimed at considering the difference caught in the female voice, and at resisting the verdict of an inferior or defective moral development of girls and young women which resulted from Kohlberg’s studies (a verdict which – as Gilligan argues in her book – was in line with the widespread representation of women’s development and role in the psychological tradition, but which is – one can add – also in line with the commonsensical

---


3 Gilligan, In a Different Voice, p. 1.
representation of women in the history of Western philosophy and Western culture).  

Resisting the idea of female minority, Gilligan tried, in fact – with what can be defined as a feminist gesture – to characterise the answers of girls and young women not as a deficient version of the male ones, but as giving shape to a different “voice”, with specific contents, and to recognize this different “female” moral development as valuable in thinking the human condition and morality. 

A crucial tenet of her work was the idea that it was indeed the representation of morality and humanity implicit in Kohlberg’s research, and not the girls’ and young women’s answers to his questions (or women as such), which was limited (this seems rather obvious to us now, but it was revolutionary at the time). More specifically, what was limited was Kohlberg’s scale itself. This scale is representative of the long history of the characterisation of human subjectivity in terms of isolation and separation, and of morality in terms of abstract and impersonal rules able to put those separate and sovereign selves in relation to each other, and also able to give an (impartial, objective) order to their moral determinations: rules and norms which are obtained by detaching oneself from one’s own particularity and inclinations and those of others, thus adopting an impersonal point of view from which to fix the representation of each moral problem as that of a relationship between “generalised others”, gaining at the same time the status of full moral agent. Gilligan’s idea was instead that what was worth inquiring into was precisely what rendered it difficult for women to give an account of their own experience that fitted these terms. Her thesis was therefore

---


5 As is well known, Kohlberg’s studies proceeded by measuring the moral development of adolescents on a scale of six stages, each characterised by a different conceptualization of justification of moral judgements, ranging from an initial stage of egotism, through one of heteronymous adherence to conventional norms, to a final stage of post-conventional universalist moral thinking.

that those characteristics which were considered as factors limiting the moral
capacity of women, such as, for instance, “their care for and sensitivity to the
needs of others” or the “emphasis on connections rather than separation”,
(characteristics she herself noted in her interviews), should instead be brought to
light and considered in positive terms. And this in order to achieve, not only, or
not primarily, a better representation of women and their experience, but rather a
more accurate view of morality and human relationships.

The disparity between women’s experience and the representation of human
development noted throughout the psychological literature, has generally been
seen to signify a problem in women’s development. Instead, the failure of
women to fit existing models of human growth may point to a problem in the
representation, a limitation in the conception of human condition, an omission
of certain truths about life.³

Here the “truths about life” worth taking into account are those concerning
the relational nature of each life and of life itself. These truths endorse a moral
standpoint which casts problems and their solutions in terms of responsible care
for relationships, an attitude acquired through solicitude and sensitiveness and
from a consideration of humans as deeply interdependent.

The truth of relationship, however, returns in the rediscovery of connection, in
the realization that self and other are interdependent and that life, however
valuable in itself, can only be sustained by care in relationships.⁹

In general terms, the proposed model can be described as characterising moral
judgement and practice as emerging from connection, instead of from detachment
(a point widely underlined by sentimentalist readers), or as binding interconnected
vulnerable selves instead of separate sovereign selves (an opposition between
sovereignty and vulnerability which opens up also to various different readings, as
for instance to some Aristotelian approaches, or to Wittgensteinian ones, or to
those connected to radical feminist stances, as we will see). A further crucial
element is the importance given to the particularity and difference of each life and
each context. This is relevant if the main concern of morality is thought of in
terms of maintaining the connection which ties us together, and of caring for our
and the other’s wellbeing or flourishing, from within this relational framework.

From different passages of Gilligan’s book, care ethics emerges therefore as a
moral model which puts at its centre the agent’s capacity to be attentive, caring
and responsive in relation to the needs of others, in their concreteness,
particularity and relational nature, and which defines moral responsibility (or the

³ Gilligan, *In a Different Voice*, pp. 18-19.
⁹ Gilligan, *In a Different Voice*, p. 127.
responsible character)\textsuperscript{10} in terms of the development and practice of these capacities.

Before coming in the following section to a more detailed account of her proposal, let me say a few words on the feminist characterization of Gilligan’s gesture.

Gilligan’s move can be considered as feminist insofar as she rejects and criticises a stereotyped verdict concerning women’s characteristics, one which is congruent with the specific and inferior position assigned to women in “men’s life cycle”\textsuperscript{11}, a typical judgement of what can be defined as the patriarchal system. In a very broad understanding, a position is considered feminist as long as it is critical toward the idea that women, \textit{qua} women, are different and inferior creatures (a critique eventually considered as leading to a more general discourse on difference/s).

But of course this feminist core issue can be, and has been, developed in different ways, and in this light Gilligan’s move can be understood as a gesture of a precise and specific feminist kind.

The aim of claiming that women are just as human as men can be achieved, in fact, in different ways. I will make reference to at least three different strategies. On the one side it can and has been sustained, that woman are substantially equal to men (at least concerning the relevant human capacities, as for instance reason), and therefore not inferior, and that if they seem different and inferior it is only because of unjust social conditions that have limited their opportunities. Women’s minority is therefore only the result of centuries of discrimination. If Gilligan’s feminism were of this kind she would have sustained that girls and young women scored low on Kohlberg’s scale only because they had been socialised in a discriminating environment; had they had the same opportunities and socialization of boys they would have achieved the same results.

But this is not what she argues.

Gilligan, in fact, following a more radical understanding of feminism, maintains that it is the scale, and the ideal of morality it enforces, that are misplaced, since they do not represent human experience and morality in their complexity: they do not, for example, take into account women’s experience or women’s ways of expressing themselves. She recognises a difference but non an inferiority in women’s development, and hence the necessity of a reconfiguration of morality. In these terms her gesture can be defined as of a specific feminist kind. Broadly speaking one can say that this second kind of feminism maintains that women are different from men, but not inferior, and that the difference they

\textsuperscript{10} Although Gilligan speaks more often of a “morality of responsibility”, I think that the reference to a “responsible character” is not misplaced in this context.

\textsuperscript{11} See Gilligan, \textit{In a Different Voice}, chap. 1, entitled “Woman’s Place in Man’s Life Cycle”.

113
represent is worth inquiring into. Valuing women’s difference opens up to a reconsideration of what it means to be human, and thus to an enrichment for all.

For the sake of my argument it is worth mentioning that although considered forceful, this kind of feminist strategy has been made the object of different criticisms.

The main criticism raised against it is that it seems to imply the possibility of offering a common definition of women or of women’s experience, which is precisely what has hitherto been contested. It is alleged that claiming the value of women’s difference, although a powerful gesture in destabilising the supposed neutrality and the false universality of patriarchal discourse, could in its turn become similarly oppressive, as it seems to convey the possibility of a universal description of women, thereby duplicating the monological structure of patriarchy.

The debate has been broad-ranging, not only on the nature of the supposed difference – whether it is essential or socially constructed – but also (and more interestingly) the definition of this difference itself has been put under pressure. Is “women’s difference” that which was defined by the patriarchal system, which should just be freed from its negative evaluation, or should it be completely redefined and explored? And could such a redefinition be seen as embracing all women? If so, what happens to the other relevant axes of difference which operate among women and more generally among humans (typically, race, class, sexual orientation, and so forth)? In the light of these problems, a third feminist strategy has emerged which can be thought of as taking on and attempting to offer an answer to these questions, thus opening up the problem of differences, and not only of women’s difference, for ethics.

This is an important issue for the kind of analysis I would like to offer, namely an attempt to consider to what extent Gilligan’s ideas, and care ethics as such, could be collocated in this latter framework. That is to say, whether care ethics should be considered as emerging from the second kind of feminist strategy, or whether it can be seen as a way to tackle these broader questions.

Let me go back to Gilligan: in assuming, characterising and valuing women’s difference and morality it has been said, for example, that Gilligan disregards differences between women and other relevant differences among human beings (social class, whether one is at the margin or at the centre of the social system, etc.). She is thereby accused of ascribing a common nature or common features to women, features which besides mirror those ascribed to them traditionally (e.g. the traditional feminine virtues), albeit in a positive light. This is a typical issue put forward by the so called power-centred feminists, but also by recent (and in my view more interesting) developments of feminist thought. I will argue, in what follows, that these critiques are disputable, with regard to Gilligan’s work or to possible interpretations of it, at least to a certain extent. But I will also argue that to situate care ethics within a different feminist framework implies (and suggests) a
very articulated set of considerations, which have not always been taken into account in the ensuing debate.

In this respect, although on the nature of the “different voice” put forward in her work Gilligan shows a sort of ambiguity, it must be said that in the end she tries to take a definite position. In the end she is explicit in claiming that while it is important to recognise that women have a voice, as a way of doing justice to them, what she is mainly interested in is that in acknowledging this different voice we are recognising the existence of differences among human voices, or – as she has recently argued – the existence of a more humane voice, which is only contingently more easy to individuate in (young) women.12

In fact, already in *A Different Voice* she claims to be interested more in the “theme” than in the “gender” of the different voice she was hearing, and she acknowledges that the association between women and the particular vision of moral development she traced there was empirical and not absolute.13 She claims not to be interested in defining women’s difference as such, but in the more general aim of giving an account of human differences that is not couched in terms of “better or worse”.14 This is intended to be the main result of her work: a criticism of the supposed neutrality of the scales commonly used to measure human (moral) development, as part of a more general critical stance on the assumed neutrality of the categories of human thought, knowledge and language (and in this respect she talks of the “relativity” of the “categories of knowledge as human constructions”, while, as I will clarify, I prefer to talk of the “instability of categories”).15 This opens up the possibility of a positive use of this criticism. Thus she claims:

> My interest lies in the interaction of experience and thought in different voices and the dialogue they give rise to, in the way we listen to ourselves and to others, in the stories we tell about our lives.16

With this kind of claim, in my opinion, she clearly commits herself to a particular radical stance in relation to the abovementioned debate within feminism, a position which is compatible with more recent forms of feminist thought, those offering an understanding of feminism not as limited to the problem of mending women’s oppression or affirming women’s difference, but as critical of patriarchy as a more generalised form of oppressive structure, as a system of power which disciplines hierarchically not only women, but both men and women, human beings in general, on the basis of supposed differences in status, whilst at the same time obscuring the value of differences that matter, that is of different

---

13 Gilligan, In a Different Voice, p. 2.
14 Gilligan, In a Different Voice, p. 14
16 Gilligan, In a Different Voice, p. 2.
voices. Gilligan seems therefore to regard care ethics as an ethics which deals with humankind in this differentiated sense.

This is clearer in her later writings (including the letter to readers of the second edition of *In a Different Voice*), for instance where she claims that “difference [is] a marker of human condition and not a problem to be solved”\(^\text{17}\) and that it is in this light that we should recast our discourses on separation and connection with regard to individuals and relationships, rules and responsibilities.

We may take it then that Gilligan is not so much interested in claiming that women have a different essence or nature as in claiming that what she finds in their development is important in order to recognise women’s possibility to express themselves and at the same time in order to reconsider a morality for both men and women, a morality which in turn is more open to differences. If this is the case, we can consider her argument as being made in three steps.

In my view, Gilligan’s contention that the scales employed to measure human morality are not neutral, and that humanity is more complex and differentiated than has been assumed in uncritical representations, is a specific feminist move which can be described in the following way: as the idea that (1) recognising women’s difference will open up to (2) a recognition of human differences which will both allow women to gain a voice and allow for a more interesting representation of human experience, through (3) a wider reconfiguration of morality and humanity. This reconfiguration will mainly be concerned with the recognition of the frailty and difference of human experiences and of the ways in which we express them. As we will see, in fact, care ethics entails care for humankind in its difference and frailty and is in itself a frail voice.

Yet this complexity of levels is not easy to express and many ambiguities remain to be clarified, but it is important – and this is my point – to vindicate the complexity of the feminist move described above, in its different aspects. Often, however – particularly in the wider debate on care ethics – this has not happened.

It should be noted, as a final point in this regard, that even in Gilligan’s writings the treatment of this complexity is not always clearly articulated. In fact, while Gilligan seems to consider the difference she recognises in previously unheard women’s voices as a new voice, only empirically womanly, at times she appears to describe it precisely as the difference that has been traditionally ascribed to women, as if it were sufficient to rehabilitate terms attributed to women, such as those of their “goodness and virtue”, to describe the most humane voice.

Gilligan notes in different places that the very same traits of “sensitivity” and “solicitude”, which she places at the centre of a different moral paradigm for humans, are the same traits which were traditionally considered as, on the one hand, characterising the “goodness of women” or as “feminine virtues” and, on the

other, as preventing women from achieving full moral maturity.\textsuperscript{18} As she says, this is a paradox that she wants to disentangle, bringing those traits fully into the light and putting them at the centre of a new vision of moral concern, in order to free women from their marginal position and, at the same time, to gain a different understanding of morality.

In this case, Gilligan’s proposal seems to involve two moves. First she recognises a moral worth to the concerns which emerged from the girls and young women she interviewed in her studies, those which coincide with the traditionally undervalued “feminine virtues”. Then, as she claims in the same book and more clearly in the following works, she recognises these “virtues” as human rather than specifically feminine, thus recognising their role in a more comprehensive account of morality. What is lacking from this description is an analysis of what we can learn from this transposition, in terms of a broader reconfiguration of morality, that is to say, in terms of the instability of categories and of the frailty of human descriptions or voices. This is what is made explicit, instead, in the three-steps framework suggested above where a more general reconfiguration of the field of morality together with a more complex representation of human condition is implied.

My point here is to ask whether the retrospective description that Gilligan gives of her own work is accurate, or whether in fact what she is proposing can be considered as something rather more complex than just moving certain questions from the margin to centre. In this movement, in fact, something is gained (or better lost): the very idea of a universal and neutral truth is lost, but the idea of a more unstable centre of morality is gained. The implication is therefore that the entire moral landscape should be reconfigured or, more radically, the notion emerges that it could never be configured once and for all, and that only in recognition of this can we try to account for human lives in all their differences.

Is there any connection between this kind of epistemological standpoint and a morality of relational sensitivity? I would try to suggest something similar.

In other words, going back to Gilligan’s claims concerning the “goodness of women” or the “feminine virtues”, one might wonder whether recognising their full value is a sufficient move from a feminist point of view, or whether something must be added in order to gain the refined understanding of care ethics as a post-patriarchal morality for both men and women along the lines we have just described, i.e. as an ethics for humans in all their particularities and differences.

To argue that this is the case, it is necessary first to consider more in details how Gilligan describes the different moral voice she has listened and envisaged.

\textsuperscript{18} See for instance Gilligan, \textit{In a Different Voice}, p. 18 but there are similar claims on many pages of this book as well as in many passages from her, \textit{Joining the Resistance}.\textsuperscript{18}
2. Interdependence and responsible care

What Gilligan finds in analysing girls’ and young women’s reactions to moral dilemmas can be described as follows: girls and (young) women are generally more aware of the role that relationships play in human life; they trust the human capacity to maintain and restore the web of interdependence which sustains human life, so that everybody and – she says – “life itself” may flourish. She claims that attention to interconnection leads girls and women to be more attentive and sensitive to the particular needs of others and to feel responsible for them, rather than abstract rules or principles of fairness or rights, as well as to the relational dimension which characterises the practice of this responsibility and not only its genesis. Women, or at least the women she has interviewed, once they are free to express themselves and to find their own voice, see moral problems in a different way, and find different solutions, which call for different concepts or different abilities. Gilligan therefore claims that if we do listen to women’s voice we will be able to envisage a different morality.

It is worth mentioning that as Gilligan notes, envisaging such a different moral voice is not an easy task, because what makes the women’s voice relevant for morality results also in its weakness: that is, its recognition of the frailty of the human condition and of the web of relationships which nourishes it, and also of the relevance of the particularity of contexts, thus it is always a tentative sound that is produced. “It depends” is the common female answer in the case of moral dilemma,19 and this can be considered either as a confused answer or as a meaningful one, depending on what lies behind it. Gilligan attempts, out of her own attentive and caring listening, to envisage a framework which makes the power of such an answer visible, namely the attention and care it shows.

Thus, in the famous example of the two children, Jake and Amy, who are asked to deliberate on the dilemma of a man, Heinz, who has to decide whether or not to steal a drug he cannot afford in order to save his wife’s life, while the boy reduces the dilemma to a conflict of rights (to life and to property) and solves it easily by adopting a detached and impersonal point of view according to which life has a logical priority over property (therefore not only should Heinz steal the drug, but if arrested the judge should reason according to the same logic and not condemn him as a thief), Amy seems at a loss and confused. She asks a lot of questions considered unnecessary according to the standard procedure adopted in such tests. It is only out of an attentive listening to her words, and by conducting the interview differently, that her voice emerges positively, and that Gilligan can claim that the girl sees the problem in a different way.

For Amy, in fact, the dilemma arises – Gilligan notes – not from the conflict between the druggist’s rights and the rights of Heinz wife, but from the druggist’s

19 Gilligan, Letter to Readers, p. xxi; In a Different Voice, p. 38.
failure to respond to Heinz and to his wife’s needs. Furthermore, it appears that this is not so much a problem that has to do with the druggist himself, as with the relationships between him and Heinz. The problem lies in a lack of communication, a communication which could render the druggist able to see the consequences of his refusal.

In Amy’s vision, therefore, the problem and the solution lay in the relationship between particular persons in a particular context. In this sense, Gilligan claims that here we are faced with a different moral language and a different moral logic: the language of responsibility and the psychological logic of relationships.

Yet the world she [Amy] knows is a different world from that refracted in Kohlberg’s construction of Heinz’s dilemma. Her world is a world of relationships and psychological truths where an awareness of the connection among people gives rise to a recognition of responsibility for one another, a perception of the need of response. Seen in this light, her understanding of morality as arising from the recognition of relationship, her belief in communication as the mode of conflict resolutions, and her conviction that the solution to the dilemma will follow from its compelling representation seem far from naive or cognitively immature. Instead Amy’s judgments contains the insights central to an ethic of care, just as Jake’s judgments reflects the logic of the justice approach.20

I think that this example is crucial in order to gain an insightful account of Gilligan’s care ethics. For this reason I will need to go into it more deeply.

Gilligan continues offering the following comments:

Her [Amy’s] incipient awareness of the “method of truth”, the central tenet of non violent conflict resolution, and her belief in the restorative activity of care, lead her to see the actors of the dilemma arrayed not as opponents in a context of rights but as members of a network of relationships on whose continuation they all depend. Consequently her solution to the dilemma lies in activating the network by communication, securing the inclusion of the wife by strengthening rather then severing connections.21

It is worth noticing that for Amy not only should Heinz’s wife be included, but that in this process the druggist is not left aside either, as he would be in the alternative case in which his rights should give way to those of the sick woman, considered in an impersonal way as having greater force (as, for instance, in Dworkin’s understanding of rights as trumps).

In fact, in the end Amy’s answer is that stealing is not the best choice Heinz could opt for, while communicating would allow the druggist to have a wider and

---

20 Gilligan, In a Different Voice, p. 30
21 Gilligan, In a Different Voice, pp. 30-31.
more compelling representation of the situation. The problem thus cannot be solved by a single person, it requires instead the maintenance of a space for a relationship in which the vulnerabilities of all the actors are considered, starting from that of the sick woman, but without obliterating the others. The moral responsibility of everybody is therefore in maintaining this space of communication, engaging in relationship and committing oneself to care for the vulnerable.

The ideal of care is thus an activity of relationship, of seeing and responding to need, taking care of the world by sustaining the web of connection so that no one is left alone.\(^\text{22}\)

This is in the end the idea of moral maturity – with its different language and logic – that Gilligan proposes, not only as a feminine one, but as an important ideal which should be recognized and valued, as indicating the way to develop an alternative and valuable pattern of moral development.

Gilligan therefore puts forward a vision of moral maturity in contrast with those based on detachment, impersonality, impartiality and universality, those which rely on norms that are abstract and acquired through reason and which purport to breach the gap between ourselves and others. Girls, not seeing the gap, are rather concerned with the responsibility emerging from relationships, and with the need to respond from within the same relational tissue, working on those same relationships from within, in their concrete particularity, in order to improve them, to render them adequate for the survival and flourishing of their participants. On this view, moral maturity does not consist, therefore, in the capacity to abstract or detach oneself from the particular context of a moral dilemma, from one’s own role in it, or in reducing it to the issue of balancing conflicting interests on the basis of impersonal, abstract and formal procedures, considered as universally valid (abstract procedures or principles once gained are superimposed on the circumstances of everyday existence). Rather it consists in the development of sensitivity and solicitude, of attentiveness and interest in the needs of the other in their concrete particularity, in the awareness that it is possible to answer to these needs only by entering the same dimension of concreteness, communication and relationality.

Finally, it could be said that, in order to grasp the core content of an ethics of care, Gilligan’s attitude is in itself an appropriate starting point: her gesture of entering into a direct relationship with the girls, in order to listen to them, abandoning pre-established interview procedures, is in fact a clear example of the

\(^{22}\) Gilligan, *In a Different Voice* p. 62. As we will see this reference to taking care of the world is present also in Joan Tronto’s definition of care, although she seems to be less interested in the relational aspects of it. See Joan Tronto, *Moral Boundaries. A Political Argument for an Ethic of Care.* (New York: Routledge, 1993) and see infra in this paper for a discussion of her position.
putting into practice of precisely what she went on to theorise as a result of her deep listening to those girls. Hers is a caring attitude, the ethical attitude in the new horizon she tries to depict.

Let us now turn to certain other aspects of Gilligan’s work in order to better characterise ethics of care as a moral paradigm.

Firstly, it should be made clear that while the “ideal of care” proceeds from the recognition of the interconnection and vulnerability of human lives, Gilligan seems to acknowledge the space for individual responsibility: the responsibility for participating in the web of relationships which sustain life, i.e. being responsive to the needs of others in their particularity. This sense of responsibility proceeds from a compelling and intimate representation of one’s own position in relation to others and to their needs.

Secondly, with regard to acquiring this intimate representation, although she generally speaks of “sensitivity” and “solicitude”, Gilligan also indicates more specific kinds of capacities as crucial, such as: verbal communication, narration and listening; sentimental communication, empathy and sympathy; attention and imagination (differentiating these from identification and generalization).23 She doesn’t establish a hierarchy among them, nor does she offer specific characterizations, but calls for all these capacities together; while of course they have been conceptualised differently and are attributed varying values in different philosophical developments of care ethics. It may be important to note that at least in her last book, Gilligan seems to give more weight to empathy (also referring to neurobiological and anthropological studies), thereby opening the way mainly to sentimentalist reconstructions of care ethics.24

Finally it is worth considering that, in her understanding, relationships can be sustained only in their particularity and should not be encapsulated within abstract rules, and therefore attention to differences and details is fundamental. Moreover, this attention to particularity is taken to be important for each (that is to say, for any kind of) relationship in which we might find ourselves, and is not limited to characterising some particular or specific kinds of relationships (such as for instance personal affective relationships). This is a significant development: at least in my understanding, the point Gilligan is making is that the needs of each participant in a relationships (that is to say those of each human being) should be met in his/her particularity, and not according to idealizations, generalisations,

23 Regarding “empathy” see Gilligan, *In a Different Voice*, p. 8 and 11, where she characterises girls as having a greater store of it; she speaks of sympathy only when quoting George Eliot, at p. 148. General reference to sentimental communication as well as to verbal communication, narratives and listening are ubiquitous in her work, as well as to attention. On imagination and generalization see p. 59.

24 See Gilligan, *Joining the Resistance*, where she refers to the works of Antonio Damasio and Sarah Blaffer Hardy.
Conversely, however, care ethics is more often thought of as a model that advocates the importance of meeting the needs of (some) particular others, namely those with whom we have particular affective ties (see for instance the considerations concerning the opposition between partiality vs. impartiality in our obligations, developed following the above interpretation). While of course there is a problem, and not a trivial one, in understanding how we can represent and feel the particular needs of persons we do not know directly, it is clear that this forms part of what Gilligan is aiming at. It is not a casual fact that the core of Heinz’s dilemma does not turn around a particular affective relationship (or around the conflict between an impartial responsibility and one emerging from a particular tie), but lies in the encounter between two strangers (Heinz and the druggist), which, as Gilligan explains, can be explored in its moral aspects in two ways: out of an abstract logic of rights or out of care for the particular needs of all involved. This is a critical point in my understanding and one of the core issue of this paper.

Thus in Heinz’s dilemma these two children see two very different moral problems – Jake a conflict between life and property that can be resolved by logical deduction, Amy a fracture of human relationships that must be mended with its own threads.

Similarly, it is important to make clear that, while appealing to the capacities and qualities of each to care for others, or to sustain the web of relationship in which one is enmeshed, in Gilligan’s account care ethics is not a form altruistic ethics, calling for benevolence or love and risking self-sacrifice. Gilligan clearly states that the different morality she has identified, in listening to girls and young women, is a morality that results from a development, just as much as that identified by Kohlberg, whose final stage results precisely from the overcoming of self-sacrifice or “maternal care”. In fact she is not merely making reference to the sort of greater openness to relationships, or willingness to communicate, or empathy, which she traces in the psychological development of girls (supposedly due to the connection with the maternal origin, a position which differs from that

25 See for instance where she claims: “only when substance is given to the skeletal lives of hypothetical people is it possible to consider the social injustices that their moral problems may reflect and to imagine the individual suffering their occurrence may signify or their resolution engender”, Gilligan, In a Different voice, p. 100. On this issue, Gilligan quotes George Eliot, from The Mill on the Floss, where Eliot lets her main character, Maggie Tulliver, claim that “the truth of moral judgments must remain false and hollow unless they are checked and enlightened by perpetual reference to the special circumstances that mark the individual lot” (Gilligan, In a Different Voice, p. 148). This is an interesting reference, for instance in the light of a sentimentalist reading of care ethics, since, as is well known, Eliot appreciated David Hume’s philosophy. Of course there are also other possible readings of the importance of particularity in different traditions, see for instance Blum’s and Laugier’s interpretations of care ethics.

26 Gilligan, In a Different Voice, p. 31.
of separation of boys), and calling for a morality of (natural) compassion, benevolence or altruism. She is rather making reference to a process which involves different stages, the latter of which, the one corresponding to moral maturity, is defined in terms of the ability to care for the other as well as for oneself through realising that “responsibility now includes both self and other, viewed as different but connected rather than as separate and opposed”. Here lies the ideal of the responsible care, which differs from a preceding stage of altruistic “maternal” care (which is in turn comparable to Kohlberg’s conventional stage).

It is worth mentioning – expanding this latter thesis – that stating that one cannot care for others at the expense of oneself is not a claim derived from the value of one’s own rights or sovereignty, but emerges again in the form of a “relational truth”: the idea is in fact that relationships will not hold if one member is only giving, and in so doing looses all her energies. In the same way, interest in the others’ suffering or flourishing is not derived from an external injunction to care for them or to be benevolent, but from the connection between our own flourishing and that of the others, between our feelings and those of the others.

Finally it should be noted that the idea of the need to overcome care of a maternal kind is a sign of what I was arguing above, namely that Gilligan is not only acknowledging the (moral or general) value of traditional feminine attitudes (or virtues), but envisaging a complex new paradigm emerging from, and able to deal with, the recognition of the human condition of interconnection.

In the same light, maintaining the web of relationships does not automatically mean that no one will be left out, as if it were a zero sum game. Of course there can be costs and hurts and harms, and even separations and failures, in the effort to maintain interconnection. Coping with the vulnerability and contextuality of (human) life and experience means coping with the evaluation of these costs, but also with the acknowledgement of their inevitability. These costs and hurts can be thought of as in some way individual but, in the light of the conception of porous subjectivities underlying this vision, they cannot be easily ascribed to one or the

---

27 In later writings she has revised this claim, disentangling the development of empathy and communication from the gender divide, and considering these capacities as “natural” for both girls and boys, but obliterated by the patriarchal superstructure which girls for a number of reasons resist a bit more than boys. These qualities could lead instead to a different human moral approach. See Gilligan, *Joining the Resistance*. Chap. 2 and 4.

28 Gilligan, *In a Different Voice*, p. 147.

29 And it is worth noticing this point, since other care ethicists instead have heavily relied on mothering as the relevant metaphor to exemplify caring attitude and to specific relationships its domain (see for instance Ruddick, Noddings and Held).

30 Again Gilligan’s quotes Eliot: “Since ‘the mysterious complexity of our life’ cannot be ‘laced up in formulas’, moral judgement cannot be bound by ‘general rules’ but must instead be informed ‘by a life vivid and intense enough to have created a wide, fellow-feeling with all that is human’.” (*Gilligan, In a Different Voice*, p. 130). Again Humean echoes are more than present.
other.\footnote{This point emerges clearly from the study on abortion, see Gilligan, \textit{In a Different Voice} chap. 4. See also Baier on responsibilities which cannot be easily divided into what is mine and what is yours. See Annette Baier, \textit{Moral Prejudices. Essays on Ethics}, (Cambridge Mass.: Harvard University Press, 1995), p. 266.} One could say that, in Gilligan’s view, the core (individual) responsibility is not that of facing and answering the other’s needs as given, or as the other represents them, but that of maintaining the possibility of an exchange (sentimental and dialogical) in which the needs emerge, are represented and cared for, as far as this is possible, and that this will not necessarily lead either to an easy solution or to the expected one.\footnote{See where she claims, commenting on the words of some interviews: “then the notion of care expands from the paralysing injunction not to hurts others to an injunction to act responsively toward self and others and thus to sustain connection.”(Gilligan, \textit{In a Different Voice}, p. 149).} Survival, flourishing and the reduction of suffering are intrinsically relational; they should not be represented as the individual claims or needs of certain separated selves that can be answered or met by certain other benevolent selves, but as issues emerging and finding answers in the context of relationships, personal and social, which keep together individualities and life. Notwithstanding this accent on interrelationship I think that, as has been said, there is space in this account to conceive of one’s personal responsibility in the shape of taking part in the interpersonal exchanges which emerge from the interconnection, and returns to it.

In Gilligan’s words:

The changes described in women’s thinking about responsibility and relationships suggest that the capacity for responsibility and care evolves through a coherent sequence of feelings and thoughts. As the events of women’s life and history intersect with their feelings and thought, a concern with individual survival comes to be branded as “selfish” and to be counterposed to the “responsibility” of a life lived in relationships. And in turn, responsibility becomes, in its conventional interpretation, confused with a responsiveness to others that impedes a recognition of the self. The truth of relationship, however, returns in the rediscovery of connection, in the realization that self and other are interdependent and that life, however valuable in itself, can only be sustained by care in relationships.\footnote{Gilligan, \textit{In a Different Voice}, p. 127.}

The central tenet of responsible care as moral maturity – that is to say of an ethic of care – is therefore that of caring for the relationships among oneself and the other/s, caring for oneself and for the other/s in the relationships, and caring for life as a web of relationships, woven from the capacities which emerge from a “coherent sequence of feeling and thoughts”. And of course a deeper analysis of this “sequence of feeling and thoughts”, at least at the personal level, is necessary from a more structured philosophical point of view.
In this light, a final point concerning the object of our care or responsibility requires a mention here, because this is again the point at which different readings of care ethics are possible. While Gilligan places a strong emphasis on the importance of personal involvement in relationships, in meeting the needs of the others or sustaining relationships in their particularity, I wish to stress that, at least in her later writings, the needs or support she refers to are not only those involved in the physical survival of each, or in relation to a physical or psychological (inter)dependence, but also – at a deeper symbolic level – those relating to the recognition of one’s own personhood and voice. In this light, taking into account the other’s and our own “suffering” can be viewed in many ways. These can include taking into account the pain caused by physical harm, the pain of being left alone or not having significant relationships, or that of not being recognised as having a voice or as being a person. The notion of care thereby shifts from the sphere of simple (naturalist, psychological) kinds of attitudes and activities, as for instance those exemplified by nurturing and the like, to that of a more complex symbolic play of interrelations that are clearly linked to the social and cultural level.

What is interesting, in my opinion, is the link between these different levels, and the idea that there is personal responsibility involved in considering all these different levels, while caring for the others, or for our relationships with the others. In other words, for example, we should cultivate a sensitivity towards different levels of violence and disapprove of them. So our being attentive to others can be configured as being attentive both to the others’ and to our own vulnerability at different levels, an attention that finds its roots in our own vulnerability, and one that can take the shape of a personal responsibility to answer to the others’ needs in order to maintain a rich texture of relationships which is in turn respectful of people’s differences and contributes to their flourishing (as we will see later a consonance can be found here with most recent feminist awareness, but also with certain refined sentimentalist considerations, such as that found in Baier’s work on cruelty in morals).34 Here again is where, in my opinion, a difference between just revaluing traditional feminine attitudes (as for instance nurturing) and the value of a wider (and multilayered) meaning of the proposed caring attitudes becomes important.

To sum up, in my interpretation, Gilligan’s ethics of care can be understood as stemming from a conception of humans as interconnected, fragile and vulnerable subjects who constitute and maintain themselves, at different levels (from their own physical existence, to their own identity, to their sense of separation and


125
independence, their flourishing and to the satisfaction of needs of different kinds),
through relationships. Porous to one another and – also through personal
relationships – to society, to culture and to various forms of discourses, they are
for that same reason in need of and capable of morality. A morality which can be
thought of, in the end, as the practice of attention to others in their particularity
and difference, woven from solicitude and care toward the relationships in which
we are all enmeshed.

The point now is how to configure this solicitude and care, so as to
accommodate the complexity just mentioned.

Here is where an account of the subsequent – and more philosophically
structured – debate on care ethics becomes relevant, along with the reference to
recent developments in feminist thought.

The general point I would like to make is that, in giving an adequate
philosophical account of the insights proposed by Gilligan, one should pay
attention to two important issues among those I have mentioned: that concerning
the existing ambiguity between caring for others in their particularity and caring
for particular others; and that regarding the different levels at which we can think
of the needs (as a way to express in brief what we should care for) of others and our
own. While the first question seems to me rather straightforward, I will offer a few
further considerations on the latter. I have tried to argue that, on the one hand,
the needs at stake in care ethics can be conceived as the pleasures and pains which
we might represent or feel or imagine in various ways. Either, for example, on the
basis of a common grammar of pain and pleasure upon which given human
faculties operate, or on the basis of a certain “knowledge” of human nature (as in
quasi-Aristotelian accounts),35 or finally on the basis of certain forms of easily
shared narration and communication. On the other hand, the needs at stake can be
considered more broadly as ranging across different levels of experience, thus
presenting the problem of the visibility or accountability of particular ways, or
forms, of living, of being human, of being a person, of having a voice or of feeling
pleasure and pain. To conceive of needs in this latter sense implies taking into
account the problem of the representation of different grammars of pleasure and
pain and that of the recognition of different voices, in a word, the radical problem
of difference and, as we will see, also that of the instability of the categories in
which we express all this.

It is my conviction that the existing debate on care ethics might be pushed
forward with regard to these two issues, and that useful contributions could be
drawn from feminist authors who are not engaged directly in the debate on care
ethics. In what follows I will be rather sketchy on the debate on care, and more
detailed on the feminist contributions.

35 Tronto in Moral Boundaries offers this kind of reference.
3. The debate on care ethics

As has been said, there have been many attempts to develop more encompassing or more philosophically structured elaborations of an ethics of care.

Among the many it is worthwhile to recall Virginia Held and Joan Tronto’s attempts to offer a comprehensive account of an ethics of care, considered as a distinctive moral paradigm, as well as the important work of Michael Slote (developed mostly with reference to Nel Noddin’s writings on care), in which he inquires into the consonance between an ethics of care and a sentimentalist approach to ethics, also in relation with a reflection on virtue ethics. His proposal of an ethics of care and empathy, his work on the breadth such an ethics could encompass, reaching out also to more traditional moral dimensions (such as deontology) and, more recently, his work on its limits (in the form of the dialectic between partial and impartial or personal and impersonal values or virtues), constitute one of the most comprehensive attempts to give shape to an ethics of care.

Another crucial reference, of course, is to Annette Baier’s work. Although Baier didn’t offer a comprehensive account of an ethics of care, dedicating her research – as is well known – mostly to the development of Humean themes (and being – as is also well known – rather diffident with regard to ethical theories), in my opinion she offers many of the most interesting clues to the possible consonance between care ethics and a distinctive and particularly refined reading of David Hume’s sentimentalism.

Finally, among the interesting philosophical researches made on care ethics, one should not forget the work on moral particularism, in part with reference to themes deriving from Iris Murdoch’s thought, as put forward by Lawrence Blum, and the recent work of Sandra Laugier in France, who offers an interesting reading

---

36 See Held, The Ethics of Care; Tronto, Moral Boundaries.
39 Baier, Moral Prejudices.
40 This is not the place to go into the debate around Hume’s ethics, though many other contributions to this volume testify to its richness. It is however clear that in attempting to give a more philosophically accurate account of care ethics in sentimentalist terms, different interpretations of Hume do become relevant. I will hint at particular points in what follows.
of care in terms of ordinary language ethics and moral perfectionism, in Stanley Cavell’s sense.  

In what follows, I will concentrate mostly on the works of Held, Tronto and Slote (leaving aside for the moment the other works mentioned, despite their relevance), as these are more widely known, and can be said to have contributed to the mainstream understanding of care ethics that is under investigation here. As already said, I will not enter into the details of these various analyses; I will instead limit my inquiry to considering whether or not, and if so in what ways, the richness and the complexities of the implications of Gilligan’s work that I have tried to show have been accounted for in these elaborations, sketching in a very schematic and external comparative analysis.

The most relevant point I would like to make here concerns the fact that – at least in some contexts (for instance in Held’s and Tronto’s development of a moral model based on care) – the nature of care itself seems to go rather unquestioned as it is defined in terms of an experience or attitude we all already share (as for instance in the case of the care exchanged in the private sphere of personal and affective relationships). In these accounts, the existence and characterisations of the human attitude and activity of care is rather unscrutinized, or considered simply in terms of a shared experience, while the analysis is centred on the issues of developing a moral paradigm based on recognition of its moral worth, that is to say, on the value of the well know human ability to care, and of considering the breadth of its domain of application (and this holds true also for Slote’s analysis, although it is different in other respects).

Held for instance in her book, The Ethics of Care. Personal, Political and Global, makes clear that an ethics of care as a normative perspective is based “on

---


42 It is worth mentioning, as an aside, that both Held and Tronto insist on considering both the dispositional aspects and the practical ones of caring, they see caring in a moral sense as involving not only being solicitous but also acting, engaging in practices of solicitude. For them this characterization poses a difficulty to the possibility of reading care ethics in terms of virtue theories. Although interesting this is not an issue I will discuss, partly because I do not think it impossible, from within some virtue ethics approaches, for instance those derived from sentimentalist approaches, such as Humean ones, to take into account actions and their consequences when considering virtuous traits of character of individuals; and partly also because I am more interested in the question I am trying to define here, which seems to me rather more fundamental, that of the possibility of to meeting the others’ needs.

the truly universal experience of care”. An experience common to everybody, at least, as she says, in that every human being: “has been cared for as a child or would not be alive”. The ordinary and daily nature of the practice of exchanging care is evident, and this is witnessed for instance – she claims – by the North American expression “take care” as a common way to take leave of people.

While for Tronto care should be understood in terms of a group of activities which characterise our human functioning, at least at an experiential and ordinary level. Her definition of care is often quoted:

A species activity that includes everything that we do to maintain, continue and repair our ‘world’, so that we can live in it as well as possible. The world includes our bodies, our selves, and our environment, all of which we seek to interweave in a complex life-sustaining web.

For both these authors, care is thus an activity and a competence which is already present and functioning in our lives, although it is often invisible or wrongly considered as confined to certain spheres of experience (that of particular relationships characteristic of family context, education or health care etc.), or else as delegated to certain individuals (women or minorities). They present it as an activity and competence whose practice and value should instead be recognised and placed at the centre of a more encompassing moral vision (and social transformation), suggesting that the qualities and competences characterising caring persons should become more widespread, and that the value of care should be recognised and translated from the personal to the social level, informing social institutions.

In both their analyses the value and the possibility of this type of expansion of an already given capacity and activity of care seems to be the main issue upon which to exercise a philosophical effort, while the characterisation of the qualities, or virtues, constituting the practice of care, as well as the inquiry into the human

44 Held, The Ethics of Care, p. 3.
45 Held, The Ethics of Care, p. 3
46 Held, The Ethics of Care, p. 29.
47 The ordinariness of care is defined also by Tronto in terms of the common presence in our language of the formula “I care” in opposition to “I don’t care”, as denoting “some kind of engagement”. See Tronto, Moral Boundaries, p. 102. For an interesting analysis of a possible ontological level of Tronto’s analysis, with an Aristotelian resonance, see Stephen K. White, “Care and Justice. Ontological, Ethical and Political Dimensions”, (paper presented at the International Society for the Study of European Ideas (ISSEI) Conference, Utrecht, 1996).
48 Tronto, Moral Boundaries, p. 103.
49 Although both authors do use the term virtue in this context and also make use of typical expressions of virtue theories (such as for instance flourishing, excellences and so forth), it could be worth mentioning, for those interested in the possibility of giving an account of care ethics in the shape of a virtue ethics, that both Held and Tronto are sceptical concerning this possibility. In their opinion, it is a mistake to reduce care ethics to the form of a virtue theory because such
capacities and faculties involved, is rather hastily undertaken. In their books this effort amounts to a small part of the whole. In Held it is devoted to underlining the role of attentiveness, sensitivity, and the ability to respond to needs, basing this on a sentimental dynamics to be corrected in relationships by dialogue and communication, or to be refined, on the model of exemplar practices, for example, that of “mothering”.  

In Tronto this effort amounts to the identification of four “moral qualities”, namely attentiveness, responsibility, efficacy and responsiveness, which should be integrated into our moral practice. 

It is moreover in this context that, while recognising a continuity between the qualities characterising the caring person, such as attentiveness, sensitivity, responsibility, relational competence, and responsiveness, and the qualities traditionally ascribed to women (and devalued), both authors strongly deny that there is anything essential about this connection. An ethics of care is such precisely in as far as it puts pressure on such an essentialist connection between care and women, and recognises the human value of those qualities, which are seen as only ascribed to women.

Both Tronto and Held seem mainly interested in this latter claim. In their understanding, it is necessary to free care from an essentialist feminine connotation, which is characteristic of the old patriarchal framework (and according to Tronto also of the modern moral point of view). Achieving some

an account would consider care only for its dispositional and intentional traits, undermining the value of the actual engagement in the practices of care. Such an account would be unable to cope with the relational and social nature of care and is pervasively patriarchal (see Held, The Ethics of Care, chap. 2-3; Tronto, Moral Boundaries, chap. 4-5). Of course in my opinion these are generic and disputable claims, but to argue in this direction would be the object of a different paper.

50 See Held, The Ethics of Care, p. 20, where she claims that an ethics of care, with “its alternative moral epistemology”, “stresses the sensitivity to the multiple relevant considerations in particular contexts, cultivating the traits of character and of relationship that sustain caring, and promoting the dialogue which corrects and enriches the perspective of any one individual”. For what concerns care moral epistemology held makes a rather syncretic reference to Annette Baier’s and Margareth Urban Walker’s works. The reference to mothering as the moral experience involving “feeling and thought experienced together” is instead in Held, Feminist Morality, p. 30.

51 Tronto, Moral Boundaries, pp. 126-137. For a possible Aristotelian reading see White, “Care and Justice”.

52 The socially constructed nature of the gender ascription of these qualities is shown, for instance by Tronto, when considering the fact, that in recent times, at least in North American society it is not even true anymore that women are those in charge of care work, which now weighs upon other minorities. On this basis Tronto criticises Gilligan’s work, especially with reference to the gender divide she retracts in the answers to the moral dilemmas. See Tronto, Moral Boundaries, chap. 3.

53 Tronto argues in the second chapter of her book, that the stabilization of the modern impartialist paradigm – during the 18th century - has rendered necessary the invention of
distance from the “old frameworks” is therefore linked to the possibility of recognising (an ethics of) care not as distinctively feminine but as distinctively human, and also to the possibility of enforcing the transformation of social institutions that is necessary for this recognition, interlacing in some way considerations on care with those on justice. In a word, the main point in their understanding is the need for the multiplication of the practices of care and the recognition of their value. But, in all this, the functioning of care as such, or the functioning of the human qualities it requires, remain somewhat unexplored.

I’d like now to take up the two points made at the end of the previous section. On the one hand it could be said that in these readings the issue of delineating care not only in terms of recognising the special value of “particular relationships”, but also in terms of the attention paid to the “particular needs of all” has been resolved through the transposition of the value of care from the personal to the societal level; while on the other hand, the difficult issue of the visibility and the invisibility of needs has not been adequately examined. The epistemological complexity revealed in the re-elaboration of the concept of care hinted at before, and present in Gilligan’s work, seems to have been left aside.

Similar, but not identical, considerations can be put forward in relation to other developments of care ethics.

In Slote’s interpretation, for instance, care ethics is developed as a specific form of sentimentalism where what is crucial is the psychological mechanism of empathy, whose functioning and central role in moral development he retraces both in recent psychological literature, and in eighteenth century moral sentimentalism. His analytical effort is therefore more far-reaching in this sense. Yet, although Slote offers a more structured philosophical reconstruction the core functioning of an ethics of care, his main interest seems nevertheless to be in the problem of how, and to what extent, one can give account of more traditional normative and political distinctions, such as obligations to distant others, deontological distinctions and the like, from within such a paradigm. He illustrates the way in which care ethics can offer an account of these distinctions and where instead it diverges and conflicts with impartialist accounts, since, in his understanding, care ethics privileges (and this is a point of difference with the readings previously discussed) mainly particular relationships.

Still, it seems to me that, while the issue of particularity and partiality is tackled, although not in the same way as Held and Tronto, there is again less attention to the problem of the different levels in which the reflection on personal

“feminine morality”, and the definition of “feminine virtues” in order both not to abandon an important dimension of human life, that of care, and to control women.

54 See Slote, Ethics of Care and Empathy and Slote, Moral Sentimentalism.

55 See Slote, Ethics of Care and Empathy. See for instance his reference to Williams’ problem about the integrity of agents, p. 33 and chap. 5.
responsibility in caring for the well being of others, and for the flourishing of the web of relationships sustaining life, should be articulated.

In one of Slote’s most recent publications however he develops this point in a particular way. In his recent book *The Impossibility of Perfection* he argues that it would be opportune to adopt a balance between different philosophical methods and ideals, accounting for “‘masculine’ concepts like autonomy and justice”, and “‘feminine’ ideals such as caring about and personal connection”. This can be thought of as one way, although not the one I would suggest, to deal with this problem, i.e. that of recognising a limit to the paradigm of care, that is to say the partial value of both care and justice.  

Of course Slote is not assuming that “feminine ideals” means, in this context, that what is socially constructed as feminine pertains only to women, and in fact he adheres to the feminist idea of considering these ideals not only as morally relevant, but also as characterising the moral thought and practice both of men and women. Interestingly, moreover, in this most recent book, relying more on the work of Gilligan than in the previous ones (where he relied more on the work of Noddings), Slote defends as particularly feminist the idea he is proposing, that we should think of ethics as seeing “partial values that are equally relevant to men and women”, and argues in this sense for imperfection.

This epistemological assumption, for which ethics in general consists of partial and different values and methods, although interesting in the light of the considerations I have put forward, nonetheless seems to me, in its application to care ethics, still to limit the understanding of this latter to a rather direct possibility of knowing how to care for particular others (a possibility which, in my opinion, should be subjected to theoretical scrutiny).

While I agree that some feminist thought does suggest the idea of the partial and incomplete nature of our values and of our moral theorising, or alternatively the idea of thinking of moral life as tragic and imperfect (although Slote rightly indicates Berlin as the champion of this latter point), I think that on the basis of a certain feminist awareness we might push forward this idea of instability. From this perspective, we might challenge the idea of care as a “partial value” with relation to justice, but as a complete value in relation to certain interpersonal settings, e.g. when we are caring for our beloved. This latter is, in my opinion, an option which does not account for all of the implications which can be derived from Gilligan’s insights, on the basis of which a more encompassing vision of an ethics of care could be developed, both regarding its interaction with justice (i.e. the problem of distant others), and in relation to a less romanticised vision of the limited domain of personal relationships: in other words, a vision requiring a more complex reconfiguration of the moral domain.

---

56 See Slote, *The Impossibility of Perfection*, p. 4 and the entire volume.
57 Slote, *The Impossibility of Perfection*, p. 34 but see the whole first chapter.
What can usefully be derived from Gilligan’s work, at least in my opinion, is that the emphasis on interconnection or interdependency goes together with that on vulnerability, which in turn can be also expressed in the form of the difficulty of meeting the multilayered needs of others or even one’s own. In this sense, as I have tried to show, caring does not mean only meeting the needs of our nearest and dearest, construed as a relatively easy task, or caring for them and also being able (through processes of refinement or artificial processes, conventions or institutions) to care for the needs of distant others with the same or sufficient attention (or in the most tragic version being split by the two partial values of care and justice). Given the different levels of human vulnerability and needs we have considered before, it is in fact impossible to consider the needs of our beloved as more easily identifiable and as severed from those of the others. Caring means therefore engaging in enough care to keep the fragile fabric of connections together and in the difficult task of recognising the difference of the other in its many aspects, the particularity of each life and the specific voice of each individual. This brings into question in both cases (near vs. distant others) the limits of our sensitivity and imagination, and of our listening, although these are fundamental and necessary resources. If – as I believe – the solution does not lie in calling for a (complementary or conflicting) impartialist methodology to confront these problems, since this is blind to particularity and differences due to its very structure, it is from within the same resources of sensitivity and imagination that we shall find a way to represent this kind of multilayered caring.

I think that some works on Hume’s reflective sentimentalism could be of interest in dealing with these issues, as well as considerations derived from Murdoch’s ideas on perception, imagination and the relevance of frameworks, as developed for instance in Blum’s moral particularism, or in Laugier’s most recent works on moral perfectionism and ordinary ethics. But I think that the clearest illustration of the problems I am trying to represent, together with some indication of the way forward, come from recent developments in feminist thought. So, even though some of these feminist reflections have been developed in dialogue with philosophical traditions that are a long way from those considered in the debate on care ethics, in the following section I will offer a brief excursus on recent feminist thought, and in the final section, return to the philosophical arena of the debate on care ethics.

---


I believe that offering even such a brief analysis will serve to clarify the different senses in which care ethics can be said to be feminist, and to render visible the ambiguity often present between feminine and feminist, as in the case of the recurrent considerations regarding “feminine virtues” already hinted at. And finally, I hope to draw from this analysis, some indication – to say it with the language of virtues – as to which traits of character might be characterised as “virtuous” from within a recent feminist viewpoint, traits which, and this is the suggestion I will put forward in the final section, might be appreciated also from within a more refined version of care ethics.

4. The parabola of feminist thought

If feminist thought has anything important to offer to moral reflection in our times, it has not, in my opinion, or not only, to do with the rehabilitation of certain human experiences linked to the traditional activities of caring, considered rightly or wrongly as feminine (as in Held’s claim concerning the “truly universal experience of having been cared for as a child”), but something more subtle that we can learn from the history of feminist thought.

What I find interesting in this history (but I do not claim that this is exclusive to this line of thought) comes from the considerable range, or parabola, of different positions and competing claims within the tradition, and in particular from the oscillation over time regarding the categories of equality and of difference in the vindication of women’s subjectivity and freedom and from recent elaborations which take into account the complications that results from this alternating trend.

As is well known (and as mentioned in the first section), in order to vindicate women’s full subjectivity and freedom, feminists have put forward a series of different strategies. These can – in very broad terms – be said to characterise different phases of the history of feminism, intended both as the political movement and as the theoretical reflection which has developed around it. As we will see, one way of giving an account of these different strategies is by referring to their different treatments of the question of “feminine virtues” – as to whether or not they exist, whether or not this is important.

The beginning of the feminist movement and thought is commonly thought of as coinciding with the suffragist movement of the second half of the 19th century (although one might go back also to Wollstonecraft and De Gouges), and for the sake of this argument we will consider this as a valid claim.60

60 On the history of feminism and the possibility of mapping the differences between Wollstonecraft and De Gouges I have written elsewhere. See Caterina Botti, Prospettive femministe. Morale, bioetica e vita quotidiana. (Milano: Mimesis, 2014).
Following the line of argument of this paper, it can be said that the suffragist movement was asking for a fuller recognition of women as human beings and as citizens, while at the same time claiming recognition of the worth of the “feminine virtues”. The suffragists argued that the inclusion of women would enrich humanity and citizenships precisely because of their specific virtues. So the suffragist claims, whilst being a plea for equality in status, resulted also in a sort of acknowledgment of an important and not negligible difference among men and women.

Coming back to care ethics, it is noteworthy that many authors, for instance Tronto, hold that some theorizations of care ethics can be reduced to this same schema, and can in this sense be criticized.

Why criticized? Because it has been suggested that, since the requested acknowledgment of feminine virtues is in line with patriarchal framework, such a vindication, in the end, would not lead to the acknowledgment of the full human, moral and political status of women. As Tronto – among others – argues, the feminine virtues have here been defined by men and imposed on women, obtaining the double result of having a part of humankind dedicated to the necessary work of care, and at the same time not recognizing this work as being of any moral worth or even as being characteristically human. In this light, if women are asking to be included as moral subjects, or as citizens, on this basis, they will not substantially change their situation: they will continue to be the ones in charge of care and will continue to be somehow devalued as subjects (although they may be able, say, to vote) since care and all its inherent problems will still be considered as their lot. (Moreover, according to Tronto, this essentialist claim also results in the invisibility of the fact that there are other marginalised groups besides women, and that there are substantial differences in power amongst women themselves). A change in this situation will come about only when the shape of social life and of political institutions is changed, such that care will be recognised not only as a fully human practice (culturally and morally laden), but also as a fundamental social and political value, that is to say when the public/private divide is radically rearranged. Until then, the caring attitude, even if it is re-established as worthy, will continue to be marginalised and marginalising.

The suffragist strategy has in fact been abandoned, but not in the direction suggested by Tronto and others.

In the 20th century, after the two wars and the social changes that occurred, what can be defined as the true feminist movement distinguished itself from the previous suffragist movement. Exploding in the sixties, the feminist movement – which finds one of its early theorizations in Simone de Beauvoir’s Le deuxième

61 See Tronto, Moral Boundaries, chap. 2. But see also Held’s interesting analysis, in Feminist Morality, chap. 6, on how death has been considered as a “distinctively human” experience, and giving birth has not.

62 See Tronto, Moral Boundaries, chap. 2.
sexe\textsuperscript{63}, – claims equality between men and women in much more radical terms. What is vindicated is no longer inclusion in difference, but the recognition of a more substantial equality, in terms of the same open-endedness in the unfolding of life. These (first) feminist vindications are based, in fact, on the leading idea of the equal nature of men and women and demand the possibility for women to decide upon their own lives as men do. From a more theoretical point of view, the crucial tenet of this feminist strategy is that of disconnecting the possibility of subjectivity from bodily differences, which are considered to be irrelevant, and on which a discriminatory destiny has been socially constructed (according to the sex/gender distinction).\textsuperscript{64}

In this light, “feminine virtues” are feminine only in the sense that they have been imposed on women, not because there is any direct connection between having those virtues and being born anatomically of the female sex (and therefore, for instance, able to give birth to children). What is more, as attributes they should be rejected in so far as they prevent women from gaining full subjectivity, which is broadly considered as the ability to transcend one’s own corporeal determinations. The ideal of subjectivity characterising this kind of feminist thought is in fact the incorporeal, reason-centred, transcendent idea of the subject of the modern tradition.

According to this kind of feminism the crucial aim for women, in order to gain freedom, is to deconstruct the discriminatory scaffolding of patriarchal society, which prevents them from flourishing in the same way as men can. In other words, it is the aim of eliminating prejudices together with all the social structures which enforce those prejudices, considered as unjust discriminations.

To clarify the point, it could be said that care activities are not only not considered as feminine traits in this perspective, but they are also devalued in general as human traits, to be substituted by social institutions or technologies which could liberate humanity from the most animal traits (or from bourgeois institutions such as the family). Think for instance of the enthusiastic reception of


\textsuperscript{64} Of course many references should be offered for this very general claim and for the whole overall and sketchy presentation (and personal interpretation) of the history and characteristics of the different feminist strategies I am offering. For general references see for instance: Linda Nicholson, ed., \textit{The Second wave. A Reader in Feminist Theory} (London-New York: Routledge, 1997); Sarah Gamble, ed., \textit{The Routledge Companion to Feminism and Post-feminism}. (London-New York: Routledge, 2001). For the personal interpretation I am offering and its references, see: Botti, \textit{Prospettive femministe}, chap. 1-2.
new reproduction technologies which characterised the work of certain American feminist thinkers in the 1970s.65

As I have already said, if this had been, which it was not, the kind of feminism guiding Gilligan’s research, she would have claimed that a discriminatory socialisation prevented girls and young women from proceeding to what Kohlberg had defined as moral maturity, and that women’s minority, though correctly measured, found its roots not in a different and inferior feminine nature (which was mistakenly sustained), but in the unjust and discriminatory treatment of women.

But this is not the point made by Gilligan. Gilligan’s feminism distinguishes itself from that vindicating equality, and is instead indebted to, as well as contributing to, a different kind of awareness.

In fact, starting from the mid seventies of the last century, a sharp change occurred in feminist thought, or at least in part of it. For the sake of this argument, I am offering here a very simplified and generalised account of a more complex process involving geographical, political and theoretical differences.

What at least some feminists came to realise and to theorize was that in vindicating equality in the form of women’s equality to men, the earlier feminist strategy had adhered to a model of humanity that – while supposedly neutral – was in reality partial and therefore inadequate: inadequate, for example, in terms of representing women’s experiences and subjectivity, or their freedom, and therefore, or more in general, also in terms of representing humankind as such. On this view therefore, the conflict should be brought to this new level.66 This is a move which is, as we will see, substantially different from that of going back to the suffragist idea of the rehabilitation of traditional (i.e. patriarchal) feminine virtues.

Women’s experience of pleasure, the conflicts of the sexes in the private sphere, the experience of pregnancy and of making decisions about it, for instance, were seen to be misrepresented, or as impossible to represent, within the bounds of the categories to which women had adhered in, for example, the general and abstract notions characterising liberal or Marxist conceptions of justice. Feminists of this persuasion therefore recognised the partiality and non neutrality of the categories and concepts used in scientific and political discourses, or in philosophy, and felt the need to develop new ones in order to describe a more comprehensive ideal of humanity, able to acknowledge that humankind is made of men and women, or to focus on the partiality of women’s experience and subjectivity. In a

66 Again, it is impossible to offer here even a small bibliography, I will mention a couple of authors that in my opinion witness this turn in a peculiar way: Carla Lonzi, *Sputiamo su Hegel, la donna clitoridea e la donna vaginale e altri scritti*. (Milano: Scritti di rivolta femminile, 1974); Luce Irigaray, *Speculum. De l’autre femme*. (Paris: Les Editions de Minuit, 1974); Luce Irigaray, “Égales à qui?”, *Critique*, 480 (1987), pp. 420-437.
nutshell: what was put forward in this feminist strategy was not a claim for the recognition of women as having the same nature as men, but the need to question the whole idea of considering men as the measure of humanity or subjectivity, and question at the same time the different nature ascribed to women by men. In other words, the way in which that founding notion of human nature was represented, namely its being accorded universal status, was challenged. This meant finding new words and new categories to account for women’s existence and experience, and to account for human existence and experience. It is in this framework, that Gilligan’s elaboration of a different moral point of view can be understood, as both representing the moral development of women and at the same time, as envisaging a moral viewpoint that is able to cope with differentiated subjectivities.

So again the problem of women’s difference had been posed, but in a different way from that of the suffragists. What is at stake in this more radical strategy, is the idea of looking for new words, new concepts and categories, or for general symbolic frameworks, in which to represent women’s and human experience, as well new forms of political practices and relationships in which to develop them. If for de Beauvoir the feminist objective was not to become a woman in order not to be the “second sex”, the target considered here is to become at last Woman, the once obscured one, now a strong one, who is different from the “women-Other” considered as the second sex, but also different from the “man-Subject”.

While claiming once more that humankind is made up of men and women, these radical feminists demanded and took for themselves time and space (to be spent mainly among women, in consciousness raising groups, at least this was so at the beginning), in order to find the words to signify women’s and human subjectivity in a new way, putting aside the old patriarchal considerations on women and men and their difference. And this meant, of course, not only offering new meanings for women’s existence but also for human existence, and new descriptions of the world.

Although there are many differences between the French and Italian versions of “sexual difference theory”, and developments of this kind in the USA, relating also to their different philosophical backgrounds, it can be said that, in general terms, in this kind of feminism the central question is not to recuperate and sustain the value of what has been considered feminine and devalued in previous

situations, but instead to vindicate a completely new consideration of women, one which will allow them to name and to think their own experience. Such a rethinking will result in a different description of human experience or of the world. This entails a reconsideration not only of social structures (or social powers) but also of symbolic ones, and therefore a reconsideration of the way in which the human condition, subjectivity, knowledge and morality should be thought of and enacted.

Care ethics has been read at times as this kind of effort, and in the previous pages I have tried to show how Gilligan’s works might be read in this way. But, it has also at times been criticised precisely for not being an effort of this kind: being read only as a way of reasserting the patriarchal definition of women. This criticism remains valid even when care is assumed as a human value or virtue and yet not substantially requalified, when no major reconfiguration of the moral point of view is offered. This holds true for authors who, like Noddings, insist on asserting the feminine nature of an ethics of care; but also for those who, like Held, while identifying care as the basis for a human morality, see in practices such as mothering, construed as typically feminine, exemplar practices to be expanded to other contexts. The problem here being where the definition of what mothering is, is not distinguished from its patriarchal configuration.70

Care ethics can in fact be read in terms of this ambiguity: either as a new paradigm offering new contents to moral reflection, or as an old content (that of the traditional feminine virtues) brought to light and revalorised. And of course, when conceived as a new paradigm, also the extent of its domain has been configured in different ways, as we have seen in discussing Held, Tronto and Slote, as has its feminine or human characterisation. In fact many ambiguities do persist.

While many feminists have suggested abandoning this paradigm for these reasons,71 I personally do not consider these ambiguities to be sufficient grounds for abandoning it. In my opinion, the claim that humanity in its entirety is interdependent and vulnerable, and that this fact should mark morality, which ought therefore to be thought of more in terms of responsibility, care and relationships, than in terms of sovereignty and respect, or fair or contracted rules, is indeed an important new content and new starting point for moral reflection. It is clear however that these claims should be clearly framed in order to avoid such

ambiguities. Their novelty depends on their articulation, which for this reason should be very accurate and thorough (which is not the case, in my opinion, in many existing accounts of care ethics).

One way to articulate and to disambiguate care ethics might be — this is my suggestion — to set up a dialogue in which its own insights are confronted with those of the feminist parabola. In other words, it would be interesting to take into account the most recent considerations put forward by those feminists who have inherited the tension existing between the awareness of the instability of our categories and the need to argue for a transformation of society, of our self-understanding as humans, and of morality.

I will (briefly) describe this third strategy as that which has radicalised the notions of interdependence and vulnerability in such a way as to involve not only our physical or relational existence but also our abilities to think, to feel and to care. I am referring here to what has became known as the third feminist wave.\textsuperscript{72}

This form of reflection (here again I am simplifying and reducing to an ideal type a wide range of opinions and theories, that are indebted moreover with different axes of reflection according to which 20\textsuperscript{th} Century philosophy has challenged the modern conception of the subject, of knowledge and morality) can be said to have arisen in reaction to the move described above of assuming and claiming the value of a different feminine subjectivity, particularly when the latter has been defined in new terms with respect to patriarchal definition of women.

Although the idea of recognising a different female subjectivity has been acknowledged as a powerful one, and although it has been modulated in different ways, (for instance as a mimetic-strategic move necessary to contrast the binary and hierarchical patriarchal order, or as able to open up to the recognition of many differences, and not necessarily to the definition of a real difference in essence), it nevertheless implies — and this is what has been contested — a common definition of all women (as subjects). On this view, forms of feminism based on sexual difference theories run the risk (common to any identitary position) of obscuring relevant differences among women, and of conveying a stereotyped and partial description of the feminine, yet again taken as a norm which renders invisible other forms of subjectivity, even of feminine subjectivity, replicating the same problem of the false universalism of the monological patriarchal order.

Examples of criticism of this kind can be found in the critiques and elaborations of lesbian feminists, black feminists and post-colonial thinkers, in the rich literature produced from the 1980s onwards. And incidentally, these critiques have faced in turn the same problem concerning the “different” identities.\textsuperscript{73}

\textsuperscript{72} The first reference to the idea of a “third wave” occurs in Rebecca Walker, ed. To Be Real: Telling the Truth and Changing the face of Feminism (New York: Anchor Books, 1995).

\textsuperscript{73} See for instance the works of Adrienne Rich, Monique Wittig, Gayle Rubin, Teresa de Lauretis and Judith Butler, for the criticism developed in terms of lesbian and queer positioning, and
The result of this turmoil is a new evolution in feminist thinking. Aware of these risks, both the project of claiming equality, and that of asserting women’s difference as something essential or definable in terms of its contents, have been abandoned. Different authors have come to consider feminism as an (epistemologically aware) undertaking which moves from and tries to account for the many differences between subjects, opening up the space for each individual to live and flourish, and considering difference not only as a principle through which to describe human (and non human?) existences, but also as a space and a mode of thought.74

From this perspective, the relevance of different axes of differentiation – as characterising each individual – has been recognised, together with the fragmentation this implies in the representation or self-representation of each individual, generating fragments which may have differential attribution of value in relation to power. This complicates the pattern of one’s own position in terms of dominant or subaltern status (for instance my being a white, Western, woman…).75

At the same time, the same binary structure which characterises these axes, has come under pressure from the deconstruction of binary oppositions such as woman/man, white/black, western/non western, etc. A binary structure suggests erroneously the idea that it is possible to account for human variability in a number of fixed positions through these binary options, not recognising instead the existence of a continuum of differences which eludes any such rigid definition.76

In these terms, the possibility of representing human beings once and for all, in their particularity, even in the first person (my own particularity), is radically challenged, not only because there is this multiplication of elements, but also because it is difficult to grasp, to maintain and to order these elements in a meaningful and definitive way. Since the possibility of accounting for this multiplicity and these fragmented selves varies according to time, to experience, to relationships and to social and cultural structures, it is difficult to individuate those elements which are meaningful for one’s own account, and at the same time

---

74 See Braidotti, Nomadic Subjects, chap. 5. For difference as “a mode of thought” see Maria Luisa Boccia, La differenza politica. Donne e cittadinanza. (Milano, Il Saggiatore, 2002), chap. 2, where she interprets Lonzi’s reflections on difference in this way, as an existential principle. See Lonzi, Sputiamo su Hegel, p. 20.

75 See for instance the elaboration of the category of intersectionality developed among others by Kimberly Crenshaw; Patricia High Collins, and Nira Yuval Davies.

there is no possibility of evading this multiplicity, or reducing humanity to a few constitutive elements.

Who we are, what we need, and even our sex, gender, sexuality, is only a fragmented and contradictory experience, composed of internal drives and external constraints, of which we can offer only partial and temporary descriptions. Descriptions which maintain a degree of opacity as compared with a possible access to an authentic nucleus of identity.

Opacity is, in fact, the notion I am coming to: that opacity which places a limit on the possibility of giving an authentic, complete or even accurate account of oneself and one’s own needs, and therefore on the possibility of representing the identities and needs of others. A notion which of course is crucial in my understanding of the possibilities and of the difficulties of care ethics.

Of course, I am not claiming that this notion is a unique result of feminist thought. I am only interested in the way in which it has occurred in the feminist parabola. One could, in fact, analyse the deep consonances between these feminist claims and many philosophical and not only philosophical – think of psychoanalysis – developments in the thought of the 20th century, to which feminism is indebted, but this is not of relevance for the aim of the present paper.

What is interesting in this most recent feminist line of thought, for the purposes of this paper, besides its usefulness in underlining the importance of acknowledging the plurality and variability of human experience (which is what is defended by all care ethicists), is that it also suggests the need to acknowledge the laborious process of representing human experience in its particularity, even in the first person; that is to say, the laborious process of representing one’s own experience. To give an account of oneself, or of each life in its particularity, Judith Butler would say, is to engage in a continuous process of “doing and undoing” the human, in order to open a space for the recognition of the particular unfolding of each life; a never ending process in which it is impossible to give a final word, or to grasp a source of authenticity, once and for all. A process which is collective and individual, since dominant or innovative representations of human experience, of what is important to maintain and let flourish each human life, are determined by each single repetition of those shared or alternative representations. A process which is, finally, according to many feminist thinkers intrinsically relational, as we define ourselves only in relations with others, as Gilligan had also recognised.

Each singular identity, its needs and characteristics, will emerge in the interplay between the permeability and resistance of the inner material and relational drives to language and constraints. In other words, one’s own subjectivity is defined only temporarily and partially in this play of resistance to – or adoption of – given descriptions, a play in which the different levels of corporeal

---

77 Similar considerations, though in a different philosophical context, can be found, I would like to suggest, in Laugier’s works on ordinary and care ethics.
materiality, psychic dynamics, interpersonal relationships and socio-cultural dimensions, all have a role that can be in turn positive and negative for one’s own well being.

What is interesting here, and this is the core issue of this paper, is the question – already reiterated a few times – of whether the consideration of this radically multilayered, vulnerable and fragile subjectivity can be accommodated in an understanding of care ethics, or – better – how it can be accommodated, since I think that this is possible, and that care ethics is one of the chances we have to envisage an ethics for such subjects.

My suggestion is that care ethics should be more clearly articulated in the light of this complex framework. It is clear, for instance, that viewed in this way talking of “feminine virtues” no longer makes sense, while it is still possible to think of care and solicitude or attention, even in the form of virtues, although their meaning should be developed on the basis of the abovementioned elaboration. From this perspective, in a nutshell, the core ideal of care ethics of meeting the others’ needs, or sustaining the web of relationships in which we are all enmeshed, should be seen in a more complex way than as just an appeal to the kind of attention everyone has been made the object of as a child. This is, by the way, what I have suggested above in giving my account of Gilligan’s work.

If then relatedness, interdependence and vulnerability are considered as the fundamental dimensions of humanity, dimensions in which care ethics finds it roots and to which it offers guidance, (as for instance Held claims when she writes: “It is the relatedness of human beings, built and rebuilt, that the ethics of care is being developed to try to understand, evaluate and guide”), then they should be articulated along the lines of the same radically complex understanding considered above.

Yet, this is not a simple question.

Given the intrinsic limits of our capacity to apprehend our own needs and those of the other, as we have seen, not only may universalist approaches encounter difficulties, but also some alternative accounts, and some versions of care ethics.

Actually, considerations such as those developed in recent feminism do not constitute a difficulty only for universalistic or impartialist ethics, which are blind to human particularity and interdependence, as many care ethicists and many other critics of universalist accounts of ethics have claimed; they are also difficult to deal with from within any moral model which, while putting at its centre the agent’s responsibility to be solicitous toward others’ needs, assumes in a rather unproblematic way the agent’s (psychological, sentimental or dialogical) capacity to recognise and meet those needs. In this sense, the idea of considering care ethics as an ethics based on the possibility of meeting others needs, letting them and us,

---

78 Held, The Ethics of Care, p. 30.
together with our shared world, flourish in relationships, might be weakened —
even when this effort is limited to particular others who are closer and more
connected to the agent.

If even one’s own identity and needs, one’s own feelings of pleasure and pain,
are accountable only against the background of a series of linguistic, historical,
socio-cultural, psychic and relational ties, it might seem impossible to meet the
needs of others in their particularity, at least in a simple way: for example, by
assuming the existence of a common grammar of pleasures and pain, or a common
structure of functionings and needs, or an easy and direct way to communicate on
a sentimental or verbal or corporeal basis. In other words, it might be thought to
be impossible to overcome the intrinsic limits of our own and the others’ ability to
speak, think, know, or even feel and imagine.

Yet I think that care ethics can evade this problem, but this is true only if it is
accurately designed in its forms, or — to use the vocabulary of virtue ethics — in its
identification of traits of character considered as virtuous, or else in its definition
of the architecture of the virtuous character.

In the next paragraph I will try to offer a few suggestions, and some open
questions, on the way in which it might be possible to read care ethics in this light
(still relying upon some feminist suggestions, but also going back to the debate on
care ethics). The basic idea I will focus upon is that of transforming the limits of
one’s own accountability and that of the other, recursively, in a positive effort to
meet the others’ needs, or better, to keep and maintain the web of relationships
which nourishes life.

4. Feminist virtues? Open questions and tentative conclusions

Resisting the temptations of a nihilistic or a relativistic conclusion based on the
implications of the just described feminist parabola, different feminist authors – in
different contexts – have tried to envisage an ethics that is able to take into
account these implications and yet to offer some normative indications. In many
cases the ethical dimension is even seen as a necessary supplement to politics, as it
allows — notwithstanding all the difficulties — to recognise the singularity of each
life and to open the possibility of interconnection and encounters, as permitted
only by the concrete and embodied personal involvement experienced in
relationships.

In fact, even when converging with post-modern or critical theories, and even
appreciating the difficulties of defending shared aims, many third wave feminists
have claimed that, although epistemologically weakened in its theorisation of the
subject of politics and morality, feminism is characterised by, and should not
abandon, the aim of transforming society, cultures and ways of living. And this in
such a way as to cope with, ease and take care of the vulnerability and frailty of
humans, of forms of coexistence and of the planetary life. In this sense it should be able to give shape to an ethics and a politics.

In this view, one of the shared convictions is that this aim can be maintained and fostered, not by finding a way to overcome the instability, partiality and variety of (feminist) categories, but, on the contrary, by considering (the awareness of) the instability, partiality and variety of categories as a resource for politics and ethics.

Interesting attempts to overcome such an impasse, are for instance those put forward by Rosi Braidotti or Judith Butler, despite the difference in their elaborations, or by Gayatri Chakravorty Spivak in the context of her post-colonial feminism.

Broadly speaking, in their reflections, they distance themselves from conceptions of ethics which are mainly defined in terms of the activity of giving judgements on others’ behaviour or proposing norms and rules for the behaviour of all, and also from attempts to find a shared vision of the good according to which we should live, or upon which we should base our benevolent attitude toward the other. They are even more critical of ethical models taking the shape of the defence of one’s own or universal rights. Conversely, in their own reflections, ethics is conceived mainly as an exercise operated upon one’s own mind in order to recognise the other/s, that is, the other/s to whom we are continuously exposed and connected in the context of concrete relationships, whilst at the same time not forgetting that one’s own mind is porous to relational, social and symbolic environments.

From this perspective, even if the arguments put forward by these authors are different, what is crucial is the capacity to criticise and to abandon one’s own prejudices, and, to pay attention to and to deconstruct what one takes for granted when trying to meet others and their needs, or when defending one’s own rights. And it is important to note that this critical attitude, that is to say, the possibility of overcoming prejudices and the possibility of meeting the others, should not be seen as a search for convergence on the basis of shared or general comprehensions of humanity, or on fundamental values; on the contrary, the aim of this exercise is precisely that of the representation and maintenance of differences, which alone can render it possible, in an encounter, to communicate and to collaborate. What is required, then, is not the exercise of detachment from one’s own partiality to gain some fixed and stable standpoint from which to look at one’s own and the other’s suffering or flourishing; what is required is rather an exercise of humility and of awareness of one’s own limits and partiality, which can be thought of either as emerging from the interplay of the differences which inhabit our own multilayered or fractured subjectivity, or from the experience of difficulty itself, from the feeling of puzzlement experienced in encountering others we are not able to understand, or even from our resistance to imagining ourselves as that other, or as bound up with that other, to whom we are nevertheless linked. It is in this feeling of the
impossibility of a meeting, Spivak would say, that, paradoxically, the possibility of opening up our minds so as to meet the other lies.\textsuperscript{79}

Crucial for an adequate conception of morality, according to these thinkers, is therefore an awareness of the limits and of the partiality of one’s own feelings, categories and norms. Feelings, categories and norms which are nevertheless necessary to the having of a morality, but which at the same time cannot be considered immune from limits. The suggestion is then, neither that of surrendering to the partial nature of these feelings, or categories, and ending up in a relativist position, nor that of insisting on searching for a way to overcome their partiality completely through convergence on presumed shared understandings, but that of being able to transform these limits themselves into a positive moment for morality, using the awareness of the limits as a way to transform ourselves and open up to the possibility of meeting the other, of communicating, to achieve at least the temporary possibility of understanding each other. Morality, it has been said, resides more in this process of self-transformation and reciprocal transformation than in the possible results; in this sense it is an open ended process.

According to all these authors, the moving force of such a process is not to be found in (abstract) knowledge or in forms of reasoning: but rather in the concrete experience of encounters with others and in the concrete experience of our own resistance to given definitions. It stems from the practice of “positioning” oneself (namely, the practice of recognising – as far as this is possible – the specificity of one’s own position),\textsuperscript{80} as well as from the feeling of one’s own vulnerability and of interconnection (that is, need and desire of the others)] in the end, from emotions, feelings and imagination. These are very intriguing suggestions, I think, and worth taking into consideration in a reconfiguration of care ethics.

It is clear that the philosophical frameworks in which many of these suggestions are developed are often derived from philosophical paradigms rather distant (to say the least) from those characterising the debate on care ethics. Accounting for them in this context would therefore require detailed analysis, not possible here, but I think, nonetheless, that it might be possible to find points of contact and to translate the main indications which are suggested – mostly that of


a recursive use of partiality and instability – into the language of philosophical accounts more traditionally associated with care ethics.81

Of course it is not the aim of this paper to provide a detailed account of such a reframing, its main aim was in fact more that of indicating a problem, than of solving it. Nevertheless, I will try to offer some suggestions, and open up a few further questions.

In fact, in so far as one wants to consider care ethics as a feminist approach to ethics, or to invoke feminism to justify certain claims concerning the importance of care in an account of morality (and possibly in order to consider care ethics, in more general terms, as an adequate account of morality), there is only one point that – in my opinion – might be proposed as a result of the analysis developed in the previous pages, and this is that one should take into consideration the abovementioned implications of the feminist parabola; that is to say, one should consider the different levels involved in our attentiveness to others and to the web of relationships in which we are (all) enmeshed. In other words, the different levels – bodily, psychological, social and symbolic – in which our own and others’ vulnerability unfolds should be taken into account. Furthermore, in this light, it might be interesting to consider the feminist suggestion of a positive role for the awareness of one’s own limits, that is, the awareness of the instability of each viewpoint and of personal landscapes - however enlarged and rendered accommodating through imagination, attention to details or through listening and dialogue - since biases, misunderstandings and opacities are always present.

Going back to title of this paper and playing with it, it might be said that if we consider recent feminist insights concerning the value of the awareness of one’s own point of view and the partiality of one’s own standpoint in morality as giving shape to a form of “feminist virtue”, as a kind of epistemological virtue that has to do with instability, critique and humility, then such a “feminist virtue”, more than the old “feminine virtues” (of nurturing, caring for the household, etc.), should be put at the centre of a sound care ethics. This is the main point here. It is in this sense, this is my opinion, that care ethics should be considered as a form of feminist ethics.

With regard to the possibility of a major articulation of such a reframing of care ethics, I will offer only some tentative final considerations and a few suggestions.

As I have already argued, we can retrace as far back as Gilligan’s work the idea of care ethics as an ethics mainly centred on the cultivation – as a form of personal moral responsibility – of a form of sensitivity or attentiveness to the vulnerability of others and to our own, at different levels. Even in Gilligan’s writings, in fact, caring means something more than simply being porous and responsive to the suffering of others at what is seen as an easily apprehensible

81 I have argued in more detail for a similar reconfiguration in Botti, Prospettive femministe.
physical or psychological level; that is to say, a suffering that – using a figure already hinted at – might be reduced to a common or universal grammar of pains and pleasures. Gilligan herself has underlined that caring – as an ethical ideal or practice – also involves caring and being attentive to the violence which derives from one’s not having been recognised as a person, from not having a voice or not being considered fully human, or from lack of recognition of the plurality and the opacity of the grammars of pains and pleasures. And it is in this sense that I have suggested that her rather complex gesture of recognising a moral voice to girls and young women shows the caring attitude.

Viewed from this angle therefore, there is a possible convergence between the tenets emerging from recent feminist stances and the development of an ethics of care.

Of course, if the major tenets of care ethics are those of considering, as the most appreciable attitudes or traits of character, awareness of the relational fabric which binds together humans and personal sensitivity and solicitude toward the others’ flourishing (or to the flourishing of given relational contexts and of the web of relationships sustaining life), then something should be added to define this sensitivity and solicitude. Here are some few suggestions.

In order to offer a more philosophically structured account of an ethics of care, care ethicists should not only – this is a first point to be made – make reference to sympathetic or empathic psychological mechanisms, but also – as many refined sentimentalists do suggest – to the fundamental human capacity of imagination. Imagination, that is meant as a faculty able to transform and widen our sensitivity to the suffering and needs of others, or of the poverty or richness of a relational context, and to open up our mental landscape and overcome the limits of those same psychological mechanisms, so as to render us able to recognise even the small details relevant for (or in contrast with) the flourishing of given relationships and single lives.

In this connection, and this is a second point worth mentioning, if we are trying to reconfigure ethics in the light of the considerations put forward here, such an account of ethics must be reframed to deal with the texture of our lives in their concrete and particular aspects. From this perspective, it is important to stress that even minimal actions or interactions, and not only major dilemmas such as those concerning, for example, which lives to save (often considered in the analytical tradition of impartialist morality), should be considered as of moral relevance. Even small gestures, or the words we use in our everyday life, become important if contrasting dominant oppressive representations of humanity – i.e. of single human beings – is to be part of the framework of our ethical attitude. Reflecting on the kind of images of humanity that are hidden in our judgments, in our behaviours, in our language, even in apparently innocuous contexts, is in fact important, in that – as has been argued – oppressive or less oppressive accounts of
humanity are made or changed through each single repetition of these representations.

Finally, in reflecting on the connection between collective and individual dimensions of these processes, it is also possible to consider whether part of our moral responsibility is to contribute to a collective effort to produce representations and images that convey a larger sense of the possibility and variety of human forms of life. In this way we may nourish our imagination; of course not regarding those representations and images as definitive, but simply as enlarging, and not exhausting “the realm of the possible”, to say it with Pindar. In this sense even establishing an alternative lifestyle might be considered as invested with moral responsibility.

Of course, neither the general claim concerning the role of imagination, nor the few considerations above may seem particularly innovative in the contemporary arena of philosophical reflections on morality, although they are not so present in the mainstream debate concerning care ethics. In effect, there have been many attempts to envisage, for instance from refined sentimentalist points of view (e.g. those which, in the wake of Hume’s ethics, focus on the possibility of returning reflectively on one’s own first level sentiments), similar accounts of the role of imagination, attention to details, collective responsibilities, and narratives and dialogues, all viewed as means able to enlarge the possibility of a mutual understanding and of solicitude toward the others, notwithstanding the contingency of our categories. I think there is room here for interesting consonances. But interesting consonances can also be retraced with the other philosophically more structured developments referred to above as potentially interesting for reframing care ethics. These issues can be dealt with, for example, from within perfectionist accounts centring on the importance of the effort of transforming oneself and one’s own images of humanity and of paying attention to “what is important” for oneself and for others. There are also numerous possibilities of exploring the role and the importance of communication, dialogue and narratives, as these too are fundamental resources to which one could appeal to in order to develop an interesting reframing of an ethics of care. I will not comment further on these consonances here; instead I will concentrate upon a single final aspect, characteristically feminist, which – in my opinion – is worth to be taken into account in order to provide a better account of care ethics (and eventually in order to develop it in tune with sentimentalist or with perfectionist approaches).

As we have already seen, feminist thought offers, in my reading, a final and different suggestion. This is the above-mentioned idea of considering as an ethical attitude, or a resource for ethical attitudes, the importance of developing an awareness of our own limits in being able to feel, to sympathise and to imagine, and also to listen and being attentive. This is something which is distinctive in recent feminist thought, and is less represented, at least in my opinion, in
sentimentalist and perfectionist developments, and even less present in the mainstream debate on care ethics. Both approaches focus more on the positive possibility of expanding one’s own point of view or capacity of attention, while here the emphasis is, on the contrary, on being aware of the limits of this possibility, as a recursive form of attention.

All these instances – sentiments, feelings and sympathy, imagination and attentiveness, communication and dialogue – are in general seen as important resources that may serve to widen our comprehension or experience, and therefore enable us to participate in the suffering of others and to correct our egoism and biases; in a word, to allow us to be in contact with the other in a positive sense. As such, they are considered as a means to add something to our mental landscape that will allow us – without recourse to reason or to a metaphysical understanding of human nature and values – to acquire a more stable and wider point of view from which to decide how to act and live morally, or from which to approve or disapprove our own behaviour or that of others, to make our ethical judgements. Conversely, feminist attention in the reading I have tried to outline here is more dedicated to the partial nature of our achievements in this sense. In fact, as we have seen, feminists have cast doubts on the possibility of achieving fully positive results through these resources, although these resources are the only ones we have. So one can wonder if, together with the idea of using all these resources to enlarge our sensitivity, one might also envisage a different use of them, so as to render continually present to us the limits of that same sensitivity, also - recursively and paradoxically - as a way to foster it.

What I am suggesting is the idea that the caring individual should cultivate not only his/her sensitivity, understood in the more traditional terms seen above (i.e. developed through psychological mechanisms, sentiments, imagination or communication), but also cultivate an awareness of the instability of his/her mental contents or feelings; that is, one should keep open the aching sense of one’s own limits, while seeking out all the possible ways to overcome them.

Such an awareness of our limited capacities might be thought of as a resource moving us toward the recognition of the tentative nature of our doings. This is not to deny the urgency we may feel to act, but to question the certainty of the results, thus rendering us more careful, in the continuous search for new ways to maintain the relationships we are in, and allow them to flourish. We can think of it also in the shape of a feeling of puzzlement. Such an awareness may play several important roles in the genesis of moral attitudes: as the engine of imagination; as a resource which pushes us to suspend our judgemental attitudes without suspending our care for others, and to engage ourselves in the effort of searching for more adequate judgments; as a drive for caring about the way we care, but also as the resource which pushes us, all things considered, to continue to care, to be interested in the wellbeing of others, as it is linked to a shared sense of vulnerability.
Coming to a characterization of this awareness, two very tentative suggestions can be offered. The first is that, as already underlined by many feminist theorists, it should not be viewed as resulting from knowledge or abstract reason, but from experience and feelings. It can thus be thought of as a form of sentiment, or as a sentimental progress or reflection upon more basic sentiments.

A second and more interesting idea, I think, is that concerning the possibility of linking this awareness to the passion or sentiment of humility, thinking of this passion as an architectural bastion of the moral character (in opposition to the often quoted passion of pride). A possibility which might be of interest, for instance, for anyone wishing to articulate a sentimental reading of care ethics.

In conclusion, in drawing together the threads of this discourse, I would say that to bring care ethics back to its original feminist matrix, but at a higher level, an attempt must be made to read care ethics as a reflection on morality centred both on care for others in their particularity, and on the recognition that such care implies the questioning of pre-established views of both self and the other. Care for others, attempting to meet their needs, has its roots in a sentimental and relational characterization of subjectivity, but there may be obstacles in the way of such care, in the form of distances, or of opacity of vision. Though it may never be possible to fully overcome one’s self and one’s limits in this sense, the attempt to do so, to discover and engage with one’s limits, and become more clear-sighted, should be explored as a dimension of the ethics of care.
MONOGRAPHICA II

_The Legacy of Bernard William’s Shame and Necessity_
Guest Editor’s Preface

Alessandra Fussi
Università di Pisa
Dipartimento di Civiltà e Forme del Sapere
alessandra.fussi@unipi.it

The idea of this monographic issue originated in a workshop that took place at the University of Pisa on October 3, 2014 on the legacy of Bernard Williams’s *Shame and Necessity*. The goal is to reflect from different perspectives on one particular problem on which Williams’s book gave a decisive contribution, namely, on the nature, history, and moral relevance of shame in contrast with other emotions such as guilt and fear.

*Shame and Necessity* was published in 1993 as a result of the Sather Lectures that Bernard Williams gave in 1989 at the University of California, Berkeley. When the book came out it attracted the attention of both philologists and philosophers. With a polemical stance towards what he called the “progressivist view,” Williams proposed a philosophical critique of some fundamental moral concepts in light of what he found illuminating in the Greek ethical world.

By “progressivism” Williams meant to refer to the view, accepted by several scholars, that in ancient Greek literature, and especially in epic and tragedy, one would find an essentially primitive conception of psychological motivation, of action and of responsibility. According to such interpretation, the Homeric world lacked the distinction between body and soul, without which one could not attribute responsibility to an agent for his or her intentional actions, and conceive of moral action as based on the will. In turn, the tragedies were populated by heroes at the mercy of divine forces, and subject to forms of necessity that were clearly incompatible with free will and autonomy.

The simplest way to express the contrast between the Greeks’ ethical world and the modern moral conception is to claim that the advent of Christianity

---

1 Bernard Williams, *Shame and Necessity*, Sather Classical Lectures, Volume 57, University of California Press, Berkeley 1993. From now on *Shame and Necessity* will be abbreviated as SN.
allowed for progress from a culture of shame (based on appearance and fundamentally heteronomous) to a culture of guilt (based on interiority and the autonomy of the will).

In itself Williams’s polemical stance towards progressivism qua historical interpretation would not have been very interesting if it had not been part of a very complex project. Indeed, in *Shame and Necessity* one can see several interrelated issues being addressed:

a) Accuracy in the history of thought. The question is how to develop a view of the past free from a prejudice in favour of modern concepts and at the same time also immune to the temptation of believing in the possibility of an absolute point of view.

b) The Greek conception of ethical and psychological phenomena. This includes the question of whether Homeric characters are capable of deliberation, how they ascribe responsibility, if and how they act on the basis of their deliberations, how Homeric and tragic characters see themselves in relation to their actions and intentions.

c) Necessity and shame. When the Homeric and tragic characters affirm certain actions as necessary, they are neither thinking in terms of hypothetical imperatives, nor autonomously choosing on the basis of categorical imperatives. What kind of necessity are they invoking when, like Ajax, they affirm that there is only one path that remains open to them and they claim that they must take it? How does shame work as a motivation for action? On the one hand the Greek characters who appeal to shame seem to express concerns which involve their sense of personal identity (performing certain actions appears to them as a betrayal of who they really are), while on the other hand shame is also a response to concrete social expectations. How are we to understand the relationship between activity and passivity, autonomy and heteronomy in the specific case of shame?

d) Moral concepts belonging to the contemporary philosophical discourse (such as will, autonomy, spontaneity, moral responsibility, guilt, regret). Are they as clear as the progressivist view claims they are? Can they withstand criticism when contrasted with alternative explanations of actions and emotions as they emerge from an analysis of Greek literature that does not take those very concepts for granted?

Because the answers to these questions are dependent on each other, *Shame and Necessity* can function as a critical reflection on contemporary ethics because it is at the same time a reflection on paths that ethics never took or left behind.

Williams found such paths hidden behind the ideological reconstructions provided by the progressivist view. In order for his project to work he had first to show that there was a peculiar combination of theoretical, historical and philological problems in the premises governing the distinction between the Greek culture of shame and the post Christian culture of guilt. He concentrated on the
shortcomings of progressivism and argued that it led to misunderstanding some major ethical aspects of Greek life. Nietzsche’s influence is evident in his critical reflection on the relationship between history, philosophy and literature, which sets the premises for a philosophically interesting and philologically informed reading of the texts.  

The ultimate goal was to show that the most common explanatory concepts employed in contemporary moral philosophy were not adequate. Ascribing responsibility is not just a matter of recognizing intentions. It involves different ways in which an agent can be identified with or dissociated from his or her actions, the interaction between what someone does and what just happens to him or her, and the complex web of personal and social expectations that make one emotionally respond in certain ways to certain actions.

The cultural gap between the Greeks and us is significant if we consider contemporary and ancient theories of action, while it becomes considerably smaller if we lay bare the structures that actually govern our behavior. To accomplish this Williams moved beyond theory to literature and legal texts. There is a difference, he argued, between “what we think and what we merely think that we think” (SN, 7), and he thought that his reflections on Greek literature helped uncovering it. This is why he explored the ways of attributing responsibility in the Anglo-American legal system, rather than concentrating only on contemporary philosophical moral theories, while with respect to the ancients his attention was devoted mainly to Homer and the tragic poets (with a preference for Sophocles) rather than to philosophers like Plato and Aristotle.

By focusing on how the epic and tragic characters deliberate, blame themselves or are proud of what they do, Williams challenged the most common views concerning the similarities and differences between the Greeks and us. This in turn helped him show the gap between our theories and the attitudes we ordinarily have towards our actions:

If our modern ethical understanding does involve illusions, it keeps going at all only because it is supported by models of human behavior that are more realistic than it acknowledges. It is these models that were expressed

---


4 Cfr. SN, 14: “Most of the texts that I consider in this book do not even look like philosophy, and my aim is not to make them do so. Tragedy, in particular, is important to many of the questions I want to ask, but its importance is not going to be discovered by treating it as philosophy, or even, rather more subtly, as a medium for discussion that was replaced by philosophy. By the same token, to point out the obvious fact that these plays are not works of philosophy tells us nothing at all about what their interest for philosophy might be.”
differently, and in certain respects more directly, in the ancient world. In these relations there is [...] a two-way street between past and present; if we can liberate the Greeks from patronizing misunderstandings of them, then that same process may help to free us of misunderstandings of ourselves (SN, 11).

Since his turning to the history of thought was motivated by theoretical questions, one interesting consequence of the project Williams undertook with *Shame and Necessity* was that the book spoke to analytic and continental philosophers alike. Williams was impatient both with analytic jargon not informed by historical sensitivity, and with a philological practice not conscious of its philosophical premises. With respect to philologists and historians his critical attitude is obvious in *Shame and Necessity*, but if we turn to “Philosophy as a Humanistic Discipline” we can see very clearly the kind of criticism he leveled against some argumentative habits common among analytic philosophers:

Paul Grice used to say that we ‘should treat great and dead philosophers as we treat great and living philosophers, as having something to say to us.’ That is fine, so long as it is not assumed that what the dead have to say to us is much the same as what the living have to say to us. Unfortunately, this is probably what was being assumed by those who, in the heyday of confidence in what has been called the ‘analytic history of philosophy’, encouraged us to read something written by Plato ‘as though it had come out in *Mind* last month’—an idea which, if it means anything at all, means something that destroys the main philosophical point of reading Plato at all. The point is not confined to the ‘analytic’ style. There is an enjoyable passage by Collingwood in which he describes how ‘the old gang of Oxford realists’, as he called them, notably Prichard and Joseph, would insist on translating some ancient Greek expression as ‘moral obligation’ and then point out that Aristotle, or whoever it was, had an inadequate theory of moral obligation. It was like a nightmare, Collingwood said, in which one met a man who insisted on translating the Greek word for a trireme as ‘steamship’ and then complained that the Greeks had a defective conception of a steamship.

Of course, *Shame and Necessity* had such ambitious aspirations that it was exposed to criticisms from each specialized field it touched upon, and twenty-two years after its publication it is not easy to ascertain whether it had the impact it deserved on ancient philosophy scholarship.

Bernard Williams has often been accused of being primarily a destroyer of theories rather than someone willing and capable to offer alternatives. The scathing irony with which he sometimes addressed theories he found inadequate gave the impression (to some) that he was fundamentally a skeptic who enjoyed dismantling anything that looked like a system (and most famously what he used
to call the morality system).

The characterization of Williams as a skeptic is mostly correct. Nonetheless, he had seen something crucial. There was indeed a problem that he approached from different points of view, a source of confusion which he tried to warn us against in quite a few of his writings. What I take to be his fundamental insight can be expressed negatively, though it was meant to have positive consequences: philosophy ought not to aim at total awareness. He took this point to be important both with respect to historical accuracy and with respect to ethics. What had to be avoided by the historian was the mistake of thinking that by giving up the belief in an absolute standpoint one was going to surrender historical accuracy altogether. He argued for this idea several times. One clear example was his reply to Putnam in “Philosophy as a Humanistic Discipline”:

But the second assumption I have ascribed to Putnam, that if there were an absolute conception, it would somehow be better than more perspectival representations—that is simply false. Even if it were possible to give an account of the world that was minimally perspectival, it would not be particularly serviceable to us for many of our purposes, such as making sense of our intellectual or other activities, or indeed getting on with most of those activities. For those purposes—in particular, in seeking to understand ourselves—we need concepts and explanations which are rooted in our more local practices, our culture, and our history, and these cannot be replaced by concepts which we might share with very different investigators of the world. The slippery word ‘we’ here means not the inclusive ‘we’ which brings together as a purely abstract gathering any beings with whom human beings might conceivably communicate about the nature of the world. It means a contrastive ‘we’, that is to say, humans as contrasted with other possible beings; and, in the case of many human practices, it may of course mean groupings smaller than humanity as a whole.5

In ethics the impossibility of total awareness is due to the fact that we always think from dispositions. What we see, the way we see it, our best insights come from our dispositions to see things in certain specific ways, to illuminate our actions according to priorities that we receive from the world and cannot undo (clearly an Aristotelian thought). This is how Williams states this point in the “Postscript” to Ethics and the Limits of Philosophy:

I hope I have made it clear that the ideal of transparency and the desire that our ethical practice should be able to stand up to reflection do not demand total

explicitness, or a reflection that aims to lay everything bare at once. Those demands are based on a misunderstanding of rationality, both personal and political. We must reject any model of personal practical thought according to which all my projects, purposes, and needs should be made, discursively and at once, considerations for me. I must deliberate from what I am. Truthfulness requires trust in that as well, and not the Obsessional and doomed drive to eliminate it. 6

In Shame and Necessity the warning against the temptation of absolute transparency takes the form of a defense of shame, both as a moral emotion and as a distinctive trait of Greek culture. Shame is contrasted with guilt, which is normally associated, especially by the progressivists, with a more developed moral sense, based on psychological depth, autonomy and self-sufficiency.

In the Greek conception of shame Williams sees preserved some fundamental aspects of guilt, but without the virtues that progressivism attributed to guilt. Shame looks at who I am, and for this reason it encourages attention to what I did or omitted to do, like guilt. But, in contrast with guilt, shame keeps me more strongly connected with others: when I feel ashamed I am not just concerned with what I did to someone, but also with what the action reveals about me, and, as Williams argues, this question cannot be properly addressed without taking into account how others I respect would react to it. (The other of shame can be concrete or internalized. Fussi and Montes Sánchez address this point in the articles they wrote for this issue).

From the point of view of shame I can be critical of the social expectations to which I am responding, but I cannot sever all ties with them: I cannot go so far as to consider my reflections as stemming from a pure form of rationality, or to consider myself free from my received dispositions of character. Williams’s warning to those who take this route is that they run the risk of motivational solipsism (SN, 99-100). His retort to those who accuse the Greeks’ reliance on shame of leading to selfishness on the one end and to heteronomy on the other hand, is to point out that this kind of critique ultimately leads to contradictory requests concerning the role others should play:

When it is complained that the Greek ethical outlook, or at least that of the archaic Greeks, is both egoistic and at the same time heteronomous, because it rests conventionally on the opinion of others, there is a constant and powerful tendency for these two complaints to turn against each other. Which is supposed to be the trouble, that these people thought too much about others' reactions, or too little? (SN, 100).

If I understand him correctly, Williams thought that shame forces us to acknowledge that there is always a place and a time from which we think. The fact that shame is bound to local practices and expectations makes it ultimately opaque to reason, but this limitation has the important consequence that it keeps philosophical thinking away from embracing illusory hopes. It drives us to realism.

Williams admired Plato enormously, but he thought that precisely because Plato had not given shame the place it deserved in his psychological theory, he believed that philosophy could or should be a liberation from humanity. As others before me have noted, this accusation was misguided. Actually, shame played a central role in Plato’s ethical and psychological reflection, and this is true both from the point of view of the arguments explicitly defended by Socrates, and from the point of view of the behavior exemplified by him in the dialogues.

On the argumentative side one should consider the importance of the political and ethical role attributed to spiritedness in the *Republic*, since that part of the soul plays a key role in education, involves social recognition and is responsible for emotions like shame, pride, anger and indignation. On the dramatic side one should mention that Thrasymachus’s blushing is clearly a turning point in the first book of the *Republic*. Furthermore, as Laura Candiotto shows in the paper she presents in this volume, the shame felt by the different characters in the *Gorgias* contributes not just to the efficacy of their refutation, but involves the audience as well. That Williams’s criticism of Plato was not correct of course does not make his connecting the recognition of the importance of shame with philosophical realism less worthy of reflection.

The essays collected in this volume take their bearings from Williams’s discussion of shame, both from the theoretical and the historical point of view. The discussion concentrates on the nature and phenomenology of shame, and on its relationship with emotions like fear or anger, and virtues like courage. The papers follow two main lines of inquiry: two of them (written respectively by Alessandra Fussi and by Alba Montez Sánches) focus on Bernard Williams’s account, while for the last two papers (by David Roochnik and by Laura Candiotto) *Shame and Necessity* is the background for further inquiries on the role of shame in ancient thought.

Alessandra Fussi concentrates on Williams’s defense of the Greek conception of shame against some commonly held critical views: 1) that shame expresses selfish concerns and it ignores the needs of others; 2) that it belongs to the so-called objective attitudes and inhibits second-person responses; 3) that it is dependent on social conventions and is therefore both superficial (since it mostly relies on appearance) and heteronomous (since it depends on external values). In the last part of her paper she concentrates on the role of others in shame. She examines the problem of the other’s attitude (do we feel shame only when others are critical of us?), and finally turns to Williams’s distinction between the concrete
Precisely the notion of the internalized other is the main focus of Alba Montes Sánchez’s contribution. She examines in detail Williams’s position in *Shame and Necessity*, contrasts it with recent accounts of shame and finally presents and endorses a Sartrean view. In order to clarify the role others play in the feeling of shame Montes Sánchez initially differentiates shame from embarrassment and disgrace. In the second part of her paper she discusses Calhoun’s objection to Williams: if respect for the other’s judgment is necessary for shame, then ultimately I am my own judge, since I cannot respect judgments that I would not myself share. Calhoun thinks this is wrong: sometimes we are shamed by people we don’t respect and by judgments we don’t endorse. In her view, this is not a sign that others have contaminated our value system. It simply shows that shame is not about evaluations; it is about their practical impact on the life we share with others. Montes Sánchez observes critically that Williams and Calhoun are both too restrictive in giving others the power to shame us. It is not true, she affirms, that others cannot shame us unless we allow them to do so by respecting them or entering into shared practices with them. Rather, “everybody has, to a higher or lesser degree, the power to shame us unless we withdraw it from them through contempt or disengagement, for example.” It is on the basis of this view that Montes Sánchez proceeds to retrieve the Sartrean account of shame. The role the other plays in shame is not that of evaluating me, but, more fundamentally, that of allowing me to focus on myself as the object of someone else’s perception. The other is fundamental not *qua* internalized other but, rather, as constitutive of the possibility of inter-subjectivity, and hence of transforming me into a subject capable of feeling shame.

The fact that shame is aroused when we feel judged by people we respect (and not just by anybody who happens to be our witness) is a point that Williams shares with Aristotle, who first formalized it. It is to Aristotle’s account of the relationship between courage and shame in the *Nicomachean Ethics* that David Roochnik devotes his paper. The problem he addresses is the relationship between courage (which for Aristotle is a virtue) and shame (which is a quasi-virtue: shame is an emotion, not a disposition like virtue, but it is instrumental to virtue because it is appropriately felt by young people in the process of becoming good men). Aristotle affirms that there is a difference between those who exhibit authentic courage and those who exhibit only “political courage” (1116a17). While courageous men “act on account of (*dia*) the *kalon*” (1116b30-31), those who act out of political courage (and are not motivated merely by fear of their superiors) face death in battle because they would be ashamed if they did not. The difference between such men and the truly courageous appears very thin, since both are motivated by what is beautiful and noble. Citizen-soldiers, Aristotle affirms, act “on account of virtue (*dî’ aretên*)”. For they do so on account of shame (*aidô*) and a
desire (orexin) for the kalon, since they have a desire for honor, and in order to avoid blame, which is aischron” (1116a27-29). Roochnik distances himself from Irwin’s interpretation, for whom the difference rests on the fact that only the truly courageous are altruistic. In facing death they consider the good of others and nothing else, while the citizen-soldiers aim at their selfish interest: they seek honor and the glory of a noble death. According to Roochnik, in Irwin’s interpretation the kalon is unduly moralized, and it loses the connotation of visibility that is essential to it: something kalon is not just noble, but beautiful and fine. Roochnik proposes an interpretation that takes into account two points made by Williams: 1) The internalized other is essential to shame. 2) Shame makes us relate not just to others but to ourselves.

If we remember that shame is an appropriate emotion for young people, we can imagine young soldiers motivated by a desire to be appreciated by their fellow soldiers, by their superiors and by the citizens at large. They act nobly on account of virtue because they desire their deeds to shine in glory. When they get older they will have internalized the sense of the kalon for which it is worth fighting. At that point their actions will not just be altruistic (virtue for Aristotle is never disjoined from a concern for one’s own happiness), but motivated by a sense of such actions as are appropriate to the kind of people they deem to be. Acting for the sake of the noble will mean, at that point, to act in such a way that one can see himself as kalon and hence maintain a sense of his own identity as the kind of person who does certain things and would blame himself for doing something aischron. Ultimately, in Roochnik’s paper we can see an Aristotelian developmental account from the capacity to feel appropriately ashamed of oneself to that of feeling appropriately proud of oneself.

The volume ends with Laura Candiotto’s paper on the role of shame in Socrates’s refutation of his interlocutors in Plato’s dialogues. More specifically, Candiotto observes that in the dialogues that portray conversations with important public figures (for example politicians, sophists, generals or rhetoricians), Socrates’ interlocutors react to their refutation with two alternative attitudes. They either accept the shame of a defeat in argument as an emotional recognition of ignorance (as is the case of Charmides in the dialogue that takes his name), or they try to hide their feelings (as does Critias in the same dialogue, or Callicles in the Gorgias) for fear of losing face. In this second instance, the audience plays the role of an extended mind. Those who are present at the refutation become aware that Socrates’s interlocutors are unable to acknowledge their failings, and this is for the audience an occasion to see through the social masks exhibited by these public figures and realize what kind of people they really are. That the refutation reaches out to the audience has two consequences: 1) The audience is allowed a critical stance with respect to public figures who tend to use rhetorical defenses in order to hide their shortcomings. 2) Those shortcomings, once perceived by the audience, bounce back to the refuted interlocutors, who are
therefore forced to revise their self-image.

Candiotto’s conclusion is that the extended elenchus involves cognitive and emotional means. Its visible aspect is aporia, which is the result of a process whereby Socrates, interlocutor and audience through the mediation of shame realize a form of catharsis. The audience is purified from the superficial opinions it initially entertained about the interlocutors and led to consider the subjects under discussion under a different light. The interlocutors, in turn, receive from the audience the possibility of experiencing the shame they initially refused to accept, and this initiates a process of transformation in their attitude towards truth.

The authors whose contributions appear in this volume approach Shame and Necessity from different perspectives. To some the book reveals important aspects about Williams’s attempt to understand ethics from a genealogical perspective; others find inspiration from the way historiographical categories are questioned. Some are bent on finding connections between Williams’s theory of shame and those of authors who preceded or followed him, while others find in this book the invitation to a dialogue with contemporary theories. Williams used to say that philosophical arguments do not need to form a system; it would be good enough if they hung together, “like conspirators perhaps.” This is an interesting key to Williams’s work. One might hope it can also characterize the essays presented here.

Williams’s Defense of Shame as a Moral Emotion

Alessandra Fussi
Università degli Studi di Pisa
Dipartimento di Civiltà e Forme del Sapere
alessandra.fussi@unipi.it

“Nobody can be vulgar all alone”¹

ABSTRACT
Section 1 examines four reasons most commonly adduced to support the claim that guilt is superior to shame, both psychologically and morally: a) While guilt expresses a concern for others shame is a self-centered and selfish emotion. b) While guilt appeals to autonomy shame is linked to heteronomy. c) Shame is not a reactive attitude, like guilt, indignation, blame, resentment, but an objective attitude, like disdain or disgust. d) While guilt invites us to second-person responses, shame inhibits them. The second part of the paper (sections 2 and 3) addresses Williams’s analysis of the role of shame in ancient Greek literature and philosophy. Section 2 is dedicated to Williams’s response to the objections concerning selfishness and shallowness and to discussing his reply to the charge that since shame belongs to the objective attitudes it tends to inhibit second-person responses. Section 3 concentrates on Williams’s reflections on heteronomy by focusing on the attitude of others in shame and on the role played by the internalized other.

KEYWORDS
Bernard Williams, Stephen Darwall, shame, guilt, objective and reactive attitudes, internalized other

Introduction

It is still a widespread belief in moral philosophy that shame is a more primitive and less reliable emotion than guilt. As Stephen Darwall puts it in a recent essay,² shame and guilt belong to two different spheres of recognition: shame is the typical emotion of honor societies, while guilt is at home in societies in which respect for one another is mediated by a mutually accountable public space. Honor is bestowed and can be taken away by those in a given society who have the power to do so, while dignity is not something that can be taken away or, like honor, diminished by the behavior of others.

Those who consider emotions like shame and guilt from a theoretical point of view often do not distance themselves from the familiar story according to which the ancient Greeks failed to put the concept of the will at the center of their moral theories, and had a conception of human life shaped by a “culture of shame.” This story is supported by the so-called progressive view of historical development, presented in a classic way by Dodds in his book *The Greeks and the Irrational* and by Adkins in *Merit and Responsibility*. Dodds and Adkins detected in ancient Greek culture a progressive if slow detachment from a culture of shame towards a culture of guilt, which was thought to have reached its best articulation with the advent of Christianity. In such historical reconstruction Plato and Aristotle play the role of intermediary figures in the progressive development from shame to guilt: they anticipate some fundamental discoveries concerning human agency and autonomy, while at the same time still expressing in several ways the culture of shame to which they originally belonged.

Contrary to this line of thought, Williams is skeptical of a psychology based on the separation between body and soul (SN, 25-26), the notion of the will as a mental action in-between a decision and the ensuing action (SN, 41-46), the idea that we can be responsible only for actions that derive from our intentions, and the thesis that guilt is more morally relevant than shame (SN, 75-102). He maintains, rather, that psychological and ethical theories can benefit from an understanding of agency that includes agent-regret, moral luck, necessity, and takes into account the role of shame. In his view, if we are open to such concepts we will no longer look down on the ancient Greeks as if they were the representatives of a primitive moral outlook.

Williams is not inclined to historical nostalgia: his point is not that the Greeks were right in their approach to the most fundamental ethical questions. Rather, he refuses to assume with progressivism that modernity, whether in the shape of Kantianism or Utilitarianism, made substantial progress in tackling problems that had supposedly not been properly addressed by classical Greek philosophy and literature. He asks us to distance ourselves from such a theory. He thinks that

---


4 See, Bernard Williams, *Shame and Necessity*, Sather Classical Lectures, Volume 57, University of California Press, Berkeley 1993, 64: “Progressivist writers refer to a concept of moral responsibility that we supposedly enjoy and the Greeks lacked, but it is unclear what they have in mind. Their thought seems most typically to be that the Greeks, or at least archaic Greeks, blamed and sanctioned people for things that they did unintentionally, or again—though this distinction is often neglected—for things that, like Agamemnon, they did intentionally but in a strange state of mind. We are thought not to do this, or at least to regard it as unjust. But if this means that the Greeks paid no attention to intentions, while we make everything turn on the issue of intentions, or at least think that we should, this is doubly false.” From now on *Shame and Necessity* will be abbreviated as SN.
by accepting the possibility that we can learn something from the Greeks we will also be ready to assume a skeptical attitude towards moral concepts that we tend to take for granted.5

Williams’s strategy in Shame and Necessity takes two directions: 1) He tries to bring back to life a view of the Greeks’ ethical life as free as possible from the prejudices of progressive history. 2) He argues that by examining our ethical concepts in the light of the psychological and ethical patterns articulated by the Greeks we can draw different maps of our emotional life.

If we understand well what aidos or aischyne were for the Greeks, our own understanding of the demarcations between shame and guilt will be significantly affected, and this in turn will shake our assurance that we really know what we mean when we characterize ancient Greek culture as a culture of shame.

In section 1 I will consider four reasons most commonly adduced to support the claim that guilt is superior to shame, both psychologically and morally. I will focus on the arguments mentioned by Williams, and, in two cases, on Darwall’s own version of them.

In sections 2 and 3 I will address Williams’s analysis of the role of shame in ancient Greek literature. In section 2 the focus will be on Williams’s response to the objections concerning selfishness and shallowness, and on his position regarding the charge that shame belongs to the objective attitudes and inhibits second-person responses. In section 3 I will concentrate on Williams’s response to the charge that shame, as opposed to guilt, is heteronomous and on his account of the role of others in shame. Here I will discuss three points: a) the distinction Williams finds necessary between the concrete other and the internalized other; b) the attitude of the audience before whom we feel shame; c) the internalized other as the focus of real social expectations.

1. Four Objections to Shame

In Shame and Necessity Williams argues that the opposition between guilt and shame is often rests on ideological preconceptions, but he also concedes that the progressivist position is based on arguments. In order to understand Williams’s

5 SN, 5-6: “These stories are deeply misleading, both historically and ethically. Many of the questions they generate, of when this, that, or the other element of a developed moral consciousness is supposed to have arisen, are unanswerable, because the notion of a developed moral consciousness that gives rise to these questions is basically a myth. These theories measure the ideas and the experience of the ancient Greeks against modern conceptions of freedom, autonomy, inner responsibility, moral obligation, and so forth, and it is assumed that we have an entirely adequate control of these conceptions themselves. But if we ask ourselves honestly, I believe that we shall find that we have no clear idea of the substance of these conceptions, and hence no clear idea of what it is that, according to the progressivist accounts, the Greeks did not have.”
own defense of shame, it may be useful to identify first the criticisms that are most commonly addressed to this emotion. Here are some of the reasons why many maintain that guilt is morally superior to shame:

1) Shame is charged with being a self-centered and selfish emotion. Adkins maintains that in Homer shame serves the competitive virtues of the warrior society. Courage in war, the ability to do heroic deeds, personal success and victory are the qualities with which the best man (the *kalos kai agathos*) is identified. Failure and defeat make someone feel ashamed of himself. Success and victory are of course public values, and this is why being good means, ultimately, to be spoken well of, while being bad is to be despised or ignored. Seeking fame is therefore the main goal, and fear of a bad reputation the central preoccupation. I care for my reputation: as Bernard Williams points out, when he gives voice to the critics of shame, “it is simply my face to save or lose, so its values are egoistic” (*SN*, 78). Guilt, on the other hand, expresses my preoccupation with the sufferance of others, and is therefore considered an other-regarding emotion.

2) Guilt invites us to look into ourselves, to discover our deepest intentions and responsibilities: it is an emotional expression of our being autonomous agents. We feel guilty when others ask us to give an account of our actions or omissions, and if we believe that their indignation or resentment is justified. By contrast, shame is felt when we feel exposed to the wrong people in the wrong situation (*SN*, 78). Since it is a response to how others see us, shame is connected with the idea of losing face, and for this reason it is often charged with being a superficial emotion. Shame gives central stage to appearance and the opinions of others: hence it is the emotion of conformism. So, while guilt is deep, shame is shallow. While guilt presupposes autonomy, shame points to heteronomy.

3) Differently from shame, guilt implies equal footing with others, i.e., our mutual accountability. We do not lose but rather affirm our dignity when we feel guilty: by recognizing that others have a right to consider us accountable, we acknowledge our right to speak in our defense, or to acknowledge our faults. This is not the case with shame. As Darwall maintains,* shame does not belong to a second-person standpoint: it is third-personal. When we feel shame we are not interlocutors of those who call us to respond of our actions, but objects of their gaze. Shame is not a reactive attitude, like guilt, indignation, blame, resentment, but an objective attitude, like disdain or disgust.*

---


imaginary spectators have an *authority over us*, not, as in the case of guilt, an *authority shared with us*. While with guilt we are on an equal footing with the other, with shame we are seen qua inferior. Darwall here follows Sartre:

> By the mere appearance of the Other, I am put in the position of passing judgment on myself as on an object, for it is as an object that I appear to the Other. Yet this object which has appeared to the Other is not an empty image in the mind of another. Such an image in fact, would be imputable wholly to the Other and so could not “touch” me. I could feel irritation, or anger before it as before a bad portrait of myself which gives to my expression an ugliness or baseness which I do not have, but I could not be touched to the quick. Shame is by nature recognition. I recognize that I am as the Other sees me.8

The moment I become ashamed my freedom escapes me, because I become an object for another and at the same time I recognize that I am that object.

4) Typical expressions of guilt are in the second person: confession, apology, reparation are ways to keep the relationships with others alive. Shame, on the other hand, makes us desire to hide: we do not want to meet the other’s gaze, we do not want to reciprocate, we do not feel called to respond, we are not addressed but merely looked at. We feel ugly and despised; we would want to disappear from view. Hence, while guilt invites us to second-person responses, shame inhibits them.

2. Shame revisited

Williams responds to the criticisms leveled against the so-called culture of shame, with which the Greeks are identified, by following a rather complex strategy.

First of all, he questions the stark opposition between shame and guilt and shows that *aidos*, the Greek term commonly translated as shame, can be properly understood only if one realizes that it contains some fundamental aspects of guilt. Secondarily, he addresses the objections to shame as a shallow, heteronomous and selfish emotion, and argues that they are misunderstandings due to a superficial reading of ancient texts. Thirdly, he explains that shame responds to concerns that are wider than the concerns of guilt and, to a certain extent, include them (in some instances, when shame fails to include the concerns of guilt Williams argues that it is for the best).

His response is conducted by analyzing literary passages: from Homer (who was the main target of Adkins’ theory concerning the culture of shame) to the tragic writers. Only occasionally does Williams comment upon classical philosophical texts, and never in great detail. This is one of the main limits of

---

8 J.-P. Sartre, *Being and Nothingness*, 222. See also 261.
Shame and Necessity, and it is especially relevant with respect to Plato and Aristotle, who, as I mentioned earlier, were mostly viewed by progressivist historians as improving upon the original Greek culture of shame by distinguishing between autonomy and heteronomy, interiority and exteriority, conventional or political values on the one hand, and moral values on the other hand.

Perhaps it is because of the priority given to the polemical stance against the progressivist view that in Shame and Necessity one does not find the sensitivity one might have expected from Williams to the significance of the literary aspects in Plato’s dialogues. Williams is clearly aware that the characters and the dramatic setting in Plato’s dialogues deserve keen philosophic attention:

In contrasting philosophy and literature, we should remember that some philosophy is itself literature. Philosophers often suppose that the kinds of difficulties raised for them by a literary text are not presented by texts that they classify as philosophical, but this idea is produced largely by the selective way in which they use them. We should bear in mind how drastically some of these texts are being treated when they are read in this way. [...] One philosopher with whom the cost of these processes is especially high is one who will be relevant to this inquiry, Plato (SN, 13).

Unfortunately, Williams subsequently sets aside the problem of the proper interpretation of the literary aspects in Plato’s dialogues, and in chapter 4 proceeds to raise his criticisms to Plato’s conception of autonomy as if he were addressing a proto-Kantian theory (and one should add that Kant himself is given a far too schematic approach in this book). Williams’s criticism of Plato’s stance towards shame in the example of Gyges’ ring is directed to the theory presented by a character, Glaucon, whom Williams arbitrarily identifies with Plato himself. Yet, Glaucon is only one of several characters in Plato’s Republic. Williams does not seem to appreciate the difference between theories held by characters in a Platonic dialogue and the interpretation of the dialogue as a whole.9

With respect to Aristotle, the problem is somewhat different. It would have been helpful to find in _Shame and Necessity_ a treatment of Aristotle’s conception of shame, with which Williams’s theory shares some fundamental aspects. Aristotle’s discussion of this emotion (the terms he employs are _aidos_ and _aischyne_) can be found in the _Nicomachean Ethics_ and in the _Rhetoric_. It is debatable to what extent the meanings of _aidos_ and _aischyne_ overlap, but, as David Konstan helpfully pointed out, in Aristotle’s ethical works _aidos_ never refers to past events.\(^\text{10}\) _Aischyne_, on the other hand, can refer to past, present and future events, as is evident in Aristotle’s definition of shame in the _Rhetoric_.\(^\text{11}\) The reference to the three temporal dimensions allows _aischyne_ to play a role similar to guilt with respect to actions that one blames oneself for having done, and to function as a prospective and inhibitory emotion with respect to actions that one finds debasing but attractive.

Williams’s analysis of _aidos_ in the Greek world is very close to the Aristotelian view both for what concerns the intersection between the ancient conception of shame and the modern conception of guilt, and for those aspects of shame which embrace ethical phenomena wider than those relevant for guilt. For example, as Aristotle makes clear, shame (_aischyne_) can be felt not just concerning actions and behaviors for which one feels responsible, but also in situations of disadvantage with respect to one’s peers, or when someone is subjected to violence and humiliation. This is a point that from Williams’s perspective can be understood and appreciated in all its importance and it is a pity that it is not explicitly discussed in _Shame and Necessity_, where Aristotle’s conception of shame is only left in the background.\(^\text{12}\)

Let us now turn to Williams’s response to the charges leveled against shame. Points 1) and 2) can be summarized by saying that shame is accused of being a selfish, shallow and heteronomous emotion, especially in contrast with guilt. We can of course imagine the kind of shame felt by someone who is solely driven by a

---

\(^{10}\) D. Konstan, _The Emotions of the Ancient Greeks: Studies in Aristotle and Classical Literature_, University of Toronto Press, Toronto 2006, 95. See also the following observation: “The emotion, as Aristotle understands it, is uniform; what varies is simply the timing of the perceived ills. The lexicographers are thus wrong to split _aischyne_ into sub-definitions, for there is nothing to disambiguate” (99). On the debate concerning the relationship between _aidos_ and _aischyne_ in EN, IV, 1128b15-23, see W.M.A Grimaldi, Aristotle, _Rhetoric II. A Commentary_, Fordham University Press, New York 1988, 105-107.


preoccupation with appearance and the opinion of others. For such a person a critical stance with respect to conventional values will be out of the question, and his feeling ashamed at his failures to adapt to the expectations of others will indeed express his conformism. However, this is not the way shame works most of the time for most people, and it is certainly not how it was meant to work in ancient Greek literature.

Let us begin with the charge of egoism. As we have seen, Adkins claimed that shame in Homer served competitive values, and was therefore a selfish emotion. If this had really been the case, Williams responds, we would find only instances of shame in the face of defeat or failure to overcome others. Yet, in the Homeric works characters are shown to be blamed for actions and situations that exemplify breaches both in competitive and in collaborative virtues. Williams invites us to remember that the terms *aidos* (shame) and *nemesis* (indignation) form a system: I will be ashamed of actions that would make me indignant or angry if they were performed by others. If we pay close attention to the kinds of actions and situations that are the object of shame and indignation in the *Iliad* and the *Odyssey*, we will realize that selfish concerns, such as personal success and victory in battle, are not the only objects of such emotions. One can feel indignant at someone running away in battle, but indignation can be an adequate response also to such actions as giving poison for arrows, sending one’s mother away, or behaving like Penelope’s suitors in the *Odyssey* (*SN*, 80). Shameful actions include failures to behave in a generous way, to respect one’s parents, to have a sense of what I can do to others and others can do to me in such a way that we both preserve our self-respect.

Adkins’ distinction between competitive and collaborative virtues does not help us to isolate those actions that in the *Iliad* and the *Odyssey* would be stigmatized as shameful. Failures in generosity are as blameworthy and shameful as failures in courage. The opposition between shame and guilt from the point of view of selfishness versus altruism is therefore due to a prejudice in favor of guilt rather than to the actual analysis of Greek texts.

We should note here that if *nemesis* and *aidos* form a system of reciprocal expectations and responses, then the charges raised in points 3) and 4) fall:

The reaction in Homer to someone who has done something that shame should have prevented is *nemesis*, a reaction that can be understood, according to the context, as ranging from shock, contempt, and malice to righteous rage and indignation. It should not be thought that *nemesis* and its related words are ambiguous. It is defined as a reaction, and what it psychologically consists of properly depends on what particular violation of *aidos* it is a reaction to. As Redfield has put it, *aides* and *nemesis* are “a reflexive pair” (*SN*, 80).
In the interpretation that Williams draws from Redfield, nemesis is clearly a reactive response: the indignation felt at failures of aidos calls the other to answer for his actions and omissions rather than treat him as an object to be judged from a third-personal point of view. And if aidos and nemesis form a “reflective pair”, then aidos itself is a reactive attitude. The shame felt by Telemachus, when he realizes that he forgot to close the door to the storeroom from which the suitors are now taking out armors and spears, entails, among other things, a recognition of responsibility and a desire to make amends (SN, 50-52). Telemachus’ reaction is not a desire to hide from view: rather, it is a response that recognizes the authority of others to blame him for what he (albeit unintentionally) did. Williams analyzes the concept of responsibility in the Homeric poems and shows that it entails regret and the need for reparation. The opposition between objective and reactive attitudes, with shame belonging to the former and guilt to the latter, does therefore not apply to aidos, which seems rather to entail traits that the opposition in question attributes to the sphere of guilt (SN, 90-91). The stark contrast between shame and guilt from which the progressivist view took its bearing in distinguishing us from the Greeks ought to be reconsidered.13

3. The Other in Shame

Let us now turn to discuss how Williams addresses the charge that shame, as opposed to guilt, is heteronomous, i.e., that the person feeling shame is dependent on the opinions and values of others. In this context Williams asks us to concentrate on three points:

a) We should distinguish between a concrete other and an internalized other.

b) We should consider the attitude of the other. Do we need a critical audience in order to feel ashamed?

c) We should ask ourselves if the other who elicits our shame can be anybody or needs some further characterization.

Williams argues that it is a trivial mistake to think that shame is only triggered in the presence of others who witness our actions and find them wanting.

13 In his review of Shame and Necessity (Mind, New Series, Vol. 104, No. 413, Jan., 1995, 214-219) Colin Allen asks whom Williams is referring to when he speaks of “we” and “the Greeks.” Williams anticipates the objection to the usage of the first person plural by explaining that it is an invitational “we”: “More than one friend, reading this book in an earlier version, has asked who this ubiquitous ‘we’ represents. It refers to people in a certain cultural situation, but who is in that situation? Obviously it cannot mean everybody in the world, or everybody in the West. I hope it does not mean only people who already think as I do. The best I can say is that ‘we’ operates not through a previously fixed designation, but through invitation. (The same is true, I believe, of ‘we’ in much philosophy, and particularly in ethics.) It is not a matter of ‘I’ telling ‘you’ what I and others think, but of my asking you to consider to what extent you and I think some things and perhaps need to think others.” (SN, note 7, 171).
If we understand the Homeric culture as indeed a “culture”, and not just as a heap of facts from the past or a collection of actions and speeches lacking any internal structure, we have to acknowledge that the characters we encounter in the *Iliad* and *Odyssey* exhibit the culture of shame in their distinctive manners: they typically do certain things, and would never do other things. Now, the idea that shame is fear of being seen by others while doing blameworthy actions fails to account for the fact that someone like Achilles would find certain actions simply below himself, regardless of whether someone or no one saw him. Can we conceive of Achilles stealing by night the gifts that he had arrogantly refused to accept from the embassy in the clear light of day? Would such an action satisfy his sense of honor? Petty stealing, doing in secret what he could not do openly, is simply not part of Achilles’ character.

Achilles is not Gyges, who, having found a ring of invisibility, commits all sorts of crimes. Before finding the ring, Gyges looked like a good person. So tells us Glaucon in Plato’s *Republic*, who uses the example of Gyges to prove that people follow the law not because of justice, but in fear of punishment. Williams takes issue with Glaucon’s thesis precisely because the example assumes that norms cannot be internalized. 14

If shame were operative only in the presence of concrete witnesses Achilles, one of the most important representatives of a culture of shame, would not be inhibited from stealing when nobody could see him. If such a picture is unthinkable, it is because, as Williams points out, understanding Achilles means understanding his manners, and this implies that his sense of shame is not reducible to paying lip service to the expectations of those who happen to be present in the scene of his life at the time (SN, 81).

If shame entails reference to another, the other must be capable of playing a role independently of his concrete presence: the other of shame is an *internalized other*. Shame plays an inhibitory role even when one is alone and certain that nobody will find out what he is tempted to do. If this is the case, then the opposition with guilt is even in this case less stark than one was initially led to believe. Shame is far from a shallow emotion: like guilt, its roots are in our interiority. The concrete other can trigger my shame, but is not what shame can be reduced to.

---

14 SN, 98-101. The problem with Williams’s objection is that he attributes Glaucon’s thesis to Plato. In my view, Socrates responds to the example of Gyges when, in book IV of the *Republic* he introduces *thumos*, the part of the soul responsible for the internalization of norms. *Thumos* makes us respond with anger and indignation to the injustices committed by others, and with shame when we are to blame. I have discussed Williams’s objection to Glaucon’s example, as well as what I take to be Socrates’ response, in A. Fussi, “La critica di Williams alla Repubblica di Platone,” *Méthexis*, Volume 22, 2009, 39-59; see also A. Fussi, “Inconsistencies in Glaucon’s Account of Justice,” *Polis, The Journal of the Society for Greek Political Thought* (UK), vol. 24.1, Spring 2007, 43-69.
Let us now focus on the other’s attitude. Do we need a critical audience in order to feel shame?

In shame one feels exposed. In *Being and Nothingness* Sartre discusses two examples: in the first example, a person making a vulgar gesture suddenly realizes that he is being seen. Now the gesture, which had previously clung unreflectively to the subject, abruptly becomes a matter of judgment, a vulgar gesture, something that makes him shudder in shame. In the second example (which Williams takes up in his discussion) someone is induced by jealousy to peep though a keyhole. He is entirely absorbed in the action when he is abruptly brought to self-reflection by the sound of some steps in the hall. In both cases shame comes as the painful realization that the person is doing something vulgar at the same time as he realizes that he is being observed. 15 Shame is seeing ourselves through the eyes of another whose gaze is critical of us.

From Sartre onwards, scholars assumed that most cases of shame would involve an audience taking a negative attitude towards a subject, who, in turn, shares the audience’s critical view. Yet, as Gabriele Taylor has pointed out in her pivotal study on the emotions of self-assessment, one can feel ashamed of being admired by the wrong audience or for the wrong reasons. 16 If I feel contempt for someone, his admiration can trigger my shame. If the other has reasons to feel that I should be flattered by his admiration, I may, in turn, have reasons to feel ashamed of it.

Williams, like Taylor, refers to Max Scheler’s example of a model who, having been sitting naked for a painter, at some point realizes that the painter is no longer absorbed in his work but looks at her body with desire. All of a sudden the model is no longer protected by an impersonal relation with the painter: she feels exposed and ashamed of her nudity.

Of course shame might not be her only reaction: she could feel anger at the painter’s unprofessional attitude, or fear if she thought that he might assault her. Two paintings come to mind, one by Rembrandt and one by Artemisia Gentileschi, in which we see a woman realizing all of a sudden that her nudity is exposed to the gaze of two lecherous men. 17 The scene portrayed is the famous biblical episode of *Susanna and the Elders*. Two corrupt judges blackmail Susanna, the virtuous wife of a wealthy man, having introduced themselves into her garden while she was bathing, and demanding sexual favors.

---

15 See J.-P. Sartre, *Being and Nothingness*, 221 for the first example and 259 for the second example.
17 I owe to Peter Hacker, who let me read part of his book manuscript on the emotions (*The Passions – a Study of Human Nature*), the reference to Rembrandt’s *Susanna and the Elders* as an “an archetypal representation of the primal feeling of shame.”
Rembrandt van Rijn, *Susanna and the Elders*, 1647. Oil-on-panel painting, 76.6 by 92.8 cm
Gemäldegalerie, Berlin, Germany.


In Artemisia’s painting, the two well-dressed men are very close to Susanna’s naked body, so that the stress is not as much on their gaze as on their threatening proximity. They are whispering to each other, while Susanna averts her face in revulsion.

Rembrandt portrays Susanna while she is looking away in front of her. We can imagine she is meeting the gaze of the painting’s viewers, who are thus involved in the voyeuristic scene. One man is grabbing the cloth which barely covers part of her body; more than excited by her nakedness he looks reflective and malicious, as if he were pondering how he could profit from the situation. The other man, more at a distance, is leering at her with a sort of amused lust.
Artemisia portrays a woman who feels in danger — the scene is one of terror — while Rembrandt focuses our attention on Susanna’s shame at being exposed to the two men’s intrusive and lascivious gaze. In both paintings the men show no sign of uneasiness about what they are doing. They are shameless.

Even though the two painters represent differently Susanna’s emotional reaction to the men’s voyeurism, she certainly does not face a critical audience. Susanna (or, rather, her body) is represented without any doubt as an object of desire in the men’s eyes. If what we recognize in Rembrandt’s painting is shame, then we have to wonder if a critical gaze is relevant at all in generating this emotional reaction.

A similar point can be made with respect to the model in Scheler’s example. The painter is certainly not critical of her. Actually, she might even feel that from his perspective she should consider herself flattered by his sexual attention. The shame she feels cannot derive, then, from her identification with the attitude of the audience. Taylor suggests that what is involved here is a more complex reflective structure:

The model need not see herself as the artist sees her. But as the result of realizing her relation to him she sees herself in a new light. The point can be expressed by introducing a second, higher order point of view from which she is seen not as an object of sexual interest, but is seen as being seen as such an object. With this point of view she does identify, and this point of view is a critical one. The adverse judgment, however, comes not from the artist, but comes from herself. It is critical in that it pronounces it wrong for her to be so seen, at least at this time and by this audience.18

Williams, in turn, does not think that we necessarily need to imagine two kinds of judgments — the positive judgment belonging to the concrete other (the painter’s desire for the model), and the negative judgment, belonging to the model (that she should not be seen as a sexual object by such a man). What is relevant in the scene is that the model feels suddenly exposed. Previously her nakedness did not make her feel vulnerable: as Williams puts it, “she had previously been clothed in her role as a model; that has been taken from her, and she is left truly exposed, to a desiring eye.” (SN, 222). In other words, the feeling of shame is a reaction to the consciousness of her loss of power.

The idea that shame may be connected to a loss of power is interesting, though in this particular case, as we have seen in the two different pictorial interpretations of Susanna and the Elders, the sense of being exposed could cause fear rather than shame. What is it that makes Susanna’s awareness of her loss of

power turn into shame rather than fear?

In the case of shame the loss of power is not linked, as in the case of fear, to a concern with the consequences for one’s safety, but, rather, to a concern with one’s worth. The model feels ashamed at the thought that the painter might think it appropriate to look at her like that. If this is the case, though, one can see why Taylor thought it necessary to introduce a second, higher order point of view. Williams, on the other hand, introduces the higher order point of view as a second step in the process by which a certain figure is internalized. In his view, what initially appears as a loss of power in the eyes of a concrete observer, can become a loss of power or a failing in the eyes of an internalized other:

The root of shame lies in exposure in a more general sense, in being at a disadvantage: in what I shall call, in a very general phrase, a loss of power. The sense of shame is a reaction of the subject to the consciousness of this loss: in Gabriele Taylor’s phrase, quoted in the text, it is “the emotion of self-protection.” [...] More generally, the loss of power is not actually constituted by the presence of a watcher, even though it is still a loss of power “in the eyes of another.” A process of internalization is now possible, and “bootstrapping” can proceed in terms of an increasing ethical content given to the occasions of shame (SN, 220-221).

In other words: what in the experience of nakedness is a loss of power caused by a concrete observer, becomes, with internalization, a loss of power or failing in the eyes of an internalized figure, and being actually seen by someone while doing something that one would consider a failing is no longer necessary. Suppose I am a writer. Realizing that my novel appeals to people whose tastes I despise makes me feel like a cheap writer. Someone I despise, however, can be right about my failings. In this case my shame is compounded: I agree with the negative assessment, and I find it even more painful because it comes from a person whose views I normally do not take into account.

Williams takes up Taylor’s point concerning the attitude of the audience, and applies it thus to his theory of shame:

Shame need not be just a matter of being seen, but of being seen by an observer with a certain view. Indeed, the view taken by the observer need not itself be critical: people can be ashamed of being admired by the wrong audience in the wrong way. Equally, they need not be ashamed of being poorly viewed, if the view is that of an observer for whom they feel contempt. Hector was indeed afraid that someone inferior to him would be able to criticize him, but that was because he thought the criticism would be true, and the fact that such a person could make it would only make things worse. The mere fact that such a person had something hostile to say would not in itself
necessarily concern him. Similarly on the Greek side of the war, the opinions of Nestor carried weight, and those of Thersites did not (SN, 82).

Shame is not just in the face of someone, but about something, and, as we have now seen at some length, some actions do inspire shame even if those whose gaze is upon us applaud what we are doing or believe that we should be pleased by what they are doing.

This implies that we should distinguish the perspective of the agent from the perspective of the audience by focusing on the content of the judgment that plays a role in shame. The concrete other whose gaze triggers our shame may or may not judge things the way we do. If we despise someone’s judgment we will not be ashamed in front of him, unless we realize that what we are doing is indeed shameful, and we come to realize it not because our witness is reliable, but because having a witness makes us take a distance from what we are doing.

In sum: there is a difference between the internalized other and the concrete other. The internalized other is someone whose judgment we share; the concrete other is someone who may be critical, admiring, indifferent, and whose judgment we need not share.\(^{19}\)

If being admired by someone we despise makes us feel ashamed, then the internalized other, the other whose judgment we share, does not coincide with the concrete other who happens to approve what we are doing. It is not true, then, that shame is fear of losing face before any kind of audience, or that it is a superficial emotion which relies only on external appearance.

To some extent it is true that shame exposes our being dependent on the opinions of others, but we should also add that the others in question must be

---

\(^{19}\) See SN, 82: “Even if shame and its motivations always involve in some way or other an idea of the gaze of another, it is important that for many of its operations the imagined gaze of an imagined other will do. It is not so, of course, with the most elementary case, the shame of exposure when naked; someone who was afraid in that case of being exposed to a merely imaginary observer would be afraid of his own nakedness, and his fear would be pathological. But the imaginary observer can enter very early in the progression towards more generalized social shame. Sartre describes a man who is looking through a keyhole and suddenly realizes that he is being watched. He might think that it was shameful to do it, not just to be seen doing it, and in that case, an imagined watcher could be enough to trigger the reactions of shame.” Sartre is claiming that what forces the man to move from being completely immersed in the activity of looking through the keyhole to becoming conscious of himself is the presence of another. Williams may be moving too quickly here from exposure to the gaze of another to the imagined gaze of the imagined other. For a critical view of the role of the internalized other in the explanation of shame, see J. A. Deonna, R. Rodogno, and F. Teroni, In Defense of Shame: The Faces of an Emotion, Oxford University Press, Oxford 2012.
people whose opinions we respect. This was a central point in Aristotle’s treatment of *aischyne.*\(^{20}\) Williams finds it expressed in Homer:

Nausikaa is afraid of what people will say if they see her with the handsome stranger, and that there will be a scandal; but she adds: ‘And I myself would think badly of a girl who acted so’.

[...] An agent will be motivated by prospective shame in the face of people who would be angered by conduct that, in turn, they would avoid for those same reasons. Nausikaa is conscious of how she shares with others the reactions that they might have to her. [...] There has to be something for these interrelated attitudes to be about. It is not merely a structure by which I know that you will be annoyed with me because you know that I will be annoyed with you. These reciprocal attitudes have a content: some kinds of behavior are admired, others accepted, others despised, and it is those attitudes that are internalized, not simply the prospect of hostile reactions (*SN*, 83-84).

Let us now turn to the third point highlighted by Williams. By distinguishing the views and the attitude of the concrete other from those of the internalized other we come to realize that the problem of heteronomy is more complex than the usual charges raised against shame may lead us to believe. However, even if we grant that the other of shame is not just anybody who happens to be our witness, we could still object that the values on the basis of which we respond with shame are heteronomous in the sense that the internalized other is a representative of the society, of the neighbors, of others with whose moral criteria we uncritically identify.

Here Williams warns us about the risks of a Manichean attitude. It would be tempting to defend shame from the charge of heteronomy by assuming that the internalized other is free of any influences derived from the contingent factors of our social, historical and political life. In this perspective the other, in so far as he or she plays a role in shame, would simply be someone whose judgment we share. The important point is not *who* the other is, but *what he or she thinks.* If he thinks what we think, we preserve our autonomy of judgment: the charge of heteronomy vanishes, or is in any way considerably weakened. The internalized other, if we follow this line of thinking, is just an echo of our conscience.

If we assume this purified sense of the internalized other, however, we face a problem: in what sense can we still talk of an “other,” if all that remains after

\(^{20}\) "Since shame is imagination (*phantasia*) about a loss of reputation and for its own sake, not for its results, and since no one cares about reputation [in the abstract] but on account of those who hold an opinion of him, necessarily a person feels shame toward those whose opinion he takes account of” (*Rhet.*, ii 6.1384a24-27).
Williams’s Defense of Shame as a Moral Emotion

having purified the other of its accretions is the content itself (that this or that action is vulgar), and not the gaze that makes us aware of our vulgar gesture, and the perspective from which it appears vulgar?

We cannot have it both ways. We cannot hope to save shame from the charge of heteronomy by grounding its content in our own judgment, while at the same time holding on to the phenomenological insight that in shame we feel exposed, i.e., that we feel shame in the face of another. If we think that the real or imaginary other does indeed play a significant role in the emotion we cannot allow it to turn into a ghost empty of all determinations.

Williams’s solution is not to defend shame from the charge of heteronomy at all costs, but to make us aware that an attack on shame in the name of autonomy may be misguided:

It is a mistake to take that reductive step and to suppose that there are only two options: that the other in ethical thought must be an identifiable individual or a representative of the neighbors, on the one hand, or else be nothing at all except an echo chamber for my solitary moral voice. Those alternatives leave out much of the substance of actual ethical life. The internalized other is indeed abstracted and generalized and idealized, but he is potentially somebody rather than nobody, and somebody other than me. He can provide the focus of real social expectations, of how I shall live if I act in one way rather than another, of how my actions and reactions will alter my relations to the world about me (SN, 84).

Once again it is by reference to Greek literature that Williams invites us to see why it is important to hold on to the idea that the internalized other is “potentially somebody rather than nobody.” Someone who, like Ajax, is led by shame to consider suicide a necessity, may appear at first sight irrational. Williams’s subtle analysis makes us listen to the other whose disappointment Ajax cannot face. He cannot continue to live in a world in which those he respects can no longer respect him.

One can try to fill the concept of respect with purely abstract content, or claim that the voice of autonomy is pure and our dignity unshakable. Williams leads us down the opposite path: by paying heed to shame we become sensitive to the reciprocal expectations that bind us with each other, and willing to question the idea that concepts like dignity or autonomy are really helpful in establishing clear-cut historical boundaries.
Shame and the Internalized Other

Alba Montes Sánchez
University of Copenhagen
Center for Subjectivity Research
xjf783@hum.ku.dk
cfs.ku.dk

Abstract
In Shame and Necessity, Bernard Williams engages in a forceful vindication of the ethical significance of shame. In his view, shame is an extremely productive moral emotion because of the distinctive connection that it establishes between self, others and world, through a self-evaluation that is mediated by an internalized other. In this paper, I examine Williams’ conception of the internalized other and contrast it with other ways of conceiving the role of others in shame. I argue that, although Williams’ views contain many important insights, much is to be gained by conceiving the role of others in Sartrean terms instead. The other’s perspective is not merely internalized; it is constitutive of the kind of selfhood that has a capacity for shame.

Keywords
Bernard Williams, shame, internalized other, moral emotion

Introduction

In Shame and Necessity, Bernard Williams engages in a forceful vindication of the ethical significance of shame. In his view, shame is an extremely productive moral emotion because of the distinctive connection that it establishes between self, others and world, through a self-evaluation that is mediated by an internalized other. In this paper, I examine Williams’ conception of the internalized other and contrast it with other ways of conceiving the role of others in shame. I argue that, although Williams’ views contain many important insights, much is to be gained by conceiving the role of others in Sartrean terms instead. It allows us to better understand the experience of shame and its moral significance. In Sartre’s account of shame, the other’s perspective is not merely internalized; it is constitutive of the kind of selfhood that has a capacity for shame. According to this view, the role of others in shame is at the same time thinner and more fundamental than the one advocated by Williams. It is thinner, because it does not presuppose the learning or endorsement of any substantial set of values and norms in order to feel shame. But it is also more fundamental, because it makes the relation with others, or rather, the capacity for it (relationality), constitutive of the self that can feel shame. This means that the structure of relationality is prior to any internalization of norms, and it establishes the ground for shame to become an emotion informed by norms and standards.
This paper starts by sketching a fairly typical definition of shame according to the relevant literature, a description that is in line with Williams’ own account of this emotion. In order to clarify why he postulates the notion of an internalized other to explain shame, I take a step back and analyze the role of others in shame by contrasting it to other emotions in the same family, which shows that audiences are not necessary for shame. The question then arises: what is the role of others in shame? The following section returns to Williams’ views on this matter, and specifically to his hypothesis of the internalized other. In the final section, I discuss my own Sartrean view and argue in its favor.

**Shame**

Among emotion researchers from various disciplines, the word ‘shame’ is often used in at least two senses. Some use ‘shame’ to refer to an “affect,” i.e. a hard-wired, innate response,¹ that underlies a whole family of hedonically negative self-conscious emotions, including embarrassment, humiliation and some types of guilt. Many others use it to refer to a specific emotion belonging to this family. I will be using it in this latter sense, and comparing it to other related emotions in the following section, while retaining the idea that all these emotions form an inter-related family.

Shame as an individual emotion is characterized by a feeling of exposure, inferiority and vulnerability. Typical bodily manifestations include blushing, averting the gaze, adopting a collapsed bodily position and so on: in shame, one feels smaller or wishes to become smaller and hide from view. In the relevant literature, shame has been described through several labels: a self-conscious emotion, an emotion of self-assessment, a social emotion or a moral emotion. It has been labeled as a “self-conscious emotion” because it is directed back at myself: the intentional object of the emotion is the individual that feels it, not the situation or action which gives rise to the shame episode. In shame, I focus on myself and see myself as small, faulty or inadequate. As such, it has also been called an “emotion of self-assessment” because it involves a negative self-evaluation.² This negative self-assessment can be due both to active and passive aspects of selfhood: to actions and omissions of all kinds (lying to a friend, acting or speaking against one’s values in order to maintain status in a particular social group), to things that befall us (victims of abuse typically feel it), to character traits, physical features, social background and so on.

The label “social emotion” refers to another key aspect of shame: exposure. Many authors claim that shame is a response to being exposed to the censoring gaze of a real, an imagined or an internalized audience.³ The unpleasantness of

---

¹ Cf. e.g. Elison, ‘Shame and Guilt.’
² Taylor, *Pride, Shame, and Guilt*.
³ Cf. e.g. Williams, *Shame and Necessity*; Maibom, ‘The Descent of Shame.’
exposure of a negative trait would explain why in shame we often experience a wish to hide and disappear from the view of others. However, the connection of shame to exposure to an actual external gaze or judgment is a rather controversial point, and this is what this paper focuses on. Finally, the label “moral emotion” mostly refers to its role in mechanisms of self-censorship and regulation of behavior according to norms or standards. This paper deals mainly with the social aspects of shame, but the other dimensions will come to the fore, especially the moral one, which was of special interest to Williams.

Shame and audiences

According to Williams, shame is an unequivocally social emotion. Now, to say that shame is social amounts to much more than saying that we learn the codes and standards of what is shameful from other people, that those standards are encoded in culture, or that shame serves social functions. As social psychologists Hareli and Parkinson⁴ and philosophers Deonna, Rodogno and Teroni⁵ explain, these claims are obviously true, but they do not tell us anything especially interesting about shame in particular, or even social emotions in general. They are far too broad to distinguish shame and the like from other, non-social, emotions, because all human emotions are partially governed and shaped by cultural codes and most of them serve social functions.⁶ Characterizing shame as social implies attributing to others a specific crucial role that is not exhaustively covered by the above general assertions. So what role do others play in shame?

In order to pin this role down, it is helpful to divide the possible answers into a taxonomy devised by Deonna, Rodogno and Teroni.⁷ They identify three different strands of the claim that shame is social: (i) shame as heteronomous, i.e., informed by values that do not belong to the ashamed subject, but to other people; (ii) shame as involving “an evaluation in terms of one’s appearance vis-à-vis an audience”; (iii) shame as the result of adopting an external perspective on ourselves.⁸ In short, others can provide the values, the gaze or the perspective. Williams, for his part, falls within the second strand: for him (and many others) the negative self-assessment of shame is triggered by the disapproving evaluation of an audience. But would he not fall within the first strand too? After all, what exactly causes my shame? The gaze of the other or her evaluation of me? Williams maintains that both are important, by arguing that this audience is an internalized

⁴ Hareli and Parkinson, ‘What’s Social About Social Emotions?’
⁵ Deonna, Rodogno, and Teroni, In Defense of Shame.
⁸ They disagree with all three strands and offer a thought-provoking non-social account of shame, which I cannot engage with here.
other with very specific characteristics: it is, as we will see, a respected other. This dampens the heteronomy of shame to a very large extent. But how does Williams come to this conclusion? Let me spell out some of the considerations at play.

The first consideration is that, if shame requires an audience, it cannot possibly be an actual audience that is present in all instances of shame. After all, it is not difficult to think about examples of shame felt in solitude, when we think about our failures (even those that nobody else knows about), or when we remember certain past situations. This can be seen more clearly by comparing shame with embarrassment.

One might think that embarrassment is merely a mild form of shame, and indeed the two emotions are related, but there is an important experiential difference between them. Shame clearly concerns our sense of self; it burdens us with an “unwanted identity”9 and impacts our self-esteem and self-respect.10 As Williams brilliantly put it, shame helps us “understand how a certain action or thought stands to ourselves, to what we are and to what realistically we can want ourselves to be.”11

Embarrassment does not seem to have this impact, and empirical studies have confirmed this.12 Nussbaum points out the differential features of embarrassment, as opposed to shame: although both typically take the subject by surprise, embarrassment is “momentary, temporary and inconsequential,” while shame lasts longer and is more serious.13 This is the case, according to Nussbaum, because embarrassment does not involve, like shame, a sense of being flawed and defective, but merely a sense that something is socially out of place (marked social attention, often in the form of praise, can cause embarrassment). As such, it is social and contextual; it always requires an actual audience in front of which we are performing awkwardly. Solitary shame is possible, since one can feel faulty or inadequate when thinking about one’s flaws or remembering one’s failures, but solitary embarrassment (a solitary feeling of social awkwardness) makes no sense. The audience may turn out not to be there - perhaps you simply mistakenly thought that someone had seen or heard you - but it must be part of the story. And as soon as we are on our own, or we have ascertained that there was nobody looking or listening, embarrassment disappears without leaving a feeling of degradation. We typically feel embarrassed in front of others of things that do not embarrass us when we are alone, such as bodily functions, or of failures that are conceivable as such only because others are present, such as telling a joke that nobody else finds funny. This need not be the case in shame, which often concerns

9 Ferguson, Eyre, and Ashbaker, ‘Unwanted Identities.’
10 Cf. e.g. Rawls, A Theory of Justice, para. 67; Taylor, Pride, Shame, and Guilt; Deigh, ‘Shame and Self-Esteem.’
11 Williams, Shame and Necessity, 93.
13 Nussbaum, Hiding from Humanity, 204–5.
flaws that are perceived as more permanent and much less dependent on the social context.

A good example of solitary shame can be found in the passages of *Anna Karenina* that describe her return to St. Petersburg from Moscow, where she has just met the dashing Vronsky.\(^\text{14}\) She sits alone in the train, with no audience, trying to fight her intense andanguishing shame by telling herself that her behavior to Vronsky was proper, that she did not betray her husband and nobody can blame her for anything. It is clear, however, that she does not dare to admit, even to herself, that a respectable and decent married woman like herself could have fallen in love with a dashing young officer like Vronsky. But who is the audience that is assessing her here in light of her unacknowledged feelings? And how to explain that she felt no shame at the party, while flirting and dancing with Vronsky in front of the high society of Moscow, and she only comes to feel it while she sits alone in the train?

This contrast between the two moments might lead one to think that audiences are irrelevant to eliciting shame: all Anna Karenina had to do was adopt a different perspective on herself and the situation. This might seem even clearer in an example proposed by Goldie in a different context:\(^\text{15}\) a man gets drunk at an office party, he climbs on top of a table and starts singing “Love is like a butterfly” at the top of his voice. At the time, in his drunkenness, he may have only been enjoying the music and the general merriment, but when he remembers the episode the next morning, he will see the situation in a different light, he will realize that his colleagues were laughing *at* him and not *with* him, and only then feel ashamed of himself. This would seem to indicate that the gazes of others, even when coupled with their mockery, are insufficient to cause shame: something is needed on the part of the subject as well, a change of perspective on himself. But does this completely rule out the need for an audience? Both Anna Karenina and the man in Goldie’s example exposed themselves publicly before an audience, and felt ashamed of themselves in solitude when they re-examined the situation, so the audience *was* a part of their memories. These examples might therefore suggest that both the audience and the change of perspective are necessary.

Furthermore, one might want to argue that the essential element is the negative evaluation of others: the man at the office party only becomes ashamed in the morning, because only then does he realize that his colleagues found him ridiculous. Similarly, Anna Karenina was engrossed with Vronsky and caught up in the excitement of falling in love with him at the party; while only in the train did she have enough distance from the excitement to realize what others must

have been thinking about them. This might suggest, then, that the actual presence of others is unnecessary, but their evaluations are not.

Even so, it would be a mistake to argue that the awareness that others evaluate me negatively constitutes shame. This would amount to conflating the emotion of shame with what I would call “objective” shame, i.e., the verdict of society on what is shameful, or disgrace. Disgrace is not an emotion, but a condition. In the Online Oxford English Dictionary, it is defined as follows: “loss of reputation or respect as the result of a dishonourable action … [In singular] a person or thing regarded as shameful and unacceptable.” A similar definition can be found in the Merriam Webster’s Dictionary: “1a: the condition of one fallen from grace or honor; b: loss of grace, favor, or honor. 2: a source of shame.” Disgrace is, therefore, not an affective phenomenon, but a “social attribute,” i.e., an objective state, or a thing that can cause such a state. But the objective state of disgrace does not always necessarily cause shame: social attributes, codes and verdicts do not shape our experience of shame in a necessary and inescapable way. Shamelessness is possible and comes in many forms, including immorality and moral reformism. One may feel ashamed of things that are not disgraceful, and conversely, one may be in a state of disgrace in one’s society and not feel ashamed. One may argue that this is at some point the case of all moral reformers who actively criticize with the ways of life and the codes of shame and honor in their societies: think about Diogenes the Cynic, Jesus Christ and the sexual revolution, for example. The possibility to resist the external verdict and respond to disgrace with defiance and even pride indicates that the external evaluation is also insufficient to cause shame.

What is then required for shame to arise? Sartre writes that “my shame is a confession”: in the moment of shame, I am endorsing the evaluation of unworthiness or inadequacy. Through shame I confess my “sin”, I confess that I am thus and so. Deonna, Rodogno and Teroni argue that this is the central element. They view shame as a negative self-evaluation in terms of my own self-relevant values, where no reference to the other is necessary. According to them, if I assess that I’m not capable of exemplifying my own self-relevant values, I will feel shame regardless of what others think or see in me. Their work yields a crucial insight: that my own values and perspective also play an important role in

18 Yovel, ‘The Birth of the Picaro from the Death of Shame’, 1299.
19 Hutchinson, ‘Facing Atrocity.’
20 Cf. ibid. for an interesting discussion of cases of shamelessness and their meaning; also Calhoun, ‘An Apology for Moral Shame.’
21 Sartre, Being and Nothingness, 261.
22 Deonna, Rodogno, and Teroni, In Defense of Shame.
eliciting shame, but in my view they go too far. Evidence from empirical psychology seems to suggest that the role of audiences is crucial for intensifying and eliciting shame experiences. Furthermore, looking at developmental psychology, it seems clear that all experiences of self-conscious emotion first appear in infants in the presence of audiences, which indicates that shame felt before others is much more primary than solitary shame. Solitary shame depends on self-reflectivity and a well-developed self-concept, but not all shame is like this. Therefore, an account of shame that only accords a peripheral role to others seems to be missing something important.

The internalized other

Let me now turn to Williams’ account of the role of others in shame. He is very well aware of the complexities of the issue that I addressed above, and in order to do justice to them, he (along with many other authors) defends the idea that shame is caused by the internalization of an audience, which doesn’t necessarily involve explicitly imagining or remembering the audience every time. Williams agrees to a large extent with Wollheim’s Freudian account, which explains shame as caused by the introjection of an external authority figure, which becomes an internal “criticizing agency,” or superego, that judges and censors the ego. Freud believed that the superego starts to emerge around the fifth year of life, as a result of the child’s internalization of the parents’ moral standards through education. For Freud, small children cannot feel shame, but during the developmental stage that he calls “sexual latency,” a transition stage that goes approximately from 3 to 7 years of age (but may be longer or shorter depending on various circumstances), several “dams” are built that restrict and block the flow of sexual drive. These obstacles are shame, disgust and morality. They arise as a form of self-protection, against a feared parental figure. According to Freud’s theory of the Oedipus complex, the little boy phantasizes his father as threatening him with castration due to the boy’s sexual desire for his mother. Then, in Wollheim’s words, “in terror, the boy introjects the father, thereby exchanging a frightening external danger for enduring internal torment. The superego now harangues, upbraids, chastises the boy according to standards that make no allowances for,

23 Smith et al., ‘The Role of Public Exposure in Moral and Nonmoral Shame and Guilt.’
24 Cf. Rochat, Others in Mind; Reddy, How Infants Know Minds.
26 Cf. e.g. Williams, Shame and Necessity; Maibom, ‘The Descent of Shame.’
27 Wollheim, On the Emotions.
28 ‘Superego | Psychology.’
29 Freud, Three Essays on the Theory of Sexuality, 43, 58; see also Metcalf, ‘The Truth of Shame-Consciousness in Freud and Phenomenology.’
indeed often expressly run counter to, the boy’s own wishes.”30 Due to this developmental history, this introjected figure possesses both authority and heteronomous force:

[I]t is because internal figures originate, through an incorporative phantasy, from external figures, that, once they have been internalized, they may well continue to address the person who now harbors them as an alien force. They may set themselves to make the person feel shame or guilt on occasions when the person finds no reason to do so.31

Now, Williams does not make any claims about the developmental history of the internalized other, but he agrees that it possesses an authority that is related to its otherness, just like the superego does. But he makes a small observation here. On Wollheim’s account, shame would share the same psychoanalytic origin as guilt: they would both be the result of different kinds of indictments by the superego. According to Williams, however, the figures that are internalized in these emotions are different, at least in terms of their perspectives and roles: in shame the internalized other would play the role of an observer or a witness, and in guilt, the role of a victim or a judge.32 The shaming audience, in any case, would be an element of our psyche, something we acquire and internalize as children, and that accompanies us throughout our lives, monitoring our emotions and behavior. Shame in these kinds of accounts is essentially in all cases a consciousness of exposure to the censoring gaze of another.

Now, this other cannot just be a literal copy of one’s actual father and his values and norms, since people can come to be ashamed of their fathers (or their mothers, or their educators) for holding and cherishing values that they later repudiate. But on the other hand, it is also problematic to assume that the internalized other can simply be a placeholder for any observer, since not every observer, witness or judge can make us feel ashamed. Williams argues, for example, that we typically are not ashamed to be evaluated negatively by people we despise. In his view, therefore, the internalized audience is someone we respect.

To illustrate this, Williams employs the tragedy of Ajax by Sophocles. He quotes Ajax’ suicide speech, where the Greek hero wonders how he can face his father after covering himself in shame.33 This is not a purely rhetorical device: it points to something deep. Indeed, in the tragedy, Ajax is surrounded by people who express support and love for him, and do not condemn him: his servants and his wife. But he is a warrior, it is his honor as a warrior that is at stake, and those

31 Ibid., 178–79.
33 Ibid., 85; Sophocles, *Sophocles II*, Ajax, 462 seq.
views have no value for him in that context, no relevant impact on his sense of who he is. They have no power to counter the fact that he would feel deep shame in front of his (absent) father Telamon, who as a young man had been a brave and celebrated warrior. For Williams, however, in contrast to Wollheim, the crucial point is not that Telamon happens to be Ajax’s father. What is essential is that Telamon functions, according to Williams, as the anchoring point that symbolizes the world where Ajax wants to live and the identity he wants to preserve (in this sense, his brother Teucer or his admired Achilles would have done just as well). Ajax is aware that the things that we do and do not do impact on who we are; that the world has certain expectations that must be fulfilled in order for us to possess and retain certain identities; that our identities, in short, are not under our exclusive control. In Ajax’ case, his final monologue expresses that if he cannot command the respect of his father and men of similar position, worth and accomplishments, he cannot keep his identity as an honorable warrior, or his self-respect. In Williams’ account, the internalized other always points towards the world I (want to) live in and its expectations of me:

The internalised other is indeed abstracted and generalised and idealised, but he is potentially somebody rather than nobody, and somebody other than me. He can provide the focus of real social expectations, of how I shall live if I act in one way rather than another, of how my actions and reactions will alter my relations to the world about me.34

As such, what is crucial about the internalized other is that it commands our respect and functions as the reference point and limit to our possibilities in the world.35 The fact that, according to Williams, respect is required implies that there is a degree of autonomy in the choice of the audience that can shame me, since respect seems to entail the recognition and admiration of certain values in the other. This means that my own values and standards are relevant to my shame and play a role in eliciting it. But at the same time, the other is genuinely other because it points towards my possibilities in the world, and those are not determined exclusively by me. Williams believes that giving respected others the power to shame us, even in the cases in which we do not share their opinions, is a sign of good moral discernment, because it entails a recognition of the limits of our own reason and the need for the help of others. Since, according to Williams, reason has its limits and moral truth is indeterminate, giving weight to the opinions of others is not incompatible with critical thinking and discernment: it is

34 Williams, Shame and Necessity, 84.
35 Ibid., 82–89.
often a result and an enhancement of them.\textsuperscript{36} Shame is an acknowledgement that we give others that weight.

Now, Calhoun agrees with Williams in this last point, but she criticizes him for not going far enough.\textsuperscript{37} According to her, his claim that, through respect, we have some power to choose the audience that can shame us is not so different from saying that in shame I am my own judge, that ultimately the only evaluation at play is my own evaluation of myself. This is so, in her view, because respect relies on shared values, and if the person I happen to respect betrays them, I would typically withdraw my respect.\textsuperscript{38} In her view, therefore, Williams’ solution means that eventually it all comes down to our individual values and norms. It wouldn’t be so far away from an account of shame as autonomous, like the one proposed by Deonna, Rodogno and Teroni, who claim that the often referred-to phenomenology of the gaze of another is a metaphor through which we sometimes seek to make sense of shame.\textsuperscript{39} It is part of some, but not all, instances of the emotion, and therefore cannot be a necessary factor for shame to arise. If this is so - their reasoning goes - would it not be more natural, and more faithful to the phenomenology of shame, to say that I am always the main judge, that the standards at play in this self-assessment are my standards?

Calhoun objects that accounts of this kind, which present shame as autonomous, have very undesirable implications when it comes to the shame that members of oppressed groups feel before the shaming of their oppressors.\textsuperscript{40} In her view, these accounts cannot explain such instances of shame without implying that the oppressed are complicit in their own oppression at some level, or that they are morally immature or self-alienated, since they let themselves be influenced by external opinions they don’t share or deem respectable. The dubious assumption that, in her view, this kind of accounts make is that “no rational, mature person who firmly rejects her subordinate social status would feel shame in the face of sexist, racist, homophobic or classist expressions of contempt,”\textsuperscript{41} and therefore, those people who do feel it are morally immature. Calhoun thinks this is unacceptable, because it shifts from the aggressor to the victim a substantial part of the responsibility for the suffering caused. Williams’ account, in her view, has the same flaw, for it implies that in this sort of cases the ashamed subject respects someone who, in her own view, is not worthy of respect, and thus the suspicion of collusion stands. This cannot be right, she claims. Her strategy to avoid this problem, while honoring Williams’ insight about the importance of others in our moral lives, consists in separating the realm of moral autonomy, reason and

\textsuperscript{36} Ibid., 100, my emphasis.
\textsuperscript{37} Calhoun, ‘An Apology for Moral Shame.’
\textsuperscript{38} Ibid., 134–35.
\textsuperscript{39} Deonna, Rodogno, and Teroni, \textit{In Defense of Shame}, pt. one.
\textsuperscript{40} Calhoun, ‘An Apology for Moral Shame.’
\textsuperscript{41} Ibid., 136.
knowledge from the realm of morality as a practice shared with others. Thus, I might recognize that the opinions of others have social weight and have an impact on me, because we’re all participants in a shared practice of morality, while believing that these opinions are false, that they carry no epistemic weight.42

Calhoun’s strategy to save the autonomy of oppressed minorities is therefore to keep it separate from shame and give shame a different ethical role. When one is shamed for belonging to an oppressed group, one might deny the truth of the insulting remark, one might deny that the minority one belongs to actually has that negative trait, or deny that being a part of that group is a shameful thing, and still recognize the negative impact that such evaluations by others have in one’s public identities, the power they have in shaping the world one will have to live in. That impact is real, and acknowledging it amounts to acknowledging a fact about the world and about our public identities, but it does not thereby imply that our capacity for autonomous judgment is compromised.

Now, there are at least two problems with Calhoun’s criticism, one of them having to do with emotions of self-assessment, the other with respect. Beginning with the latter, one might argue that respecting another person does not have to entail agreeing with everything she says, thinks or values. This is clear in debates about tolerance. Indeed, the classic liberal way of thinking about tolerance, which is an essential attitude to cultivate in a pluralistic democracy, precisely requires distinguishing between respecting persons and agreeing with their actions, opinions or judgments. Tolerance is supposed to be justified precisely because one respects the person, since she is a free autonomous agent, and as such worthy of respect, even though one disagrees with her opinions - disagrees to the point where those opinions are deeply unpleasant, perhaps even in some sense painful to oneself, and yet one maintains one’s respect for the person who holds them. Indeed, one tolerates those opinions out of respect.43 With this I do not mean to take a position in the debate on tolerance, I simply mean to highlight that there are important and widespread views on respect that do not imply agreeing or even sharing values with the respected person. They merely imply conceiving persons as intrinsically valuable in themselves.

Judging by Calhoun’s choice of examples, and from the ease with which she concedes that one might withdraw respect, it seems to me that she construes respect not in the above way, but as something quite close to admiration. Admittedly, Williams also choses “role models” as examples of respected persons, and so he seems to lean in that direction as well. But if tolerance is thought of as a sign or a consequence of respect, I think both Calhoun and Williams go too far in approaching respect to admiration: respect is a more neutral attitude than they lead us to understand. At this point, one might think that such a notion of respect

42 Ibid., 139.
43 Cf. Tonder, Tolerance.
is too wide to allow us to distinguish audiences that have the power to shame us from those that lack it, but I don’t think so. Indeed, the problem with both Williams’ and Calhoun’s formulations is that they are too restrictive about who can shame us. In a sense, they get things the wrong way around. They seem to imply that nobody has the power to shame us unless we give it to them. In my view, the reverse is true: everybody has, to a higher or lesser degree, the power to shame us unless we withdraw it from them through contempt or disengagement, for example. The power to shame is not a privilege we accord to certain esteemed others, it is a default power we all have over each other to varying degrees in social relations, and completely withdrawing it from particular individuals or groups is typically an effortful endeavor (with the exception perhaps of some pathological cases, like those of psychopathy or other social impairments).

The second of my criticisms to Calhoun has to do with her way of presenting the relation between shame and oppression. The main problem is that she only takes into account one of the varieties of shame and other shame-related emotions that come into play in resisting oppression and caving in to it. But as Deonna, Rodogno and Teroni remark, at least two other notions that Calhoun overlooks need to be considered in these cases: humiliation and stigma. Distinguishing shame from the feeling of humiliation can do part of the work of saving the autonomy of oppressed minorities. The feeling of humiliation is different from shame in several crucial aspects. First and foremost, it necessarily involves another agent, who is trying to downgrade your status vis-à-vis hers. Her actions can be extremely violent (torture typically involves systematic acts of humiliation), but they need not be: something like refusing to greet you or acknowledge your presence can be humiliating. As such, the other’s negative evaluation and downgrading of you is typically perceived as unjust and outrageous, and the focus of the experience is as much on the humiliated self as on the offending other. Taking this into account, one can argue that acknowledging the practical impact of someone else’s negative evaluation but refusing to ascribe any normative weight to it is precisely what we do when we feel humiliated, when we feel unjustifiably attacked, accused, offended or put down in the eyes of others. Humiliation is a common response to shaming; and some of Calhoun’s examples could be described as humiliation. But the moment one transitions into shame, one seems to be appropriating the negative evaluation on some level, this is why shame is considered an emotion of self-assessment.

44 Cf. Hutchinson, ‘Facing Atrocity’ on Diogenes the Cynic.
47 Cf. Calhoun, ‘An Apology for Moral Shame’, 137. The example she takes from Adrian Piper, in which Piper herself seeks to describe her feelings as ‘groundless shame’, is in my view a good example of feelings of humiliation.
Indeed, in this type of situations, when we feel ashamed of something we do not deem shameful because of pressure from others, it is not rare later to come to feel ashamed of one’s own shame, as FitzGerald argues. As an example, she uses a personal anecdote, where she reports that out of shame anxiety, she lied to a group of colleagues about shopping in a discount supermarket, and told them that she shopped in a more expensive one. Afterwards, however, when telling the anecdote to her partner, she felt ashamed of having allowed her shame anxiety to overrule her other values. In other words, she reports being ashamed of her shame. This meta-shame can often arise because one interprets the initial shame episode as a betrayal of one’s own deeply held values, as a moment of weakness in which one momentarily upheld someone else’s wrong values and self-evaluated in terms of them. One feels ashamed of having felt unjustified shame, of having caved in to external pressure. One feels that one’s own value system has been fleetingly contaminated. This is exactly what FitzGerald reports. This meta-emotion evidences the difference between shame and the feeling of humiliation, and it is proof that one felt a contaminating shame in the first place. Imagine a situation where, instead of acting like she did, she had told the truth, and someone in the group had ridiculed her for her choice of supermarket. She might have felt humiliated while enduring the mockery, but she probably would not feel ashamed of herself afterwards. This meta-shame suggests that her initial shame, as opposed to what would have been the case in humiliation, did evidence a fleeting value contamination.

Since social groups can and do exert high levels of pressure, feelings of humiliation often transition into shame, when the other’s negative evaluation infiltrates our own. This infiltration can be very fleeting, like in FitzGerald’s example above, or more permanent and insidious, as in the case of stigmatized groups that are constantly bombarded by stigmatizing messages. It is therefore important to also look at stigmatization processes and be aware of their capacity to infiltrate our values and contaminate our autonomy.

It should be noted, however, that the transition need not always be from humiliation to shame. As Morgan argues, shame can prompt us to examine our

---

49 Ibid., sec. 2.
51 I agree with Calhoun that one should resist accounts that construe members of stigmatized groups as morally immature for feeling this type of shame. But the solution is not to separate shame from autonomy and construe autonomy as impervious to emotions. A more promising route, in my view, is to construe moral maturity differently, as involving something else than autonomy and allowing corrections to it when necessary. Cf. Williams, Shame and Necessity, 100.
relations to others, and thus motivate a transition from shame to humiliation. In the case of a stigmatized group, this transition towards feelings of humiliation might be an affirmation of autonomy, but this is not necessarily the case for all individuals and groups. In some other cases, responding to shaming with shame (rather than humiliation) might be an expression of autonomy too, such as when somebody publicly accused of a reprehensible action she indeed performed accepts her responsibility and shame.

In any case, I believe that the right approach to these issues requires abandoning the idea that, at least when it comes to emotions, autonomy and heteronomy are dichotomous, that this is an all-or-nothing issue, that shame must be either autonomous or heteronomous. FitzGerald argues convincingly that it is not possible to clearly determine whether the self-assessment of shame is autonomous or heteronomous in all cases. In her view, autonomy and heteronomy come in degrees, and shame is a phenomenon where this becomes particularly clear, since it shows that the values of others can infiltrate one’s own to varying degrees.

Thus, so far it seems clear that the role of others in shame is not equivalent to heteronomy, because shame can be autonomous, or it can fall somewhere in between autonomy and heteronomy. But it seems also quite clear that this is an emotion through which others can and do exert some influence on our values. This is, arguably, one of the reasons why shame has often been regarded as especially conducive to moral learning, as in Aristotle, and it is frequently used for education purposes, as a tool to instill social and moral norms. In this sense, shame has even been called the “midwife” and the “condition of possibility” of any human education. This is also why humiliation and stigmatization can become insidious. The question now is whether we need to postulate an internalized other to make sense of the influence others can have in our shame.

The other in shame: internalized or constitutive?

Let me now address head-on the issue of the internalized audience. Why postulate it? As shown above, Deonna, Rodogno and Teroni are right in pointing out that in many cases, one cannot pin down the audience in front of which one is allegedly

---

52 Morgan, On Shame.
54 Burnyeat, ‘Aristotle on Learning to Be Good.’
55 See e.g., Heller, ‘Five Approaches to the Phenomenon of Shame’, 1024; Ferlosio, El alma y la vergüenza, 29.
56 Ferlosio, El alma y la vergüenza, 29. This is not to say that shame should be used as an educational tool. I merely mean to register the fact that it can and has been used in this way.
Furthermore, the issue of internalization raises the questions of how and when do we internalize this audience, and which audience do we internalize. According to the psychoanalytic story, the audience would be our parents and main educators, which might be in line with many of the above examples. But what about someone who came to severely question and even be ashamed of her parents’ values? What if, for example, the racist one is your father and you become ashamed of him having such views? This gets complicated even further by Williams’ requirement that one must respect the audience that gives rise to our shame. Does this mean that we internalize all the others we respect? Or a representative of all spheres of our lives that significantly shape our identities? How many internalized others do we typically have? When do we internalize them? Perhaps looking at development can give us some clues to clarify these questions.

According to Freud, shame emerges during the “latency stage” of development, between the 3rd and the 7th year of life. Modern developmental psychologists have pushed this age further back, although there is no consensus on a specific age, or on approaches and interpretations of results. According to the cognitive-developmental approach to developmental psychology, self-conscious emotions such as shame or pride emerge in normally developing infants around the second year of life. This is so because in the cognitive-developmental view, self-conscious emotions are thought to depend on the possession of a concept of self. Empirical proof that this concept is in place is linked to the mirror self-recognition test. Typically infants start to pass this test consistently from the 18th month of age onwards, and from then on, supporters of the cognitive-developmental view start to talk about the onset of self-conscious emotions, such as embarrassment, pride, jealousy and shame. There is much to discuss here that falls out of the scope of this paper, but for my current purposes, it is enough to underline that, according to the cognitive-developmental view, the condition for self-conscious emotions to arise is a concept of self that allows the child to re-identify herself from the perspective of any external observer (the image that the mirror shows is what others can see of us).

There is a growing number of developmental psychologists that criticize these views, and favor an interactive, second-personal approach instead. With this approach, Reddy, for instance, argues that starting from a self-concept gets things

---

60 Cf. Lewis, *Shame* for a prominent example; Rochat, *Others in Mind*, 96–98 endorses this view.
backwards.\textsuperscript{62} In her view, what self-conscious emotions require is self-experience and interpersonal awareness (which the infant is capable of from the start), not any conceptual abilities.\textsuperscript{63} In Reddy’s view, interpersonal engagement elicits in the infant a very basic sense of self and other (the other person as qualitatively distinct from an object), that is tied to the interaction. This basic interpersonal awareness is what prepares the ground for a concept of self, not the other way around.\textsuperscript{64} This is also what allows infants to experience self-conscious emotional reactions of coyness and showing off, which she documents already during their first year of life, sometimes as early as in the second month.\textsuperscript{65} She doesn’t call these reactions full-blown shame or pride (although some others, like Trevarthen, have no qualms about attributing shame to 3-month-olds\textsuperscript{66}), but they are its precursors and form its experiential ground. The point is that the sense of self that is required for shame, the kind of self-consciousness it exemplifies, arises immediately in and from the interaction and emotional engagement with others: it does not require a stable self-concept independent of the interaction, nor an internalized other. Indeed, these things rely to some extent on the sense of self that arises out of intersubjective emotional engagement.\textsuperscript{67} This is why Reddy suggests that we call it “self-other-consciousness”, instead of just “self-consciousness”.\textsuperscript{68}

All this suggests that the most basic structures that enable shame are social in a fundamental way, but they do not rely in any form of internalization. This is in line with the Sartrean insight that shame exemplifies a form of self-consciousness that is fundamentally different from the minimal pre-reflective form given in the first-personal character of experience. As Zahavi explains, there is a minimal form of selfhood implied by the fact that experiences are perspectival, they imply an experiencer for whom these phenomena are given.\textsuperscript{69} Sartre’s example is illuminating: if I’m crouching at a keyhole to spy on someone, at that point I’m completely focused on what I see and hear, and my experience of self is minimal, it is implicit in those perceptions as their \textit{mode}: those experiences are given \textit{for me}, but this for-me-ness is not their focus, it is not part of their content, it is rather the mode in which they are given. Now, suppose I hear a noise in the corridor, indicating perhaps that another person is approaching, and I’m overcome by shame. Now my experience has changed fundamentally: I focus on myself and


\textsuperscript{63} Reddy, \textit{How Infants Know Minds}.

\textsuperscript{64} Ibid., 144.

\textsuperscript{65} Ibid., 129–40.

\textsuperscript{66} At talks given at the University of Copenhagen and the University of Portsmouth in 2014.

\textsuperscript{67} Reddy, \textit{How Infants Know Minds}, 144.

\textsuperscript{68} Ibid., 148–49.

\textsuperscript{69} Zahavi, \textit{Subjectivity and Selfhood, Investigating the First-Person Perspective}. 195
experience myself as an object, the object of someone else’s perception. This is what Sartre calls my “being-for-Others”: a fundamentally different form of self-consciousness that arises from recognizing that there are other subjects in the world who can perceive me.

Sartre’s subsequent development of these ideas is very complex and problematic in many ways, but this is no place to engage in such long controversies. All I want to highlight is his central idea that engagement with others gives rise to a fundamentally different form of self-consciousness, that Zahavi and I have previously called social self-consciousness.70 This is the form of self-consciousness that is at stake in shame. And according to Reddy, it is present in a very basic form from the first weeks of life, when infants are aware and respond to others’ attention to them. This does not necessarily entail that shame (or pride, for that matter) can only be elicited when other people are present, it simply means that in those phenomena, the dimension of selfhood I am focusing on is the same that arises out of and is at stake in my engagements with others. A solipsistic being would be incapable of shame or pride. In emotions of social self-consciousness there is a crucial change in my self-experience: I relate to myself from the perspective of engagement with others. This does not mean that I need to take on any specific person’s attitudes towards or judgments of me, it doesn’t mean that I need to imagine a specific audience. It simply means that I focus on the dimension of myself that can be perceived from the outside and engaged with. Very young infants might only be able to experience it in direct interaction, but adults, who do possess a self-concept, can experience it in solitude. This is both thinner and more robust that internalization of an audience: it is a background constitutive condition, a feature of any self that can experience shame. But it isn’t an other with a particular face or an idiosyncratic set of values. It is a condition of possibility for “self-other-conscious” emotions to arise.

Now, this is obviously not all that shame requires. Shame is an unpleasant form of social self-consciousness, a form that foregrounds the vulnerability of self and the dangerousness of others. Sartre believes that it is more fundamental than the pleasant forms, like pride, but it is quite unclear why that should be so, and developmental psychology does not support that view. Be this as it may, my aim here is not to establish the primacy of shame over pride, or the other way around, it is rather to highlight that the social self-consciousness they evidence does not require internalization of an audience: it requires the ability to relate to others as subjects, which is there from the outset in a very basic form. All emotions of social self-consciousness get enriched and complexified through sustained engagement with others, to the point where the notion of internalizing an audience might seem plausible. The underlying structure, however, is not a product of internalization; it is a feature of engagement. What we internalize, or learn, are norms and

70 Zahavi, Self and Other; Montes Sánchez, ‘Self-Consciousness, Caring, Relationality.’
standards, and facts about the social world and one’s position in it. This makes it possible for shame (and guilt, for that matter) to disengage from the direct disapproval of others and become a personal self-evaluation, but this does not require internalizing the other. The other remains outside, as that which constitutes the dimension of self that gets evaluated: the intersubjective self.

To conclude, I want to return to Williams and his claim that shame includes a reference to real concrete others, to the world I want to live in and my possibilities in it. I think he is essentially right, but this does not require an internalized other. It requires an intersubjective self, i.e., a self that can become aware of a dimension of her being that depends on others. In other words, it requires the capacity to understand others and relate to them as subjects who can perceive oneself, and to understand oneself as an object of their experience, an object who is thereby affected and changed. It requires the capacity to experience that my identity, who I am, is not fully in my control, but depends on others. As one’s world widens and enriches through learning, and one acquires the capacity to project oneself backwards and forwards in time, among other things, this enables the kind of complex, thick experiences that Williams describes, where Ajax understands he can no longer live as the hero he had been and wanted to keep being. The experience is concrete and can involve a reference to individual others, but it does not require internalization of audiences, just the awareness that who I am does not solely depend on me. The heightened shaming power of specific others depends on how my world is built, how much I care about them, and how much influence they have on the possibilities that matter to me, among other factors. All in all, I believe that shame is social in a thin but fundamental sense: in the sense that it is only possible for a social being, who is aware that her identity is partially dependent on others. But this does not require internalizing others: it requires living with and being affected by them.

Bibliography


Courage and Shame: Aristotle’s Nicomachean Ethics III.6-9

David Roochnik
Department of Philosophy
Boston University
roochnik@bu.edu

ABSTRACT

This paper analyzes the intricate relationship between courage and shame as presented in Aristotle’s Nicomachean Ethics III.6-9. To cite the most pressing example: citizen-soldiers endure in the face of deadly risk in the hope of gaining honor and avoiding what is “shameful” (1116a29). They act “on account of virtue” and “a desire for what is noble” (1116a27-29). Nevertheless, Aristotle insists that such citizen-soldiers, however admirable, are not truly courageous men. In order to understand both the distinction between, as well as the proximity of, shame and courage, this paper draws on Bernard Williams’s account of shame offered in his Shame and Necessity.

KEYWORDS

Aristotle, Nicomachean Ethics, shame, courage, Bernard Williams.

Describing the courageous man in NE III.7, Aristotle says that he will “fear such things as he should, as reason (logos) maintains and for the sake of the kalon” (1115b12-14).1 By contrast, in NE II.4 he says that in order for agents to qualify as genuinely virtuous they must “act by choosing and choosing [the virtuous actions] for their own sake (di’ auta)” (1105a33). Terence Irwin thinks that “Aristotle takes these two descriptions of the virtuous person’s motive to be inseparable.”2 Although I believe his statement to be true, his reconstruction of Aristotle’s argument on its behalf is unsatisfying. For Irwin, the kalon, at least as used in the NE, is “necessarily connected to the good of others,”3 and a virtuous action is kalon. Hence, acting for the sake of the kalon is inseparable from acting for the sake of the virtuous action itself. Part VI will argue against this view on the grounds that it excessively “moralizes” Aristotle’s conception of virtue.

1 Kalon will be left untranslated. It is typically rendered as “beautiful,” “fine” or “noble” and is regularly used to described that “for the sake of which” a virtuous person acts. So, for example, The “liberal” person (eleutherios) “will give [money] for the sake of the kalon” (1120a24). Translations from the NE are my own.
3 Ibid., 251.
Courage and Shame: Aristotle’s Nicomachean Ethics III.6-9

In order to narrow down the vast (and treacherous) terrain of the NE, this paper will focus only on courage as explicated in NE III.6-9. Parts I-IV will analyze these chapters. Part V will investigate another pair of closely related descriptions. In NE III.7 Aristotle says that the courageous man acts for the sake of the *kalon*. In NE III.8 he states that the citizen-soldier who, while similar to is not a genuinely courageous man, acts from a desire for the *kalon*. Once again, I will argue that Irwin’s moralizing interpretation of the *kalon* will not suffice to explain the proximity of these phrases.

1. NE III.6

Aristotle announces that in taking up each of the “moral” (*êthikê*: 1103a15) virtues, “we should state what they are, concerning (peri) what sort of things they are, and how (pôs) they are” (1115a4-5). He begins with courage (*andreia*: 1115a6), defined as the “mean concerning (peri) fear and confidence” (1115a7). Strikingly, however, he quickly shifts focus away from courage and toward the person who possesses it: “the courageous man” (*ho andreios*: 1115a11). For reasons unstated, understanding the “who” seems to precede understanding the “what.”

NE III.6 then asks, with what is the courageous man concerned? This move reveals an important feature of Aristotle’s treatment of the moral virtues (with, arguably, the exception of justice). Understanding what a virtue is first requires concretization or personalization - the virtuous person must be studied - which in turn requires identifying the area of concern, the field or region of human life, in which that person activates his virtue. This procedure can be likened to cartography. Throughout his discussion of the moral virtues, Aristotle demarcates various regions of concern on a conceptual map of the ethical life. In these bounded areas the virtues, manifested in the people possessing them, will appear. In this sense, his procedure can also be likened to a phenomenology of the ethical life.

The region of concern in which courage appears is constituted by the two emotions or passions (*pathê*) of “fear and confidence” (*phobous kai tharrê*: 1115a7). As he explains in III.9, the former is decisive, and so Aristotle begins here by discussing fear and its objects. He seems to approve of the definition of fear as “the expectation of a bad thing” (*kakou*: 1115a9). This is, at least, compatible with his more detailed discussion in *Rhetoric* Book II, where he says that “fear is a

---

4 For reasons that soon should be obvious, the masculine will be used throughout this paper. It is meant to reflect Aristotle’s thinking, not my own.

5 Compare NE VI.5: “concerning practical wisdom, we could grasp it by studying (*theôrêsantes*) those whom we say have practical wisdom” (1140a24-25).

6 The mapping of the moral virtues in NE III.6-IV.11 thus develops or fleshes out the “outline,” “sketch” or “diagram” (*diagraphê*: 1107a34) presented in NE II.7.
pain or disturbance that comes from the image of an imminent bad thing that causes destruction or pain” (1382b14) and is accompanied by “an expectation of experiencing a suffering that will be destructive” (1382b22).

Human beings fear many things, including “disrepute, poverty, illness, friendlessness, and death” (1115a10-11), but not all of them belong in the region in which the appears. Aristotle must thus narrow the sprawl. Although, for example, a reasonable person should fear disrepute - in fact, it is, “kalon to do so, and aischron not to do so” (1115a12-13) - doing so does not manifest courage.

To make the region in which courage is manifested more precise (and thus visible), Aristotle restricts the proper objects of the courageous man’s fear to the “greatest” (megista: 1115a25) of the fearful things; that is, to the worst of the bad things. This is death. Again, however, since most human beings fear death, and many comport themselves admirably in the face of it, this determination is inadequate. Facing death “on the sea or from disease” (1115a29), for example, does not qualify as a proper occasion for courage. Aristotle does not elaborate, but he does give a clue. He asks whether courage is only manifested in the face of “the most kalon” (kallistoi: 1115a30) of circumstances, and then states that “such circumstances would be found in war, for here there is the greatest and most kalon of dangers” (1115a30-31). It follows, then, that “the one who is fearless (adeês) in the face of a kalon death could authoritatively said to be courageous” (1115a33-34). At this stage of the analysis, Leighton thus seems quite right in saying that “the paradigm of courage…courage in its most perfect and noble form…is to be found upon the battlefield.”

Several difficulties in Aristotle’s ethical mapping of courage emerge from these brief remarks. Most basically, many people who are not courageous will look like they are. As just mentioned, since it is kalon to fear disrepute, it is tempting to include the person doing so (properly) among the courageous. This, however, would be a mistake. Another example, to expand (considerably) the sparse language Aristotle himself employs (in 1115b1-6), can be generated by imagining two people on a ship during a storm. One is an experienced sailor. When the waves get bigger, he remains unperturbed since he knows they do not actually threaten the safety of the ship. Another, however, is a passenger who has never sailed before. Seeing the waves grow he “despairs of his survival and the prospect of such a death upsets him” (1115b2-3). Nonetheless, he does not capitulate to his fear and so, like the sailor, appears “fearless” (adeês: 1115b1). To a third person who knows neither of the two men and can only observe their behavior, they appear indistinguishable even though they are, in fact, quite different.

---

NE II.4 prefigured this difficulty. There Aristotle asserted that it is impossible to determine whether people are virtuous simply by witnessing their actions. Even if an action “holds in a certain way (pòs echêi)” - that is, “is in accord with the virtues” (kata tas aretas) - it must be performed by an agent who “holds in a certain way (pòs echôn) as well” (1105a28-31). In short, the agent must be virtuous if the action is to be counted as genuinely virtuous. Specifically, he must meet three conditions: (1) he acts “knowingly (eidôs);” (2) he “acts by choosing and choosing [the actions that accord with the virtues] on account of themselves (di’ auta);” (3) he “holds (echôn) in a stable and unwavering manner” (1105a31-34); that is, he reliably and consistently possesses a virtuous character trait. Obviously, a deep knowledge of the agent, rather than a mere observation of his behavior, is required in order to determine whether he meets these criteria.

A second and related difficulty raised by Aristotle’s exploration of the courageous man in NE III.6 is that it seems to identify courage with fearlessness. Consider this passage, parts of which have already been mentioned.

One should fear some things, like disrepute, and it is kalon to do so and aischron not to do so. For the person fearing [disrepute] is decent and modest (aidêmôn), while the one not doing so is shameless (anaischuntos). He is said by some to be courageous metaphorically, since he does bear some similarity to the courageous man. For the courageous man is someone fearless (aphobos) (1115a12-16).

Aristotle’s thinking here is expressed in typically sparse language. Perhaps it is this: disrepute is bad, and as such is a reasonable object to fear in a kalon manner. So, for example, a decent person may assist people who are trapped in a fire because of his fear of what others would say about him if he failed to act, and it will be kalon for him to feel such fear. Entering the burning house, then, he appears fearless. He is not, however, genuinely courageous. After all, he is motivated by fear of disrepute. At best, he is similar to the courageous man.

The big problem here is located in the last sentence of the passage: the courageous man is said to be someone fearless, a point Aristotle reiterates later when he says that “the one who is fearless (adeês) in the face of a kalon death could authoritatively said to be courageous” (1115a33-34). But on Aristotelian terms this cannot be. After all, courage is the “mean concerning fear and confidence.” By definition, then, it requires some measure of fear.

On the one hand, this problem can be eliminated by noting that NE III.6 is a preliminary stage of Aristotle’s analysis, and that in it he is exploring widely held opinions that he may not share. Note, for instance, the optative at 1115a33: someone “could say” (legoit’) that the fearless man is the courageous man.

---

8 There does not appear to be a difference in the meanings of aphobos and adeês.
Presumably, though, he would be wrong. For the reasons mentioned above, however, this is an easy mistake to make. Recall the earlier example: to an external observer the experienced sailor and the passenger - both of whom look calm and composed and behave accordingly - seem indistinguishable, even though one understands that the waves are not dangerous, while the other believes the ship will be swamped and that he will suffer an unbecoming death. It is impossible to detect courage through a glance.

NE III.6 provokes another challenge. Aristotle says courage is “authoritatively” instantiated only in facing a “kalon death.” Furthermore, it must occur in the face of “whatever may bring sudden death, and this occurs especially in war” (1115a33-34). Recall Leighton’s comment that “courage in its most perfect and noble form...is to be found upon the battlefield.” But does it matter on which side the good soldier is fighting? Can one who is fighting in a war that is not kalon be courageous? What about the Persian soldier who in every observable way appeared indistinguishable from his Athenian counterpart when Xerxes invaded Greece? Did he not have the opportunity to be courageous because his cause was unjust? Could he not have met the three conditions Aristotle lists for genuine virtue in NE II.4?

Aristotle’s mapping of this virtue in NE III.6 is not yet perspicuous. Its various difficulties can be encapsulated under the heading of “the recognition problem.” How is it possible to distinguish the courageous man from those who merely appear courageous?

2. NE III.7

In NE III.7 Aristotle advances beyond the preliminary reflections of NE III.6 and seems to speak in his own voice. First and foremost, he maintains that courage does require the agent to feel fear. Indeed, there are some fears that would require a “super-human” (huper anthropon: 1115b8) capacity not to feel and courage, a human virtue, cannot possibly require that. It is thus in facing normal fears that “the courageous man is, as a human being, imperturbable” (anekplêktos: 1115b11). He will “fear such things as he should, as reason (logos) maintains, and for the sake of the kalon” (1115b12-14). To reiterate, the underlined phrase differs from that required of the virtuous agent in NE II.4, where he must choose virtuous actions for their own sake. Before confronting this apparent discrepancy, another issue, also related to the recognition problem, must be addressed.

Aristotle states that the courageous man “endures” (hupomenei: 1115b23). He “experiences” (paschei: 1115b20) fear, presumably in facing a kalon death, but does not get shaken by it. As such, he is not fearless. But if that is so, then how can he be distinguished from someone who has “self-restraint?” (engkrateia: 1145a18); someone who knows what he should do, feels a desire or passion to do
otherwise, but then resists that temptation? The ethical typology of NE VII.1 makes it clear that even though the self-restrained person (the *engkratês*) is praiseworthy, he is inferior to the genuinely virtuous person. (Analogously, the *akratês*, the person characterized by “lack of self-restraint” [*akrasia*], is blameworthy, but superior to the genuinely vicious person.) The principal difference between the virtuous and the self-restrained is that the former suffers no conflict between reason and desire or passion. As Aristotle says, for example, about the moderate person (*ho sôphrôn*), “the desiring part” (*to epithumêtikon*) “ought thus to be in harmony with (*sumphônein*) reason (*logos*)” (1119b15-16). This person wants to do what he should do, and in fact enjoys doing it. As a result, he experiences no internal struggle. By contrast, the *engkratês* fights a battle against himself, against his *pathê*, and wins. He does what he thinks he should even though his desires sorely tempt him.

If the courageous man is required to feel fear, and endure in the face of it, then how does he differ from the *engkratês*? Leighton is helpful in tackling this question. He argues that, yes, the courageous man does indeed feel fear, but he does so in a uniquely virtuous manner. Most important, he does not experience it, or any other emotion, as a force that needs to be “whipped into line.” In other words, instead of being enemies that must be subjugated, the *pathê* of the virtuous man are potentially useful friends or allies. Describing the fear experienced by the courageous man, Leighton explains:

What Aristotle claims for fear, then, is a mechanism sensitive to and focusing upon certain sorts of information, leaving persons moved to fear with very specific conceptions of how their surroundings relate to them. To use a contemporary way of speaking about this, fear identifies certain features of our world as salient ones.

Viewed in these ways, fear is not something to be struggled with or overcome, but something to be sensitive to and exploited. The prospect of imminent, destructive or painful evils that fear embodies is particularly important to a creature whose ultimate aim is to live and do well. Fear on a battlefield then involves an awareness of the situation with which one has to deal, and does so in a way that allows one to prepare to deal with the situation. Feeling fear enables one to act courageously.

On this understanding of fear, not only is there little wonder that Aristotle should think that the courageous fear, but also little surprise that this passion should help define the state as an excellence rather than a form of self-control

---

[engkrateia]. Fear, so conceived, is not like [the engkratēs’] bodily appetites with which one must struggle and overcome, but rather something that is of great assistance to the courageous, something to be sensitive to and exploited […]. So understood, fear is a sign of virtue, not [engkrateia].”

To illustrate by means of an innocuous example: when riding a bicycle on a busy city street, especially one that does not have dedicated bike lanes, it is useful to feel some measure of fear. It is vital for the cyclist to be alert to the possibility of a parked car veering suddenly to the left or a car on the left abruptly turning to the right. The appropriate measure of fear, just as Leighton describes, heightens the cyclist’s sensitivity to his surroundings and helps him identify those features of it that might get him killed. Strikingly, this sense of keen alertness, this bit of excitement, is attractive, which is why the cyclist returns to it day after day. His doing so illustrates (only by analogy, since cycling is not a virtue) an Aristotelian principle: “what accords with virtue is pleasant, or not painful, or minimally painful” (1120a27). The virtuous person enjoys, or at least is not pained by, being virtuous. Aristotle makes the same point in stronger terms in NE I.8: “the person who does not enjoy kala actions is not good (agathos). For no one would say that the person who does not enjoy acting justly is just….if this is so, then actions according to virtue are pleasant in themselves” (kath’ hautas) (1099a17-21).

Consider the moderate person: he is “not pained by the absence of pleasure and for holding himself away (apechesthai) from it” (1118b32-33). At first blush, this seems odd. If, for example, the beautiful body of his neighbor’s wife becomes available to him, and he is attracted to her but knows adultery is wrong, how could “holding himself away” from his passion not be a kind of pain? On Leighton’s model, it is because the passion has been integrated into the agent’s whole being. Yes, the woman is attractive and so, yes, he is drawn to her. But acting on this desire is inconceivable to him and so the passion he feels is, once again, not a wild beast that needs to be tamed, but a source of energy that can be successfully sublimated. While it may not be pleasurable for him to hold himself away from the woman’s body, it is not a struggle nor does it hurt to do so. This notion will prove to be pivotal in Part VI below.

To sum up so far: Leighton successfully differentiates, at least conceptually, the courageous man from the one who exhibits self-restraint. The latter must struggle hard to contain the fear that is exploding in his chest. The former feels fear, but instead of having to fight it, he welcomes it as an energy source. For an external observer, however, it remains difficult, if not impossible, to distinguish the two. Like the experienced sailor and the passenger on the ship, they may appear to be identical in the face of danger. Internally or psychologically, however, they are not. In fact, we can easily imagine (today) a device that

10 Ibid., 91.
measures the physiological fluctuations that accompany fear and pleasure in a human body. By using it we could detect differences in the courageous man and the engkratês. Even as the two appear identical on a casual glance, the latter would measure low on the pleasure scale, and high in fear, while the former would be the reverse.

(Aristotle would support this thought experiment, since he identifies courage as a virtue of the “non-rational parts” (1117b22). As such, either its presence or absence could be accompanied by physical symptoms. So for example, someone fearing death “turns pale” (1128b14).

Unfortunately, even if Leighton provides us with significant resources to differentiate the courageous man and the engkratês, the recognition problem is, as we shall next see, far from resolved.

3. NE III.8

“Courage is this sort of thing,” Aristotle says at the outset of III.8, “but others (heterai) are said in five ways” (1116a15-16). He does not specify to what “others” he refers and so, not surprisingly, translations vary. Bartlett and Collins attempt to assist the reader by supplementing the Greek: “but there are also other kinds of courage, spoken of in five ways.”¹¹ This is not impossible, but it is misleading. Because heterai is feminine plural it could refer to “courages” (andreiai), which in turn could be taken as shorthand for “kinds of courage.”¹² However, as will soon become clear, this chapter describes five kinds of people who only “seem” (dokousi: 1116a20) to be, or are “similar to” (hōmoiōtai: 1116a27), or “appear” (phainontai: 1116b8) courageous, but in fact are not. This chapter, then, is not articulating five kinds of courage. Crisp, who also supplements the heterai, captures this by translating thus: “Courage, then, is something like this. But the name is applied to five other states of character as well.”¹³

The order in which the five “others” will be examined below differs from that presented by Aristotle himself. It begins with the cases that are more easily distinguished from courage, and concludes with the most difficult.

1. “Those who are ignorant also appear to be courageous” (1117a22). Such people are oblivious to the dangerous reality that faces them and so, even if they

---


¹² Sarah Broadie and Christopher Rowe, in *Aristotle’s Nicomachean Ethics: Translation, Introduction, and Commentary* (Oxford: Oxford University Press, 2002), follow this reasoning strictly when they render the phrase, “but there are also other ‘courages,’ so called.”

¹³ Roger Crisp, *Aristotle’s Nicomachean Ethics*. (Cambridge: Cambridge University Press, 2000). Terence Irwin, in *Aristotle’s Nicomachean Ethics* (Indianapolis, Hackett,1985) translates the line as follows: “but other states, five of them, are also called bravery.”
DAVID ROOCHNIK

seem as unperturbed as a courageous man would be, they do not feel the fear that is requisite in order to possess the virtue. As Leighton puts it, such virtuous fear has cognitive value since it “identifies certain features of our world” - namely, those that immediately threaten us - “as salient,” and ignorant people fail to detect these.

2. “Nor are men of good hope (euelpides) courageous, since they are confident in the face of risks because they have often been victorious over many. They are much like (paromoioi) the courageous because both are confident” (1117a9-12). In fact, however, they are “not far” (1117a23) from the ignorant; that is, they are unaware of salient features of their surroundings; namely, the ones that are really dangerous. For this reason, they are like those who are “drunk” (1117a14).

3. The person of “experience” (empeiria: 1116b3), such as a professional soldier (or the sailor on the ship mentioned in NE III.6), will also appear courageous. Well-versed in things military, he knows what events are actually life-threatening, and what are merely “false alarms” (kena: 1116b7).14 Since the latter are common, most of the time he will seem unafraid or imperturbable, and so he is easily confused with the genuinely courageous soldier. In fact, Aristotle says, such men typically flee when their lives are seriously at risk. For this reason they are quite unlike “citizen soldiers” (politiaka: 1116b18), who keep their post even in the face of likely death (and who will be discussed in 5. below.)

4. Men energized by “spirit” (thumos: 1116b31) are ready, even eager, to face real danger. Driven by anger, or the desire for revenge, or brute love of the fight, they enter the battle and face death with fierce joy. As such, they too will appear courageous. But they are not because “courageous men act on account of (dia) the kalon” (1116b30-31), while spirited men, driven by their “passion” (pathos: 1117a9), act impulsively. Nonetheless, the character described here (1116b24-1117a9) is easy to mistake for the genuinely courageous because “the courageous are spirited” (thumoeideis: 1116b26). As Aristotle puts it, in them spirit “works together” (suergei: 1116b31) with other psychological components. This seems to mean that, similar to fear, this potentially lethal passion will be integrated into the totality of their being, specifically their “reason” (logos:1117a9). To apply Leighton’s terminology again, spirit will be “something to be sensitive to and exploited” rather than struggled against. In fact, Aristotle says that if thumos is supplemented by “choice and the for the sake of which” it will actually become “courage” (1117a4-5). Presumably the “for the sake of which” is the kalon.

To sum up again: it now seems that the courageous man must face a kalon death in battle, must energetically try to kill his enemies while feeling either pleasure or at least no pain in doing so, and yet not be driven by rage or an exuberant lust for blood, but instead be guided by a deliberate choice and for the

14 “False alarms” is from Bartlett and Collins, Aristotle’s Nicomachean Ethics: A New Translation.
sake of the kalon. As the final (albeit Aristotle’s first) example makes clear, such a person is terribly difficult to recognize.

5. “Political” (politikê) courage - or rather the political inflection of apparent courage - “seems most like” (1116a17) real courage. Citizen-soldiers “seem to endure” (1116a18) in the face of deadly risk on account of both the punishment they would suffer were they to desert their post and “the honors” (1116a19) they will receive should they prove heroic. The first motivation is not impressive, since such soldiers are “compelled” (1116a30) to fight. The second, however, is. As Aristotle puts it, those who face death in battle in the hope of gaining honor and avoiding what is aischron (1116a29) do so “on account of virtue (di’ aretên). For they do so on account of shame (aidô) and a desire (orexin) for the kalon, since they have a desire for honor, and in order to avoid blame, which is aischron” (1116a27-29).

These lines bring the recognition problem to a head. First, both the courageous man and the citizen-soldier act “on account of virtue.” Second, both their actions are propelled, in some fashion, by the kalon. The courageous man acts “for the sake of the kalon” (1115b13) or “on account of the kalon” (1117a8), while the citizen-soldier has a desire for honor, which translates into a “desire for the kalon.” Both will appear fearless and either enjoy or at least not be pained in facing death. How can they possibly be differentiated?

4. NE III.9

This chapter begins by asserting that courage is far more concerned with fear than with confidence. As such, “courage is in fact painful” (1117a33). This is odd, since the exercise of virtue should be pleasurable, or at least not painful, so Aristotle elaborates. “The end (telos) that relates to courage would seem to be pleasant” (1117a35). He illustrates with a comparison to a boxer who endures painful blows “for the sake of” (hou heneka: 1117b3) sweet victory. Similarly, a courageous man endures death and wounds “because (hoti) it is kalon to do so and aischron not to” (1117b9). This phrase is similar to the description in NE III.6 (1115a12-13) of the “modest” (aidêmôn) person. It is kalon for him to fear disrepute and would be aischron for him not to. The difference seems to be that, as stated in NE II.4, the courageous man “knowingly” (eidôs: 1117b13) “chooses (hairetai) what is kalon in war” (1117b14-15), while the person animated by aidôs, principally the citizen-soldier, desires honor, which is kalon, and seeks to avoid disrepute. The difference in these formulations is subtle, to be sure, and needs to be elaborated.
5. For the Sake of the Kalon v. Desire for the Kalon

In NE IV.9 Aristotle presents two reasons why the agent driven by *aidôs* is not genuinely virtuous. First, *aidôs* “seems to be more of a *pathos* than a characteristic” (*hexis*: 1128b9). As stipulated in NE II.4, for an action to be virtuous it must be performed by an agent who “holds (*echôn*) in a stable and unwavering manner” (1105a31). Because the person energized by shame is driven by a *pathos*, and so acts impulsively, he does not qualify. This is reinforced by the fact that *aidôs*, like fear, manifests itself in physiological symptoms such as “blushing” (1128b13). Second, while it is appropriate for the “young” (1128a16) to be moved by *aidôs*, it would not be suitable in an older, genuinely virtuous person.\(^{15}\) Such a person “will never voluntarily do base things” (1128b30) nor (unlike the *engkratês*) even be tempted by doing so. Therefore, *aidôs* is simply not part of his psychological repertoire. By contrast, the young person, impelled by *orexis* for the *kalon*, strives to be, but has not yet actually become, a good man. At his age, then, the best he can do is take his bearings from what others think of him. As such, shame is his appropriate motivator. When he fights as a citizen-soldier he holds his ground even when facing the threat of imminent death because he fears looking bad, and hopes to look good, in the eyes of his comrades and elders, through which he sees himself.\(^{16}\)

By contrast, a mature and genuinely virtuous person no longer needs such external motivation. As Aristotle says in NE I.6, because honor “seems to reside more with those who bestow it” (1095b25) than with the agent himself it does not qualify as the highest good, which of course turns out to be “an activity of the soul in accord with virtue” (1098a17). Furthermore, he continues, “people seem to pursue honor so that they may be convinced that they themselves are good” (1095b28). In other words, the desire for honor manifests an internal insecurity, which younger people are naturally prone to feel, and consequent need for validation. Because the highest good must be “complete” (*teleion*: 1097a29) and “self-sufficient” (1097b7), being motivated by either the desire for honor or the fear of disrepute eliminates someone from the ranks of the genuinely virtuous.

It must be remembered, however, that the mature man was once young. In his youth he too was guided by what others thought of him. Through experience and reflection he somehow outgrew this need. He became able to see himself for what he is rather than needing the eyes of others. Alternatively formulated, he internalized those others. This notion will be elaborated in Part VI below.

\(^{15}\) Note that shame plays an important role in the educational scheme Aristotle sketches in X.9. A young person is habituated so that he regularly “feels pleased by (*stergon*) the *kalon* and displeased by the *aischron*” (1179b30).

\(^{16}\) My colleague, the ancient historian Loren J. Samons, tells me that although Athenians were liable for service into their 50s, and the average age of the Athenian soldier is unknown, it is likely that most were rather young; that is, between twenty and thirty-five.
Aristotle now seems to have mustered the conceptual resources needed to distinguish between the apparently courageous citizen-soldier and the genuinely courageous man, and not simply on the basis of age differences. The former, driven by aidôs, acts from a desire for the kalon, understood as some sort of public visibility and affirmation, and the latter for the sake of the kalon understood...understood as what?

This is a problem since, as many commentators have noted, “Aristotle never explains in the NE what to kalon is.”17 Furthermore, as Irwin notes, throughout the corpus in different contexts it clearly refers to different properties. He cites as examples things that are “pleasant through sight and hearing,” those that exhibit “teleological order” or “symmetry, definiteness and greatness,” and those that are “appropriate and admirable.”18 But he also argues that there is “one recurrent point” when the word is used in the context of ethics: something kalon is “admirable insofar as it extends to the common good.”19 He reformulates: there is “good reason to believe that the property [Aristotle] picks out in moral contexts through his use of ‘kalon’ is moral rightness,”20 understood as acting for the good of others.

On Irwin’s view, then, the genuinely courageous soldier who endures in war would be driven by a non-selfish imperative, which could be formulated roughly as follows: ‘I must stay at my post and risk my life because doing so will benefit others and as such is morally right.’ Another clause can be added in order to distinguish it from the imperative that would guide the citizen-soldier: ‘Even though honor is kalon, whether I am recognized for my actions or not is irrelevant to me.’

There is textual support for Irwin’s thesis. In the Rhetoric, for example, Aristotle states “those actions are kala that a man does not for the sake of himself. Actions that are simply good (agatha) are ones that someone does for the sake of his city, while neglecting his own interest” (to hautou)” (1366b35-67a5). Nonetheless, Irwin’s account is unsatisfying for at least two reasons. Crisp marshals one. Simply stated, “the virtuous person’s only object of concern is his own happiness,” not the well being of others. This is true even if the agent does care about the well being of others, and acts (virtuously) on their behalf. As Crisp puts it, the Aristotelian agent “may be moved by a genuine non-instrumental concern for the good of a friend. But he knows that action on such concern is part of what constitutes the best life for him.”21 If Crisp is right - and whether he is or not is obviously a complicated question - then Irwin is wrong in his interpretation.

19 Ibid., 243.
20 Ibid., 241.
of what it means to act “for the sake of the kalon.” The phrase would not signal moral rightness understood as acting essentially for the benefit of others.

Irwin’s account is problematic for other reasons, which come to light in returning to the genuinely courageous man (or soldier) of NE III.6-9, the one who acts for the sake of the kalon. On Irwin’s view this person is willing to risk death on the battlefield for a moral reason; that is, the good of others. Since Aristotle explicitly contrasts him with the citizen-soldier who is prompted by desire for honor or fear of disrepute, he is unconcerned with what those others will think of him. He wants to benefit them but does not need to be noticed by them. He is willing to act even if his agency becomes invisible. This, however, violates a primary feature of “kalon.” As Lear puts it, “visibility or ‘showiness’ is essential to” Aristotle’s conception of the kalon, which is why she translates it as “fine.”22 If she is right, then by moralizing “for the sake of the kalon” Irwin strips it of its connection to visibility. In other words, he too sharply distinguishes “kalon” from “beautiful.”

To elaborate: Aristotle says that “political courage” as embodied in the citizen-soldier “seems most like” real courage. But if Irwin were right, then the two would in fact be fundamentally different. The genuinely courageous man, on his view, is guided by a moral imperative and is animated by concern for others, while the citizen soldier is animated by aidôs and therefore a kind of self-concern. He desires the kalon for himself, and so seeks honor and avoids disrepute. Nonetheless, Aristotle states that even the young soldier acts “on account of virtue.” Now, with this phrase he may be referring to an immature version of virtue, but his use of the word suggests that far from being fundamentally different as Irwin’s account would imply, there is continuity between the young self-regarding soldier and the mature virtuous man.

Furthermore, recall that the mature man was animated by, even educated by, aidôs when he was young. But if Irwin were correct that acting “for the sake of the kalon” means acting for what is morally right and the betterment of others, then a youth spent driven by aidôs would not prepare someone for a morally virtuous life. Instead, a youth spent in the army, surrounded by (mostly young) fellow soldiers who aspire to glory, who take their bearings from what others think of them, would prepare him to become supremely self-concerned. A genuinely “moral” education would somehow inculcate selflessness in the agent, not a desire for honor. On Irwin’s moralizing interpretation of the phrase “for the sake of the kalon” a mature (virtuous) person should repudiate his own youth, for he would see himself then as having been far too oriented to others. By contrast, Aristotle states that it is “fitting” (harmozei: 1128b16) for a young person to feel, and for an older person to have felt, aidôs. The mature man should thus appreciate, even admire, who he was. After all, even back then he acted “on account of virtue” and

---

on a desire for what is kalon.

A quick rebuttal is available to Irwin here. He could follow Korsgaard in arguing that a young soldier’s desire to look good in the eyes of his comrades and superiors - that is, his desire for the kalon in the form of honor - prepares him for true virtue insofar as it makes him “receptive to the more mature state of autonomy.” This is just one element in her argument that Aristotle and Kant, the moralist par excellence, are actually closer than is usually thought. As she puts it, both “think that what gives an action more value is the fact that it is chosen for its intrinsic rightness.” Furthermore, she quotes Kant in saying that “the love of honor is ‘the constant companion of virtue’” and thereby helps teach a young person to act morally; that is, unselfishly. She does not explain how this educational process would work. In any case, by her lights, “Kant and Aristotle need have no disagreement about this kind of case [aiming for honor] at all.”

While this line of thought is plausible, it does not adequately address the startling proximity into which Aristotle brings the citizen soldier and the genuinely virtuous man. The former acts from a desire for the kalon while the latter for the sake of the kalon. If, as Irwin has it, the property to which Aristotle refers in using “kalon” in the ethical works is “moral rightness,” then both would be acting for the sake of others. But this is manifestly untrue of the (young) citizen soldier who is driven by aidôs and the self-regarding desire for honor. Another account of how the kalon enters the life of the virtuous is needed.

6. For the Sake of the Kalon and For the Sake of the Virtuous Actions Themselves

Recall that Irwin claims that acting “for the sake of the kalon” (NE III.7) and “choosing [the actions that accord with the virtues] on account of themselves (di’ auta)” (NE II.4) are “inseparable.” As he puts it, “to say that it is kalon” - which for him means morally right and for the benefit of others - “is simply to say that it is to be chosen for its own sake.” The two formulations are inseparable indeed, but not because both share a similarly moralized conception of the kalon. To explain, consider the following hypothesis.

The genuinely courageous man, the one who acts both for the sake of the kalon and on account of the virtuous actions themselves, is guided by the following sort of imperative: ‘I must stay at my post because that is what I do; that is who I am. Not to do so would be to fracture or betray myself, and as such would be

24 Ibid., 205.
25 Ibid., 220.
aischron.’ In this formulation, aischron will come closer to “ugly,” and kallon to “beautiful” than Irwin would allow. For the genuinely courageous soldier acts for the sake of maintaining himself, of presenting himself to himself as a coherent, well ordered being; one who is kallon. Hence, he acts for the sake of the kallon. He acts on account of the virtuous actions themselves insofar as he does not look beyond them to consequences or for any reason other than their worthiness; that is, they are kala.

This hypothesis helps account for the proximity between the citizen-soldier who is driven by aidôs and so is self-regarding (and who is the younger self of the mature man), and the genuinely courageous man. The kallon figures prominently in the souls of both. In order to clarify how it does so, a deeper understanding of aidôs is needed. For this I turn briefly to Bernard Williams’s well-known discussion of shame.

Williams argues against the “progressivist” view that the Greeks, specifically Homer and the tragedians, whose characters are indeed often driven by aidôs, lacked a sophisticated concept of ethical responsibility. Thinking this results from the pernicious identification of the ethical with the “moral.” For the moralist, “I am provided by reason, or perhaps by religious illumination… with a knowledge of the moral law, and I need only the will to obey it.” While a shame-based system is not that, neither is it shallow or crudely heteronomous.

Consider what Williams has to say about Sophocles’ Ajax. Having been tricked by Athena into thinking that the animals in the pen are actually the leaders of the Greek forces who have humiliated him, he slaughters them. When his madness subsides and he realizes he is surrounded by the bloody entrails of dead sheep, he feels an overwhelming sense of shame. He can no longer live with himself and so resolves to commit suicide. He expresses his conviction in the form of an imperative: “now I am going where my way must go.” This is not a moral imperative but nonetheless, according to Williams, it is categorical. As he puts it, Ajax’s words mean “that he must go: period.”

From a Kantian or moral perspective, Ajax’s imperative does not deserve to be counted as categorical. On this view, it is merely hypothetical in that it represents an action chosen “relative merely to what the agent wants to do….or to avoid what he fears.” Even worse, it takes it bearings from what others will say of the action, and as such is ethically shallow and bound by convention. Williams disagrees. While it is true that shame is fundamentally related to being seen, it is not true that this makes an aidôs-driven action superficial or crudely heteronomous. As he puts it, even if shame does require an other, that “other may be identified in ethical terms. He…is conceived as one whose reactions I would

27 Bernard Williams, Shame and Necessity. (Berkeley: University of California Press, 1993), 64.
28 Ibid., 95.
29 Ibid., 76.
In this, its most developed form, where it no longer is dependent on an actual audience (and therefore on mere convention), shame incorporates an internalised other capable of holding an “inner personal conviction.” And this makes possible the sort of categorical imperative, or necessity, that drives Ajax. As Williams puts it,

This sense of necessity lies in the thought that one could not live and look others in the eyes if one did certain things: a thought which may be to varying degrees figurative but can also be in a deadly sense literal, as it was with Ajax. These necessities are internal, grounded in the êthos, the projects, the individual nature of the agent, and in the way he conceives the relation of his life to other people’s.

For Williams, as for Aristotle, the person animated by shame can grow.

Shame looks to what I am. It can be occasioned by many things - actions…or thoughts or desires or the reactions of others. Even where it is certainly concerned with an action, it may be a matter of discovery to the agent, and a difficult discovery, what the source of the shame is, whether it is to be found in the intention, the action, or an outcome. Someone might feel shame at the letter he has mailed because it is …a petty and stupid response to a trivial slight; and the shame is lightened, but only to some degree, when it turns out that the letter was never delivered. Just because shame can be obscure in this kind of way, we can fruitfully work to make it more perspicuous, and to understand how a certain action or thought stands to ourselves, to what we are and to what realistically we can want ourselves to be.

Williams’s aidôs-driven agent advances beyond shame understood as seeing himself through the eyes of others and toward shame as seeing himself through an internalized other who holds ethical convictions he himself embraces. His agency deepens and he becomes more self-sufficient. The agent must act the way he does because that is who he is. As Williams so sharply puts it, “shame looks to what I am.”

Following Williams on shame helps explain why, according to Aristotle, aidôs is preparatory for a life of genuine virtue. To return to my hypothesis: as a youth, the genuinely courageous man was driven by aidôs and the desire for recognition by his comrades and elders. As he developed, he internalized these others. He

---

30 Ibid., 84.
31 Ibid., 95.
32 Ibid., 103.
33 Ibid., 93.
came to think for himself. But he never lost his desire for the visibility of the kalon. In acting on account of virtuous actions themselves, he seeks this visibility in the form of a self-image or representation. In short, the virtuous person is prompted to act by the desire to see himself as kalon.

To clarify this interpretation of Aristotle, which I am not attributing to Williams, consider Lear’s argument concerning the relationship between moral virtue and “the fine.” Pace Irwin, who is adamant that Aristotle’s “use of [kalon] does not express a conviction about morality and beauty,” Lear thinks Aristotle has “good reason to make beauty central to his account of virtue.” For her, the key ingredients of the kalon are “effective teleological order, visibility and pleasantness.” The latter two items are particularly important here. As she says, “the experience of one’s actions as beautiful is, we might say, the mode of the virtuous person’s apprehension of their goodness.” In turn, this apprehension is the source of the pleasure that Aristotle stipulates that the virtuous person must feel. As Lear puts it, “acting well is a proper source of self-regarding pleasure.” It is a pleasure that comes from apprehending oneself, not simply as someone who acts to benefit others (although of course one may do so), but as kalon…as beautiful to behold.

To suggest reasons why Lear may be right here, consider the following three passages from the NE.

1. In NE III.4, Aristotle asks whether the object of “wish” (boulêsis) is the good or “what appears to be good.” If it is the former, then the wish of someone who “chooses incorrectly” - that is, pursues a bad end - in fact does not have an object and so is no real wish at all. If it is the latter, then a pernicious relativism sets in: whatever “seems so to individual” would have to be counted as good. Aristotle resolves this dilemma by stating that “to the ethically serious person (spoudaios) the object of wish is what is truly [good], while to the base person it is whatever happens [to appear good]…for the ethically serious person judges each situation correctly, and in each of them what is true appears to him.”

On the one hand, Aristotle’s reasoning here may seem uninformative or even circular. The true good is what appears to be good to the virtuous person, who in turn is identified by having the capacity to discern the true good. But this offers no independent specification of what the true good is, and so provides no guidance. On the other hand, Aristotle’s reasoning here is illuminating, at least given his own standards of philosophical adequacy in the NE, whose project (as argued in Part I) is a kind of phenomenological mapping of the ethical life. While it does not disclose criteria of the true good, it does tell much about the character

---

36 Ibid., 117.
37 Ibid., 117.
38 Ibid., 128.
of virtuous agency. Simply put, the virtuous person can trust himself. He can rely on himself to determine what is good. He is supremely confident and needs only look to himself in order to determine in what direction he should go.

The above is compatible with the hypothesis offered earlier: the courageous man acting for the sake of the kalon is driven by his sense of himself as a kalon being who strives to be, first and foremost, himself. He acts on account of virtuous actions themselves because he issues to himself the categorical imperatives governing them: “this I must do: period. For this is who I am. Not to do so would be aischron; ugly, ill-becoming, ill-fitting. It would be a betrayal of myself.”

2. My hypothesis is supported by the results of Part II above. There the virtuous agent was distinguished from the engkratês. The latter struggles against his passions or desires, and is finally victorious. The former experiences no internal conflict and is instead a harmonious being. As Aristotle says, the knowledge of what good actions are “has grown into him” (sumphuênai: 1147a22) such that he is a fully integrated and harmonious being.

In this context, Williams’s characterization of the Kantian-moral view provides a useful contrast: “what I am, so far as it affects the moral, is already given, and there is only the matter of discerning among temptations and distractions what I ought to do.”39 In other words, on the Kantian view engkrateía is the highest moral achievement. Fighting temptation simply is the moral life. By contrast, for Aristotle, while fighting temptation (and winning) is admirable, it pales in comparison to true virtue. For this is a kind of wholeness. The virtuous agent acts for the sake of the kalon. He acts for the sake of maintaining and presenting to himself his wholeness, which is kalon.

3. In NE II.4 Aristotle states the virtuous agent “acts by choosing and choosing [the actions] on account of themselves (di’auta).” But it is important to note that this is only the second of the three requirements for virtuous agency. The first is that the agent acts “knowingly (eidôs),” the third is that he acts from a “stable and unwavering manner” character trait. All three relate to the kind of person he is. He is a unified being who trusts in his own judgment and as such confidently engages in certain actions simply because he thinks they must be done. Hence, the self-generated imperative of the genuinely courageous soldier is, ‘I must stay at my post: period. That is what I do. That is who I am. Not to do so would be aischron, which I am not.’

39 Williams, Shame and Necessity, 95.
8. Conclusion

The complications of the NE are legion. Indeed, sometimes the book feels like a vast and churning sea, in which a reader, at least one attempting to make sense of it as a whole, will surely drown. A sterling example of this is the problem, raised by Crisp and Irwin (and many others), of how to reconcile Aristotle’s conviction that happiness is the ultimate human good, which implies that self-concern drives the ethical life, with his discussion of the virtues, at least some of which are other-oriented. This paper makes no pretense at a comprehensive reading of the NE or a solution of this problem. It does, however, suggest, that trying to determine and distinguish the meaning of three of Aristotle’s descriptions will be useful in the attempt to do so.

1. The courageous man acts for the sake of the kalon.
2. The virtuous person acts on account of the virtuous actions themselves.
3. The person driven by aidōs and a desire for the kalon in the form of honor is most similar to the courageous man.

As Irwin says, the first two are inseparable. They are not identical. An agent acts for the sake of the kalon insofar as he insists on maintaining his identity and sense of self as kalon. He acts on account of the virtuous actions themselves insofar as he is driven by categorical imperatives. He must remain at his post even if his life is at risk; period. Not to do would destroy himself. When he was a younger soldier he also remained at his post, but he did so because he would have been ashamed to do otherwise. As a mature man he no longer feels this way. But he is still driven by the kalon, with its essential feature of visibility, even if he fundamentally concerned with how he appears to himself.
Aporetic State and Extended Emotions: the Shameful Recognition of Contradictions in the Socratic Elenchus

Laura Candiotto
University of Edinburgh
Department of Philosophy
candiottolaura@gmail.com

ABSTRACT
The Socratic elenchus is a procedure which tests out the consistency of the interlocutors’ beliefs. To this end, it is necessary to carry out, alongside the renowned Socratic strategies (questioning, examples, definitions, etc.), also an emotional process acting inside reasoning and where shame has a leading role. The aporetic state is a good example of the collaboration of emotions and reasoning, growing from the shameful recognition of contradictions. It is a cognitive and emotional acknowledgement of errors that pushes the subject to transform his/her behaviour. The use of emotions is not merely a rhetorical strategy for argumentation; emotions are the elements that embody knowledge into a practice capable of transforming life into a good life thereby determining the rational way of living for flourishing.

The recognition of mistakes does not happen just “in the head” but is “extended” in the public environment that permits the generation of shame. This is the case, not only because shame is a “collective emotion” but because the audience is a necessary component of the catharsis.

My main thesis concerns what I call the “extended elenchus”, a process based on the extended nature of the aporetic state. The first section highlights the “necessity thesis”, or the role of emotions in reasoning; the second focuses on shame as an epistemic emotion and on the cognitive role played by the audience in the implementation of the “system of shame”; the third addresses the role of cathartic and zetetic aporia.

KEYWORDS
Socrates, elenchus, aporetic state, shame, extended mind, extended emotions

Introduction

The Socratic elenchus is a procedure which tests out the consistency of the interlocutors’ beliefs. To this end, it is necessary to carry out, alongside the

---

1 This paper arises from the project “Emotions First”, which I am carrying out at the University of Edinburgh as a Marie Curie Research Fellow, thanks to a grant from the EU Commission. A previous version of the paper was presented at the 2nd Annual Conference of The European Philosophical Society for the Study of Emotions, University of Edinburgh, 15-17 July 2015. I would like to thank the organizers and those with whom I had the opportunity to discuss my approach, in particular Alessandra Fussi, John Dillon, Edwart Harcourt, Niels Hermannsson, Paolo Maccagno, Dory Scaltsas, Clerk Shaw, Jan Slaby and Luigi Vero Tarca. Their contribution enabled me to rethink and improve the first version of the paper.
renowned Socratic strategies (questioning, examples, definitions, etc.), also an emotional process acting inside reasoning and where shame has a leading role. The dialogue takes the form of a refutation, which is different from Eristic due to its moral purpose: thanks to the critique that it engenders, the dialogue enables the interlocutor to realize his own inferiority, to the extent that he must recognize that he does not really know what he thinks he knows. This feeling of inferiority aims at inducing in him the desire to respond by changing his lifestyle, recognizing in himself the truth he sought. Therefore, the Socratic refutation plays on the feeling of shame that enables the interlocutor to admit his own ignorance.

This aspect is crucial as shame played a fundamental role in Greek civilization. The Athenian citizen had to avoid all situations in which he could appear weak, or he was lost. But Socrates, arguing in the public square, does precisely that: he shows his interlocutors that they do not know what they think they know; he ridicules them and, above all, strips them of their claims.

My thesis is that the audience listening to the refutation was not a mere spectator but had an active role within an extended cognitive process that included Socrates, his interlocutors and the audience. This hypothesis, which I will explain in its main epistemological facets, is based on the recognition of the particular historical period and of the specific functioning of the dialogues.

Socratic dialogues were written not only by Plato but also by other writers. They represent a particular form of writing emerging in a period in which the oral performance was the most important. Based on the analysis of this historical context, I think it is possible to claim not only that Socrates’ dialogues took place in the public square, but also that the dialogues written by Plato and by other disciples of Socrates were performed and read in public. In writing the Socratic dialogues, Plato had in mind a specific and well defined external audience: an audience on which he wanted to impress a conceptual change, therefore a change in values and political approach.

My main thesis concerns the existence of what I call the “extended elenchus”, a process based on the extended nature of the aporetic state as catharsis in the drama, and is grounded on various premises that I am going to explain in specific sections, providing also items of textual evidence. The first section highlights the “necessity thesis” or the role of emotions in reasoning; the second focuses on shame as an epistemic emotion and on the cognitive role played by the audience in the implementation of the “system of shame”; the third addresses the role of cathartic and zetetic aporia.

---


1. The Socratic method: emotions in reasoning

My main and general thesis is that the awareness of the role of emotions in reasoning represents a cornerstone of the Socratic method. I argue that the view that Socrates propounded is the reverse of the so-called Socratic intellectualism: emotion is the more primitive guide to the discovery of the good, since it shows the way to reach knowledge, and has the power to transfer it into our lives. In this perspective, knowledge concerns every aspect of life and reality. Accordingly, not only do we reach knowledge through emotions, but emotions are also the powers through which knowledge can impact our lives.

The elenchus aims at improving the interlocutor through a process of purification that is capable of changing his whole existence: the goal of the Socratic method is to give birth to a correct mode of life, and, as we shall see, it is precisely the literary aspect of Plato’s dialogues that makes this possible. The literary form allows us to understand the performance of the dialogues: Plato was well aware that the diegetic-mimetic form of the dialogues allowed the public to participate actively in the process. This participation does not mean, in my opinion, just that the audience could identify with the interlocutor, mirroring his emotional state, but also that the audience played a fundamental role in the entire cognitive process engendered by the dialogue. For this reason, the hypothesis of the extended mind and, more specifically, of the extended emotions, as we shall see in detail in section 4, seems central for understanding this dynamic.

Plato argued that emotions are necessary to reach the truth: emotions are not sufficient by themselves but – and in this perspective we can maintain a moderately rationalist approach – they act within reasoning to enhance the epistemic process.

Plato was the first to explore and gain significant insights into the relation between emotions and reasoning: for instance, Plato’s *Sophist* 230b4-230e5 (the “noble sophistry” passage) clearly shows the bond between the logical and the

---

8 On the problem concerning whether emotions be necessary or only sufficient to produce a moral judgement, see Sauer, H. 2012. “Are Emotions Necessary and Sufficient for Moral Judgment?”. *Ethical Theory and Moral Practice* 15:95–115.
emotional levels that can be found in the Socratic elenchus.

In this passage it is possible to isolate a quotation which proves the extension of the cognitive process: Plato argues that the release is sweet for those who attend as listeners and is firmly established for those who undergo the process (the refutation). Besides indicating the presence of an audience, this passage demonstrates how the action of the elenchus affects also the listeners of the dialogue.

Refutation brings the interlocutor to the aporetic state, understandable as an embodied and embedded experience of mistakes, the first step for grasping the truth. The aporetic state is a good example of the collaboration of emotions and reasoning, growing from the shameful recognition of contradictions. It is a cognitive and emotional acknowledgement of errors that pushes the subject to transform his behaviour. The use of emotions is not merely a rhetorical strategy for argumentation; emotions are the elements that embody knowledge into a practice capable of transforming life into a good life, thereby determining the rational way of living for flourishing.

In order to be complete, the elenctic purification needs also a psychological cleansing: in this perspective, it can be obtained only through the collaboration between rationality and emotions, mainly shame. Socrates uses shame as a tool for healing the illness of one’s soul and style of life. That is the effect that Socrates aims to achieve through the elenchus, namely the state of aporia of the interlocutor. The aporia is a mental state of perplexity and being at a loss, that involves feelings, which in turn play a role in the cognitive development of the interlocutor. The aporetic state is not a purely cognitive state; it is a cognitively-motivational state involving emotive elements.

The turning point between the refutation and the maieutical production of the thesis consists in the acknowledgement of one’s own inadequacy, a sense of inferiority – a situation that is captured by the Greek terms *aidos* and *aischyne* and which unfolds as an aporetic condition, in other words as awareness of contradiction. The recognition of mistakes does not happen just “in the head” but is “extended” in the public environment that permits the generation of shame. Frustration and the feeling of shame as a result of the dialogic challenge is thus experienced by Socrates, by the interlocutor and by the audience.

---

issue of the journal *Educational Philosophy and Theory*, Routledge, forthcoming.
2. The central role of the audience for the dynamics of shame

Arguably, the kind of shame pursued by Socrates as a factor of elenctic transformation represents a productive social use of the “system of shame”, adopting the distinction proposed by Bernard Williams between a positive and a negative shame.\(^{10}\)

Shame is often defined as a primitive emotion: it plays a fundamental role in the way our personal identity is constructed through relationships with others (both with peers and hierarchically). In Greek times, dialogues were not for the most part private, but more often public conversations. Acceptance of the refutation therefore had a social value, given the way in which social status was constitutive of individual identity. In this perspective, accepting a public refutation could imply accepting a change of identity.

However, Plato’s dialogues only rarely give testimony of a successful transformation occurring in the interlocutor. This is due to the interlocutor’s attitude towards shame: the feeling of shame can be accepted as a means for self-transformation or hidden to protect a social status. Shame is frequently concealed (through the psychological mechanism of the “shame of shame”) due to social reasons.

In *Euthyphro* (12b4-c1), Socrates argues that where there is shame there is also fear of losing face, and we blush for this reason. What makes us blush is the fear of losing our reputation or, conversely, as Socrates says literally, of acquiring the “reputation of an evil man” (12c1). Shame is in fact experienced in front of other men (15d4-e2).

The interlocutor’s identity depends on social recognition, namely the social attribution of a role; therefore, the interlocutor can hardly accept to forego this safe foothold by openly admitting his errors. Arguably, by outlining the distinction between these two types of shame, it is possible to notice how the purification of the interlocutor implies a turning—or “break”—point within the dialogue, which influences the epistemic outcome of the aporia. Shame as a tool of transformation conduces to the generative phase of the maieutic process and shame as an obstacle to transformation that functions as a resonator for the audience.

In *Charmides* we find, embodied by Charmides and Critias, the expression of these two different ways of experiencing shame. Modesty is connected with shame when Charmides blushes as he does not know whether or not he is wise: thanks to shame, Charmides recognizes his own inadequacy, accepting – at least at the age when he is represented in the dialogue – that he should be accompanied by

---

Socrates in the research. Critias’ shame,\textsuperscript{11} however, hinders the research, as it triggers a defence mechanism that causes Critias not to admit his shortcomings.

In passage 164 d Critias says he is not ashamed to admit his mistakes; however, subsequently he fails to realize what he had set out to do. As a result of the inconsistency between words and deeds, and of the inability to admit the condition of being in aporia, Critias does not admit that he does not know. Therefore, he is unable to make the first step towards a sincere research: he does not recognize his own ignorance. As we can see, shame, depending on the character and the social role of the interlocutor, can either engender or hinder the process of research.

The concealment of shame, however, is not an evidence of the failure of the Socratic method, which uses shame as a tool for transformation. On the contrary, Plato uses the defence of the interlocutor (the way he conceals shame) as an element that, thanks to the involvement of the audience, backfires on the interlocutor himself.

Emotions collaborate with reason not only to purify the soul, but to deliver a message to the audience: the necessity to be aware of the inadequacy of contemporary politicians and teachers. When interlocutors try to protect their social image, their standing is unavoidably compromised. By trying to save face, they lose face. More specifically, interlocutors cannot protect their socially ratified identities insofar as, by attempting to do so, they demonstrate their unwillingness to admit their errors. The audience, realizing that the interlocutor does not acknowledge the shortcomings which, thanks to the refutation, emerged clearly in the dialogue, understands that he is not the person he believes himself to be. This mechanism, which I call “outreach elenchus”, occurs mainly when the interlocutors are politicians, sophists, and rhetors. In the outreach elenchus Socrates carries out directly the refutation of the interlocutor, but the elenchus affects indirectly also the audience.

In other words, this elenchus increases in size and incorporates the dialogic context, like a stone thrown in a pond that produces a series of increasingly larger circles. This mechanism, however, as we will be explaining in detail in section 4 by emphasizing its extension, demonstrates not only the effect of the Socratic intervention on the audience, but also the role the public plays in the refutation, functioning as a resonator and leading back the refutation to Socrates’ interlocutor.

In this perspective, the movement of the refutation is not just similar to that of a stone thrown into a pond, but also to that of a boomerang, which comes back to those who have launched it and which, when used for hunting, allows the

\textsuperscript{11} Platone, \textit{Charm}. 169 c-d.
hunter to hit an object whilst apparently being thrown towards a different direction. In this sense we could say that when Socrates points indirectly to the public, he does so in order to hit his direct interlocutor: he just needs the active participation of the public, which allows the boomerang to bounce back. The movements exemplified here represent the cognitive and emotional dynamics that develop between Socrates, the interlocutor and the audience, and that find their realization in a specific moment of Socratic dialogue, the aporetic state.

A clear example of this mechanism is present in the Gorgias and, in particular, in the figure of Callicles. In 461 c Polus says that Gorgias was ashamed to maintain certain statements, e.g. that he did not know what justice was and that, as he did not know, he could not teach it. The shame that is ascribed to Gorgias is therefore caused by the recognition of ignorance; Polus himself, later, will become a victim of the mechanism of shame by recognizing that, had he expressed his thoughts, he would have fallen into contradiction. Callicles, however, manages to avoid these consequences – the recognition of ignorance and contradiction – exposing accurately the Socratic strategies and individuating in which point of the dialogue Polus gave in and “found himself entrapped in your discourses and could no longer open his mouth, ashamed to say what he was thinking”.

Callicles is not ashamed as he does not identify with the values that are at the basis of the critique – this is why Socrates had to use another strategy with him, which is based on the extension of the elenchus. In my perspective, the tenacity of Socrates in continuing his dialogue with Callicles is not moved by the hope of changing his lifestyle – this interpretation would ascribe to Socrates a certain naivety – but by the attempt to express explicitly the consequences of such a vision and lifestyle, in order for the listeners and the audience to realize Callicles’ shortcomings and to rebound onto Callicles a critique that compromises his image.

These tools are not specific to every Socratic elenchus, but Socrates uses the public as a vehicle for the extension of the elenchus when he is dealing with those who represent the values of the society he wants to criticize. Accordingly, my claim is not that all the elenchi are extended, but that it is necessary to recognize the existence of this particular form of elenchus. The Socratic elenchus is contextual and is configured into different ways according to Socrates’ strategic purposes.

14 For our study, it is important to highlight how these changes impact also on the emotional aspect of the dialogue. Cf. Brisson, L. 2001. “Vers un dialogue apaisé. Les transformations
LAURA CANDIOTTO

This interpretation could be criticized by referring to another passage in the *Gorgias*\(^{15}\) where Socrates tells Polus that there are two types of elenchus, the rhetorical elenchus, which is usually used in courts, and the dialectic elenchus, the type used by Socrates. The sharp distinction between these two types of elenchus lies in the fact that, to have value, the first needs a large number of witnesses, whilst for the second to succeed it is sufficient that a single witness recognizes the truth. This difference is linked to the specificity of the Socratic maieutics, which addresses each time a single interlocutor, and to Socrates’ refusal to seek consensus and approval from large audiences, as did the orators and politicians of the time. The objection would thus emphasize the fact that Socrates’ intervention is usually directed to a single party.

I would reply to this objection by highlighting how the difference between rhetoric and dialectic elenchus is not only methodological, but also related to the Platonic construction of Socrates’ public role, which should be understood in opposition to the masters and politicians of his time.\(^{16}\)

Plato is aware that the elenchus has an effect on the listeners, he even uses the audience to induce the elenchus to bounce back on the interlocutor. However, he does not unmask this mechanism for two reasons: firstly, because he is drawing Socrates’ image in contrast with that of the masters of the time, and secondly because he wants the strategy to be successful (if the rules of the game were revealed they would lose their effectiveness). Moreover, it is also true that Socrates seeks the consent of only one individual: to do this, however, he needs the active participation of the public in the aporetic state. In so doing, he obtains also the political and rhetorical effect that Plato could not ascribe to Socrates, given the apologetic construction of his figure.

3. *Aporetic state*

According to Anne-Marie Bowery shame, which is linked to the physical reaction of blushing, indicates exactly the aporetic state and the difficulty of recognizing what has been discovered. The interlocutors blush when they have to admit what they would rather not admit, or when they do not know how to respond. According to Bowery, the phenomenon of blushing indicates a turning point of

---

the dialogue, the emergence of real possibilities for a dialogic development.\textsuperscript{17}

In relation to the aporetic state, it is also important to remember that the Socratic method does not reach a stable definition but, in a Daedalic fashion,\textsuperscript{18} it puts in motion all the notions proposed by the interlocutor in order to lead him to recognize his own ignorance. In my view such an outcome, often considered as a skeptic one, should not be understood as an epistemic failure: the object of the method was not to attain a stable definition – as in the case of Prodicus’ method – but to carry over the purification from error, which can be achieved only through conceptual contradiction and becoming ashamed of oneself. The Socratic method, despite having a strong sophistic connotation in the use of linguistic and rhetorical strategies, could therefore be turned against the sophists themselves, who could be accused of selling a knowledge that was not as stable as they claimed. The aporetic outcome of the Socratic method can thus be understood as a place in which rationality comes to a standstill, where the paradox replaces firm knowledge and contradiction serves as the best medication against the assumption of wisdom. Exactly in its negativity aporia provides the consciousness of errors as the necessary starting point to wisdom. In this way, aporia is not only cathartic, as pointed out by the traditional approach, but also zetetic. The zetetic aporia underlines how solving particular aporiai is part of the search for knowledge. This conception of aporiai as puzzles to be solved is not only central for Aristotle (i.e. \textit{Met.} B1. 995a34–b1) but also meaningful for the Socratic elenchus.\textsuperscript{19} The roadblock is also “a breakthrough (\textit{euporia}), pointing to the right direction in which to pursue an answer to the question posed by the dialogue”.\textsuperscript{20}

Studies on the so-called epistemic emotions are central to the cognitive phenomenon we are describing: they highlight how emotions – being conceptually vital, and emerging in the course of a practically motivated enquiry – are necessary for thinking. One aim of this paper is to demonstrate how, in the Socratic method, the feeling of shame, connected to the above-mentioned aporetic status, represents also an epistemic emotion.\textsuperscript{21} Socrates was persuaded that deep


\textsuperscript{19} According to Vasilis Politis it is necessary to distinguish these two types of aporia, highlighting the point that only the second one refers to puzzlement in itself and, therefore, to the significance of the question as the main drive for the research (see p. 107-109). Politis, V. 2006. “Aporia and Searching in the Early Plato”. In J. Lindsay, V. Karasmanis, eds., \textit{Remembering Socrates: Philosophical Essays}, 88-109. Oxford-New York: Clarendon Press.


beliefs revealed by shame were true beliefs about what was right or not right to do: shame played therefore a central role in the practice of wisdom.

The Platonic paradigm of knowledge as a vision (originating from the attribution of ideas as object of noein) cannot be found in the Socratic method, except as its negation: as a matter of fact, the Socratic discourse leads to a non-vision, to the incapability to see a way out (cf. the etymology of aporia) for thought. However, the value of the Socratic elenchus should be grasped intrinsically and in its radicalism: its possible positive value should not be sought in the generation of further knowledge (although this process does take place with the transition from elenchus to maieutics), but in the way it engenders an ethic of care that is able to transform the recognition of powerlessness in the constant search for good deeds. It is necessary to emphasize this aspect in order to grasp the intrinsic epistemic valence of the elenchus, which should not be considered only as a pre-condition to reach a subsequent state of knowledge. The aporetic outcome of the elenchus, therefore, should not be understood in a passive sense: the strength of the aporetic event requires a transformative process that allows us to find, within negativity itself, the key to imagine an otherness.

4. The extended elenchus

The study of the primary role of emotions in the Socratic elenchus, as well as the research on the epistemic nature of shame and on the zetetic character of purification through aporia, allow me to propose the thesis that the Socratic elenchus is extended. This means that not only the purpose of the Socratic elenchus is external (e.g. in relation to lifestyle), but also its genesis.

I argue that the aporetic state is achieved in the elenchus, not only in the interlocutor’s mental state; the state is the conclusion of the elenchus that is a shared cognitively-motivational state of both interlocutors, Socrates and the dialogue-partner. My position is that the elenctic aporia is the external shared dialogical embodiment of the cognitively-motivational state of the two interlocutors in a Socratic elenchus.

The theory I employ for explaining the shared state achieved through the elenchus is the theory of the extended mind and of the extended emotions.

The theory of the extended mind – a form of active externalism, for which the environment constantly drives one’s intellect in an ongoing way – refers mainly to the way in which the human mind extends itself in external technologies:

---

although this would seem to be very far from the classical age, I believe that Socrates “was using” the audience in the same way as today’s mind “uses” technologies, for a well-defined cognitive purpose: not only the purification from false beliefs, but also (and here we see the need for externalization of the cognitive phenomenon) to bounce back to the interlocutor a conceptual shift concerning his role. According to this model we regulate ourselves through other people: they are there “for me”, they are resources which I use instrumentally, for my own self.

Socrates uses the audience to achieve his objectives: an extended conceptual change, able to impact the ethical and political behaviour of the audience. The cognitive extension takes place especially via the affective channel: in this perspective, I think the configuration of the extended mind as extended emotions is even better recognized as a conceptual paradigm for understanding the phenomenon of the aporetic state. Extended emotions are therefore part of the Socratic method: Socrates uses them as a strategy and a tool to achieve the aporetic status and, through it, a cognitive transformation.

The theory of extended emotions proposed by Jan Slaby\textsuperscript{24} goes in the direction of “collective emotions”, to be intended both as emotions “common” among the members of a group, and as emotions constituted by all the members of a group at the same time. For this aspect it is fundamental to refer to Printz’s work,\textsuperscript{25} even though Slaby holds his distance from the perceptual framework and emphasizes the rich phenomenology of affective states, drawing on Helm in regard to the systematic interrelatedness of the instances of momentary feeling,\textsuperscript{26} and attaining the concept of “phenomenally extended emotions”.\textsuperscript{27} Slaby proposes examples of emotions which are very significant for our theme: in his opinion it is possible to encounter extended emotions not only in the general social-interactive domain, but also in the context of art reception and in dialogical interplay.

For Adam Carter, Emma Gordon and Orestis Palermos\textsuperscript{28} emotions extend beyond the agent’s body to aspects of its dynamic environment. Their proposal is to understand the hypothesis of extended emotions as a novel application of the hypothesis of extended cognition, claiming that, if understood within this conceptual paradigm, their characterization is less radical. Their defence depends

\textsuperscript{26} In “Relational Affect” (paper delivered at the 2nd Annual Conference of The European Philosophical Society for the Study of Emotions, University of Edinburgh, 15-17 July 2015) Slaby underlines however how, differently from Helm, he assumes that the relational affect is from the outset transindividual.
\textsuperscript{27} Slaby, J. 2014, 42.
therefore on the justification of the hypothesis of extended cognition.

For our analysis about the cognitive process pursued by the Socratic elenchus, it is enough to highlight here that knowledge does not just happen “in the head” of the interlocutor but is “extended” in the public environment, and this allows the generation of shame. It is not only a question of location but also of determining the type of knowledge that is realized: such knowledge is not just “shared” with both the speaker and the audience; it is also extended in the sense of enhanced or maximized. This also clarifies how the regulation of the self always involves the other. Shame is really experienced also by the audience; it is not a “fictional shame”. Not only because shame is a “public emotion” but because the public and dialogical context is a necessary component of the catharsis, through what I call “outreach elenchus”, i.e. the public act of purification, which I have already described in its essential features in the section devoted to that topic.

Therefore, Socrates, his interlocutors and the audience form a group: to understand this aspect we must remember that the audience of the Socratic dialogues is not a generic set of listeners, but a very specific audience that Socrates wants to influence using the instrument of the outreach elenchus to trigger an extended elenchus. The audience is composed of the Athenian intelligentsia, a group that has a great weight in the political constitution of morals and customs. As claimed by Williams, one does not depend on generic others, but only on a few others, those whose way of judging is shared by the agent. The cognitive dynamic underlying the outreach elenchus is that of an extended elenchus, which expresses the externalization of mind and emotions. The fact that Socrates aims his method – albeit indirectly – at the audience enables us to grasp not only the embodiment of knowledge, but mainly its external origin: by bouncing back the elenchus towards the interlocutor, the audience makes it more powerful. This process takes on not only a political and educational valence towards the audience, but also a cognitive and epistemic significance. It is a synergic process of transformation of both the subject and the environment. In particular, I would like to stress that this kind of elenchus not only purifies the audience, but it affects also the interlocutors who are the object of the confutation: it bounces back and obliges them to recognize the shame they had concealed. Socrates “uses” the interlocutors to ensure that the message he sends them - through the aporetic state experienced by Socrates himself and by the interlocutor - goes back enhanced.

Andy Clark describes the cognitive process as “the actual local operations that


realize certain forms of human cognizing include inextricable tangles of feedback, feed-forward and feed-around loops: loops that promiscuously criss-cross the boundaries of brain, body and world”. The movement that I described previously, referring to the expansion of circles on the water produced by a stone thrown in a pond, and to the trajectory of a boomerang, which has the ability to turn back and to hit an object without being thrown at it directly, should therefore be understood within this wide, circular process of continuous entries and exits, intersections and links between the mind, the body and the world. Furthermore, the circularity of this process is expressed by the movement of the boomerang: aerodynamic forces generate a twisting moment that causes the ‘gyroscope’ to proceed and to move on a circular path.

Moreover, the emphasis on the cathartic connotation of aporia within the paradigm of extensive knowledge allows us to understand how such connotation does not imply a passive stance but defines an immediately active source of knowledge, reinforcing therefore the active externalism model.

Therefore, the distinction between cathartic aporia and zetetic aporia we mentioned earlier, referring to the interpretation of Politis, should be considered as an expression of the strength of aporia as a tool of extended knowledge. Shame is generated and has effects in the “society of dialogue”.

This emotional knowledge is realizable just in the shared and cathartic setting of the drama.

According to Aristotle, rhetoric favours working on logos in order to lead to the truth (representing a technique of persuasion through logos) but, by doing this, it does not negate the possibility of using the emotional dimension to influence the audience.

Moreover, Aristotle was the first to identify extended emotions in the practice of dramatic catharsis and to point to its significance and value for society. He analysed the role of tragedy in the theatre, showing that the tragic events in a play are acted out in the feelings of the audience. The embodiment of the emotions in the engagement of the audience with the tragic plot becomes a deliberative corrective in the audience, balancing their feelings of pity, anger and fear, in the light of the conceived calamity. The Socratic elenchus is staged by Plato in performative settings, and is a carefully crafted counterpoise between arguments and feelings in social reasoning interactions. Both practices, as analysed by these great philosophers, give us profound understanding of the interplay and mutual support of emotions and reasoning resulting in knowledge and cognition.

33 Aristotle, Rhet., book II.
Conclusion

At the core of wisdom – which is the main purpose of the Socratic practice and, therefore, needs to be understood as practical reasoning – there are emotions, seen as forces capable of directing actions towards the good of the individual and of the context in which he operates. Nevertheless, we should not think that only “positive” emotions can be a source of improvement in personal and collective life. This paper has attempted to demonstrate, as an outcome which is secondary to the one referring to the extended elenchus, how “negative” aspects of shame and the aporetic state may acquire a positive meaning insofar as they enable us to transform our unquestioned knowledge of reality. Notably, the most effective form of purification for human knowledge is accessible in a dialogic context, in a situation which implies relations with others. In this perspective the extended elenchus takes shape within a conception of extended cognition, where a primary role is played by collective emotions. In fact, the elenchus is the main strategy of the Socratic dialogue, which – albeit often unfolding as a dialogue between two individuals – implies also the presence of listeners who serve as source and receptacle for the process of purification.

Bibliography


---


VARIA
The Trolley: a Libertarian Analysis

Walter E. Block
Joseph A. Butt, S.J. College of Business
Loyola University New Orleans
wblock@loyno.edu

ABSTRACT

In the trolley case, an individual is faced with killing one man in order to save five equally innocent people. This philosophical conundrum pits deontology (do not murder) against utilitarianism (saving lives). Numerous non-libertarian commentators have weighed in on this challenge. The present paper offers a libertarian analysis of this case.

KEYWORDS

Murder, rights, deontology, utilitarianism.

I. Introduction

Foot (1967, 1) discusses “the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed.”

Another classic statement of the trolley challenge reads as follows (Thomson, 1976, 206):

David is a great transplant surgeon. Five of his patients need new parts – one needs a heart, the others need, respectively, liver, stomach, spleen, and spinal cord – but all are of the same, relatively rare, blood-type. By chance, David learns of a healthy specimen with that very blood-type. David can take the healthy specimen’s parts, killing him, and install them in his patients, saving them. Or he can refrain from taking the healthy specimen’s parts, letting his patients die.

If David may not even choose to cut up one where five will thereby be saved, surely what people who say “Killing is worse than letting die” mean by it must be right!

On the other hand, there is a lovely, nasty difficulty which confronts us at this point. Philippa Foot says² – and seems right to say- that it is permissible for Edward, in the following case, to kill:
Edward is the driver of a trolley, whose brakes have just failed. On the track ahead of him are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Edward can turn the trolley onto it. Unfortunately there is one person on the right-hand track. Edward can turn the trolley, killing the one; or he can refrain from turning the trolley, killing the five.

If what people who say “Killing is worse than letting die” mean by it is true, how is it that Edward may choose to turn that trolley?

Killing and letting die apart, in fact, it’s a lovely, nasty difficulty: why is it that Edward may turn that trolley to save his five, but David may not cut up his healthy specimen to save his five? I like to call this the trolley problem, in honor of Mrs. Foot’s example.

The philosophers usually credited with plagueing us with this challenge are Foot, 1967; and Thomson, 1976, 1985. Since then there have been literally hundreds of discussions of this issue. But not a one of them has been written from a libertarian perspective. The contribution of the present paper is to offer a distinctively libertarian analysis of the trolley problem.

We do well, then, to at least briefly discuss this particular political economic philosophy (Rothbard, 1998; Hoppe 1989, 1993; Huebert, 2010; Kinsella, 1992, 1996). Libertarianism is not a theory of law, nor is it an analysis of rights, nor, yet, ethics. Rather, it is an attempt to discern what the proper law should be; an analysis of the just use of violence. As a first approximation, there is the non-aggression principle (NAP). According to this axiom of libertarianism, it should be legal for a person to do whatever he wants to do, provided, only, that he refrain from initiating aggressive violence against anyone else and his legitimately owned property. Libertarianism does not say people should adhere to the NAP; that it would be right for them to do. Rather, it is an attempt to discern what the proper law should be; an analysis of the just use of violence. As a first approximation, there is the non-aggression principle (NAP). According to this axiom of libertarianism, it should be legal for a person to do whatever he wants to do, provided, only, that he refrain from initiating aggressive violence against anyone else and his legitimately owned property. Libertarianism does not say people should adhere to the NAP; that it would be right for them to do. It only maintains that if they do not so restrict themselves, they are in violation of libertarian law. But that is merely a first approximation. At bottom, libertarianism is a theory of what constitutes just punishment for law breakers.


2 In the view of Rothbard (1998, p. 88, ft. 6): “It should be evident that our theory of proportional punishment—that people may be punished by losing their rights to the extent that they have invaded the rights of others—is frankly a retributive theory of punishment, a ‘tooth (or two teeth) for a tooth’ theory. Retribution is in bad repute among philosophers, who generally dismiss the concept quickly as ‘primitive’ or ‘barbaric’ and then race on to a
It is only the first approximation of libertarianism to aver: do not violate rights. The more sophisticated version, is that libert is a theory of punishment: if you violate rights, we punish you in thus and such a manner.

This is in sharp contrast to the non libertarians who have tried to wrestle with the trolley challenge. For example, states Thomson (1976, 204, emphasis added): “Alfred kills his wife out of a desire for her death. Bert lets his wife die out of a desire for her death. But what Bert does is surely every bit as bad as what Alfred does. So killing isn’t worse than letting die.” Note that this philosopher is not concerned with proper law, and punishment, use of violence against, criminals. Rather, she, like so many others (Unger, 1992, 1996; Barcalow, 2007; Singer, 2005; Mikhail, 2007; Norcross, 2008; Otsuka, 2008; Hauser, et al, 2007), focuses on the good and the bad, the bad and the worse, what people should do and refrain from doing. It cannot be denied that of course there is a strong overlap between these two different concerns, but it is the divergences that distinguish the libertarian analysis of the trolley case from that of all others.

One more element of libertarianism. In this view, there is no such thing as positive rights. There are only negative rights (Block, 1986; Gordon, 2004; Katz, undated; Long, 1993; Mercer, 2001; Rothbard, 1982; Selick, 2014). A right implies an obligation. If Smith has a (negative) right not to be murdered, raped, robbed, then everyone else has a negative obligation not to murder, rape or rob him. It is incumbent upon all others to refrain from initiating violence against him. If Smith has a (positive) right to food, clothing and shelter, then all other people have a positive obligation to give him these goods. But, if so, then their own negative rights to their private property will have been violated.

We are now in a position to shed some light on the issue of killing an innocent person on the one hand, and letting him die on the other. Under the libertarian code, the former would be considered murder, and punished severely. However, allowing someone to die, standing idly by while another person perishes, would not be considered a crime. Now, it may not be nice to do so, it may not be moral or ethical to fail to come to the aid of a potential victim, but, qua libertarianism, that is not, cannot be, our concern. Rather, we focus, very narrowly, on whether or not an NAP violation, an uninvited (personal) border crossing has taken place, and, if so, what violent repercussion would be justified.

Pinker (2011, 328) states: “Most of us agree that it is ethically permissible to divert a runaway trolley that threatens to kill five people onto a side track where it would kill only one. But suppose it were a hundred million lives one could save discussion of the two other major theories of punishment: deterrence and rehabilitation. But simply to dismiss a concept as ‘barbaric’ can hardly suffice; after all, it is possible that in this case, the ‘barbarians’ hit on a concept that was superior to the more modern creeds.” See also Block, 1999, 2002-2003, 2003a, 2003b, 2004a, 2004b, 2006, 2009A, 2009B; Block, Barnett and Callahan, 2005; Gregory and Block, 2007; Kinsella, 1996; Morris, 1968; Olson, 1979; Rothbard, 1998, 88; Whitehead and Block, 2003
by diverting the trolley, or a billion, or — projecting into the indefinite future — infinitely many. How many people would it be permissible to sacrifice to attain that infinite good? A few million can seem like a pretty good bargain.”

What is the libertarian analysis of someone who diverts the trolley from its present track where it will kill, say, a billion people, and onto a path where it will murder a single individual? Utilitarians would speak out as one: such an act would be justified, since it would save one billion minus one lives. Most commentators would agree.³

But things are different for the libertarian. The diverter of the trolley is a murderer. We must not lose sight of this primordial fact. As such, he must be punished to the full extent of the law. And what, pray tell, is the full extent of the law? His life is forfeit, since he took someone else’s life. However, if the heirs of the victim, all of them, all five billion of them, forgive this murderer his crime, then and only then may he go free. But if even one of these heirs wants to impose the death penalty of our savior of five billion lives, he has the right to insist upon this punishment.⁴

This brings us to the libertarian concentration camp guard (Block, 2009). Here is the situation. All such criminals must murder 100 innocent Jews, gypsies, blacks, gays, other non Ayrians, per day. However, we posit that a libertarian guard can get away with murdering only 90 victims daily. If he goes below this figure, say, to a mere 89 or fewer murders, he will be found out, and himself summarily executed. Of course, our liberttarian hero engages in this dangerous pursuit not to murder 90 people daily, but rather to save 10, who, we posit, would perish were it not for his admirable⁵ and courageous acts. A week goes by, the war is over, and our libertarian murderer is in the dock at the Nuremberg trial. He has murdered 630 people, but saved 70. Should he be put to death for his evil deeds? Yes, unless all of the heirs of the 630 victims agree to forgive him his tresspasses.

Our hero may plead with these people: “I wish I could have saved your parent, your child, your spouse, but I could not. I could only save 10 people per day. That is why I embarked upon this dangerous acts of mine. If I had saved even as one more person daily, I would have been discovered, and prevented from saving

³ Thomson (1976, 206) states: “Edward may turn that trolley to save his five (people, at the cost of one life).” If so, she would certainly favor saving a billion people even though one innocent man must be murdered.

⁴ However, some of us will hold a ticker tape parade in honor of this murderer, and pin a big medal on this chest before the execution. After all, he is a hero. He lives, we may suppose, in a libertarian society and full well knows the penalty for murder: execution. Yet, he engaged in this heroic murder in order to save the lives of five billion people. We cannot do this qua libertarian, since this very narrow philosophy admits of no such actions. But, we can do this as decent human beings.

⁵ We say this not qua libertarian, which, we insist, is a very narrow philosophy, concerned only with justified punishment for criminal behavior. See Block 2001B, 2002, 2003B, 2004, 2006
any more victims of Nazi oppression.” Whether or not he prevails, he deserves, once again, a medal, a parade in his honor, and the thanks of all civilized people.

In order to further highlight the differences on the Trolley question between a libertarian and members of other philosophical traditions, I offer my responses to a popular query on this subject (http://www.philosophyexperiments.com/fatman/). My responses are in italics.

II. Should You Kill the Fat Man? Preliminary Questions

This activity is a treatment of some of the issues thrown up by a thought experiment called ‘The Trolley Problem’, which was first outlined by the philosopher Philippa Foot, and then developed by Judith Jarvis Thomson and others. But before we start properly, we need to ask you four preliminary questions so we get a sense of the way that you think about morality. There are no right or wrong answers. Just select the option that most corresponds to your view.

Question 1: Torture, as a matter of principle, is always morally wrong.

I answered no. As a libertarian, I have no views on this. No, I go further. Qua libertarian, I am precluded from having an opinion on this vital issue. Because it asks about morality, and libertarianism, in sharp contrast, deals only with what the law should be. So, I can only answer as a citizen, as a moralist. I can think of cases where it would be justified. For example, if the criminal tortured a victim, it would be just, under the libertarian code, to torture him back.

Question 2: The morality of an action is determined by whether, compared to the other available options, it maximises the sum total of happiness of all the people affected by it.

I responded in the negative to this one, too. Not because I have strong views about morality; I don’t. I answered in this way because utilitarianism turns me off, and this sounded pretty utilitarian to me.

Question 3: It is always, and everywhere, wrong to cause another person’s death - assuming they wish to stay alive - if this outcome is avoidable.

This one, too, got a thumbs down from me. Again, as a libertarian, I know of nothing “wrong” except that which violates the NAP. But, surely, given that the death penalty is justified, and as a libertarian I maintain it is (Block, 2006), it would be justified to execute a murderer, even assuming he wished to stay alive.

Question 4: If you can save the lives of innocent people without reducing the sum total of human happiness, and without putting your own life at risk, you are morally obliged to do so.
I responded positively to this question, since in my own vision of morality, it would be immoral not to save someone’s life if I could do so with trivial cost to myself. However, speaking as a libertarian, I would oppose laws requiring that we give aid to each other, as this would be a positive obligation, and libertarians support only negative rights.

Next question:

*Should You Kill the Fat Man? - The Scenarios*

You will now be presented with four different scenarios to test your moral intuitions against the answers you gave to the first four questions.

**The Runaway Train** (Scenario 1 of 4)

The brakes of the train that Casey Jones is driving have just failed. There are five people on the track ahead of the train. There is no way that they can get off the track before the train hits them. The track has a siding leading off to the right, and Casey can hit a button to direct the train onto it. Unfortunately, there is one person stuck on the siding. Casey can turn the train, killing one person; or he can allow the train to continue onwards, killing five people.

Should he turn the train (1 dead); or should he allow it to keep going (5 dead)?

- Turn the Train
- Allow the Train to Keep Going

I do not know what Jones “should” do. I do know that if he turns the train he is a murderer, if he allows it to keep going he is not. Based on the analysis offered above, if he turns the train to save four lives, he fits the bill of “heroic murderer. I had to answer, otherwise I couldn’t finish the quiz, so I filled in “turn the train,” making him into a heroic murderer.

I was then presented with the following:

Interesting, you do not believe there is any general moral requirement to maximise the happiness of the greatest number of people, yet you think that Casey Jones ought to divert the train. There’s no contradiction here, but it would be interesting to know what thoughts motivated your decision. For now, though, let’s see what you make of the scenario below.

**The Fat Man on the Bridge** (Scenario 2 of 4)

Marty Bakerman is on a footbridge above the train tracks. He can see that the train approaching the bridge is out of control, and that it is going to hit five people who are stuck on the track just past the bridge. The only way to stop the train is to drop a heavy weight into its path. The only available heavy enough
weight is a (very) fat man, who is also watching the train from the footbridge. Marty can push the fat man onto the track into the path of the train, which will kill him but save the five people already on the track; or he can allow the train to continue on its way, which will mean that the five will die.

Should he push the fat man onto the track (1 dead); or allow the train to continue (5 dead)?

- Push the fat man onto the track
- Allow the train to continue

I voted for pushing the fat man onto the track, so as to make Bakerman, also, into a heroic murderer. I might as well be consistent, I thought. But, I could as easily have gone the other way, if I wanted to reduce the incidence of murder.

Next, I was presented with the following:

That’s an interesting response. Previous research has indicated that most people disagree with you that it would be right to push the fat man off the bridge into the path of the train. However, your response is certainly consistent with your claim that Casey Jones should divert the runaway train so that it only kills one person rather than five. However, as before, there is a puzzle in that you do not think there is any general moral requirement to maximise happiness, which makes the reason why you think it is justified to kill the fat man to save five people unclear. Perhaps your response to the scenario below will help to clarify your thinking here.

The Saboteur (Scenario 3 of 4)

Okay so this scenario is identical to the preceding scenario but with one crucial difference. This time Marty Bakerman knows with absolute certainty that the fat man on the bridge is responsible for the failure the train’s brakes: upset by train fare increases, he sabotaged the brakes with the intention of causing an accident. As before, the only way to stop the train and save the lives of the five people already on the track is to push the fat saboteur off the bridge into the path of the train.

Should Marty push the fat saboteur onto the track (1 dead); or allow the train to continue (5 dead)?
I elected to shove fatty onto the track; he richly deserved to die, as he was a murderer. And, if this could save five innocent people, that is even better.

Whereupon, I was presented with this material:

Your belief that the right thing to do is to throw the saboteur off the bridge is not surprising given your previous response that it would be right to throw an ‘innocent’ (fat) man off the bridge if it had the effect of saving five people. We noted a tension in your earlier response in that it wasn’t clear why you thought it justified to kill the innocent man given that you do not think there is any general moral requirement to maximise happiness. However, this tension is less pronounced this time around, since presumably thoughts to do with culpability are part of the moral calculus in deciding whether it is justified to throw the saboteur off the bridge. It is possible that similar thoughts about culpability will be a part of how you think about the scenario below.

**The Fat Man and the Ticking Bomb (Scenario 4 of 4)**

The fat man, having avoided being thrown in front of the runaway train, has been arrested, and is now in police custody. He states that he has hidden a nuclear device in a major urban centre, which has been primed to explode in 24 hours time. The following things are true:

1. The bomb will explode in 24 hours time.
2. It will kill a million people if it explodes.
3. If bomb disposal experts get to the bomb before it explodes, there’s a chance it could be defused.
4. The fat man cannot be tricked into revealing the location of the bomb, nor is it possible to appeal to his better nature, nor is it possible to persuade him that he was wrong to plant the bomb in the first place.
5. If the fat man is tortured, then it is estimated there is a 75% chance that he will give up the bomb’s location.
6. If the fat man does not reveal the location, the bomb will explode, and a million people will die: there is no other way of finding out where the bomb is located.

Should the fat man be tortured in the hope that he will reveal the location of the nuclear device?
Yes, the fat man should be tortured

No, the fat man should not be tortured Please make a choice!

I opined that yes, the fat man should be tortured.

The reaction to my answers to this quiz was as follows:

*Should You Kill the Fat Man?* - Analysis 1

*A Matter of Consistency*

The first thing to note is that your consistency score is 100%. This is higher than the average score for this test (where higher is better), which is 78%.

It is often thought to be a good thing if one’s moral choices are governed by a small number of consistently applied moral principles. If this is not the case, then there is the worry that moral choices are essentially arbitrary - just a matter of intuition or making it up as you go along. Suppose, for example, you think it is justified to divert the train in the first scenario simply because it is the best way to maximise human happiness, but you do not think this justification applies in the case of the fat man on the bridge. The problem here is that unless you’re able to identify morally relevant differences between the two scenarios, then it isn’t clear what role the justification plays in the first case. Put simply, it seems that the justification is neither necessary nor sufficient for the moral judgement that it is right to divert the train.

You’ve done better than average in this test, but now is not the time to rest on your laurels, because let’s face it, most people don’t think very clearly about morality. However, before you embark on any further study(!) we suggest you check out the next page of analysis.

My reaction. Whoa, I am not sure that I like this even one tiny bit. If a non libertarian thinks I am logically consistent, perhaps I have erred somewhere along the line, for, as I say, there is a gigantic chasm between the thinking of those who favor, and oppose, the freedom philosophy.

The next response of the quizmaster is this:

*The Trolley Problem* - Analysis 2

The scenarios featured in this activity have been constructed to elicit contrasting intuitions about whether it is justified to end the life of one person in order to save the lives of some greater number of people.

Part of what is interesting here is what this tell us about consequentialist approaches to moral thinking. For example, straightforward utilitarianism, which holds that an act is morally right to the extent that it maximises the sum total of happiness of all the people affected by it (when compared to the other available
options), would seem to require an affirmative response to all the questions below. However, we know from previous research that such a consensus is unlikely. In particular, very few people tend to think that the fat man should be pushed off the bridge in order to save the lives of the five people stuck on the track. The fact that this option is so counterintuitive to so many people represents a significant challenge to straightforward utilitarian thinking.

I am not sure of what to make of all of this, but I report it, just to be thorough.

I opted to torture this fat pig of a man. This seems like a no brainer to me. Such an ogre deserves the most heinous punishment imaginable, and torture would appear to fit the bill. I resist the notion, however, that there is any “tension” in my answers. Yes, this is a reasonable position for a non-libertarian to take of an adherent of this position, but this is a two way street. I, too, see a “tension” in the mainstream view, not to say an utter contradiction, with the NAP.

The point of this exercise is not to cast aspersions on the quiz. It is well thought out, and interesting. Rather, my goal here is to establish that there is rather a large chasm between the thinking of non libertarian philosophers, who, I assume, concocted this quiz, and libertarians such as myself, who fit into this model as do round pegs into square holes or vice versa.

When asked about morality, I chose to call the heroic murderer “moral.” I could have easily gone the other way around, since I have no strong views on ethics or morals. (I only have an established perspective on what the law should be: to punish murderers, heroic or not, unless forgiven for their crimes). If I had indicated this, I suspect, my opinion would not have been characterized as 100% consistent by the non-libertarian creator of this quiz.

References


Block, Walter E. 2001B. “Jonah Goldberg and the Libertarian Axiom on Non-Aggression.”


Pufendorf, Samuel. 1673. Natural law and the law of nations (*De officio hominis et civis prout ipsi prae scriptur lege naturali*).


The Concept of Ethos: Aristotle and the Contemporary Ethical Debate

Irene Dal Poz
Università di Padova
Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata
dalpozirene@hotmail.it

ABSTRACT

The problem of the self has recently regained a wide interest in the philosophical panorama. The need to rethink the agent has then encouraged the rediscovery of Ancient Philosophy. Focusing on Aristotelian practical philosophy, this paper aims to demonstrate the intrinsic circularity existing between the agent and his context. To give evidence of this theory, it will be necessary to investigate the extensive concept of ethos, the meanings of which embrace both individual and collective dimension. Moreover, this itinerary through the Aristotelian thought will be the occasion for putting the rebirth of interest towards Ancient Philosophy into question.

KEYWORDS

Aristotle, ethos, moral agent, self

Introduction

“b”

In the last decades the contemporary ethical scenario, especially in analytic philosophy, has renewed the debate about the self and the agency. The shared dissatisfaction towards the ethical perspectives, as P. Donatelli observes, has led to the rediscovery and examination of the overshadowed dimension of the moral agent.

In this discussion, the Aristotelian voice has been raised by many scholars (belonging mainly to the Virtue Ethics current), who have identified the Greek philosopher as a relevant source for rethinking the moral agent and for putting the coordinates of contemporary moral philosophy into question. Given this premise and background, the paper will develop the research hypothesis that to fully understand the moral agent we have to consider him as deeply linked to his context. In order to throw some light on this issue, the essay will engage the cardinal concept of ethos in the Aristotelian practical philosophy. In particular, it

2 In this essay I will follow some simple and coherent rules for the transliteration from Greek to English. I will not be taking into account the differences between the vocals, such as epsilon and eta or omicron and omega; the two vocals’ couple will be transliterated with “e”, in the first case,
The Concept of Ethos: Aristotle and the Contemporary Ethical Debate

will attempt to analyze the relationship between its two essential meanings: individual ethos and collective ethos. While the first one states the agent’s character, the second one indicates the context, referring to shared values, habits and tradition. Collective ethos will then be named directly polis or context in the development of the paper, due to the plurality of the connotations involved.

However, if the general aim is to determine the relationship between the two meanings of ethos, then the specific themes of the paper are two crossed subsets. The first aims to show how for Aristotle moral agent and polis are conceivable as poles existing only in their relationship. Whereas, the second sought subset wants to prove whether this approach is effective in front of some specific contemporary practical challenges. The intersection between these two parts is played by a transitional passage that shows the reasons for the rebirth of the Aristotelian praxis in the contemporary scenario. Coherently with its purpose, the paper is organized into three main sections: (1) the reconstruction of contemporary ethical demands; (2) the Aristotelian answer to the relationship between the part and the whole and (3) the test of its legitimacy and limits regarding contemporary ethical experience.

Therefore, (1) the first part of the paper will focus on the contemporary ethical debate, by giving an account of what kind of contemporaneity is thought and why it has been chosen. In fact, due to the complexity and heterogeneity of the philosophical panorama, as previously mentioned the boundaries of the discussion will be circumscribed to a specific movement in the analytic philosophy, notably Virtue Ethics, and to some of its actors (such as A. Anscombe, A. MacIntyre, M. Slote). Thus, to satisfy this preliminary passage, it will be necessary to interrogate the reasons of Virtue Ethics and the main steps of the rediscovery of Aristotelian practical philosophy in this scenario.

Following the research project, (2) the second part will be shaped on the concept of ethos as it emerges in Aristotle’s practical writings. As a matter of fact, its characteristic plurivocity is the key lecture for showing the circularity between individual ethos and the collective one. This draws the attention to two cornerstone ideas: (a) the character preformation and (b) the practical agent’s active contribution. In order to elucidate them, it will be necessary to study the problem of the character education by looking at the concept of virtue. First, I will briefly introduce (a) the passive phase, by showing how the polis influences the individual character. Thanks to the concepts of induced virtue it will be possible to emphasize the process of metabolization of habits and values. Thereafter, I will demonstrate how Aristotle conceives (b) the way an agent may modify or influence his context. Indeed, although Aristotle believes in the priority of the context/polis, he does not

and “o”, in the second one. Moreover, the Greek vocal “u” will be transliterated with “y” (e.g. physis).
suppress a certain level of freedom for the agent. In this case, it will be necessary to examine the two meanings of practical reason in depth: the phronetical and the philosophical one. This will lead to the possible conclusion that the circularity between the individual ethos and the collective one cannot be defined as vicious, despite many scholars doing so.

To conclude, (3) the essay will focus on the pertinence of the Aristotelian perspective in the contemporary debate. Due to this goal, the conclusive part will problematize one issue: the impasse of collective ethos. The purpose will be to verify whether at present the concept of collective ethos has become too reductive or whether it is still functional. In other words, the attempt will be to understand whether the collective ethos still offers interpretative tools to approach the ethical experience and, if it does not, it will be necessary to explain why. In this case, the core issue will be to understand if we are bearing witness to an explosion of forms of life, that is to say to a pluralization of the contexts involved in our experience. However, as mentioned, it is necessary to begin with the reconstruction of the contemporary philosophical background.

1. The necessity of a new approach and the rebirth of Aristotelian practical philosophy

As the introductory title anticipates, this chapter aims at reconstructing a particular horizon in contemporary philosophy, Virtue Ethics, where Aristotelian practical philosophy has been the object of study and served as a model. This statement immediately raises difficulties, because it does not reckon with the many different trajectories that cross Virtue Ethics. In effect, when we talk of this movement, we are easily tempted to read it as if it were strictly coherent and rigorous. Contrarily, all the philosophers who took part in it have developed some essential peculiarities and they would deserve a specific analysis. Some of them have probably distanced their ideas so far from that they can no longer be assimilated. However, even if it is impossible (and most likely not even necessary for our purpose) to abridge all of them in a single scenario, it is unobjectionable that there are some relevant early agreements. In fact, although their thoughts have been described as described heterogeneous and irreducible parabolas to each other, they have primarily moved from some considerable affinities. My present goal will be to investigate them and their reasons.

To begin with, I would like to retrace the theoretical milieu that has encouraged the actualization of Aristotelian thought and of its cross concepts. The best way in order to do that is to understand its starting point. The analysis of two philosophers will support me in this operation: E. G. Anscome and M. Slote.

3 I am thinking mostly of MacIntyre A., Murdoch I., Anscombe G. E., Annas J. and Slote M.
Thanks to them, it will be possible to fully examine the ethical demand at the heart of the Aristotelian rebirth. By reading the Anscombe’s text, *Modern Moral Philosophy*, and Slote’s contribution, *Virtue Ethics*, we realize (a) that these authors raised their voices because of their dissatisfaction with contemporary moral philosophy and (b) that in Aristotle they found their favorite interlocutor because his philosophy offers a reliable alternative approach.

Let us start by taking a look at the analysis of the *dissatisfaction with the way moral philosophy has been done in modern times and, especially, recently* made by Slote. In particular, Slote’s claim is that the contemporary panorama is connoted by the struggle between two main ethical theories, on the one hand Kantian ethics or deontology and on the other Utilitarianism or Consequentialism. Slote remarks that neither of these approaches can give reason of the complexity of the ethical experience. As a consequence, according to Slote this shared unfitness legitimizes the assimilation of the two and, at the same time, promotes the revival of interest in ancient practical philosophy. Hence, the landscape of ethical debate has become re-articulated from two ways of thinking to three: Consequentialism and Kantian ethics, joined together, and Virtue Ethics, whose roots are generally in ancient philosophy and, especially, in Aristotle.

But Slote’s work also offers us a deeper key to understanding the interpretative battlefield, by explaining why Consequentialism and Kantian’s dyad are not able to completely satisfy the necessity of thinking the *praxis*. The reason lies in the dominant property that characterizes them: they are act-focused, in the manner that they look for *moral rules that are supposed to govern human actions*[^1]. According to Slote, this polarization of the action swallows up a fundamental part of morality, the one regarding the self and the moral agent. Conversely, Virtue Ethics is actually qualified for being agent-focused, that is to say that the *focus is on the virtuous individual and on those inner traits, dispositions, and motives that qualify him as being virtuous*[^2]. So, Virtue Ethics is born exactly because of the necessity of rethinking the importance of the self, of the agent and of his entire characteristics, in contrast with the abstraction and the focus on action of the normative theories[^3].

However, the distance between these two lines of thought is marked not only by their main-focuses (act or agent), but also by their different ways of settling in the reality. On the one side, the Kantian/Utilitarian approach privileges impersonality and objectivity, thanks to the support of an abstract and general parameter (whether it is the duty or the advantage is not important). Indeed, with this external principle they can guarantee an efficient and reliable way of

[^3]: *Ivi*, p. 177.
handling the experience. On the other side, Virtue Ethics demands a contextual or internal point of view. In this sense, the main problem is not the search for a criterion, but the analysis of how the agent is basically shaped by his form of life in order to become good. Now, the Aristotelian reference also becomes much clearer. As we will soon widely analyze, in Aristotelian practical philosophy, the character development necessarily takes place within the *polis*.

To sum up, Slote’s principal contribution came in terms of a contraposition: the act-focus approach is grounded on an external point of view, while the agent-focus approach is anchored on an internal one. Before continuing the analysis of the revival of practical philosophy it is important to bear in mind that the above-seen contrast answers to the necessity of clarifying and that a “pure” approach does not exist. For example, imagining an absolute Kantian approach is useful in a descriptive perspective, but it is not realistic. In other words, it does not matter whether the focal center is the concept of duty, one of advantage or one of virtue: in the moral approach there is always an amalgam of different considerations working together.

Anyway, in these terms the importance of an internal point of view emerges for the first time with Anscombe. In her article, *Modern Moral Philosophy*, she denounces the stalemate of contemporary philosophy and she attacks contemporary normative philosophies. The reason of Anscombe’s account lies in their detachment from their background, where they were still valid. Talking about the contemporary approaches, she says that they are survivals, or derivatives from survivals, from an earlier conception of ethics, which no longer generally survive, and are only harmful without it. In this sense, the necessity of rethinking morality is born because of the disconnection between moral concepts and the consequent form of life, which concretely gives sense to them. Again, the answer is the rehabilitation of the Aristotelian conception of ethics and, in agreement with this different paradigm, the focus shifts from the duty and the action to the virtue and the agent’s character.

But, Anscombe’s analysis, based on the importance of the form of life, also highlights another latent root of Virtue Ethics: Wittgenstein and, especially, his late writing, *Philosophical Investigations*. This clarification reveals that the perception of dissatisfaction, which moves Virtue Ethics, is not a unicum and it is also shared with other authors, who find their roots in Wittgenstein’s works and in his fundamental idea of form of life. If on the one side there are Anscombe,

---


9 See Donatelli P. in *Reshaping Ethics after Wittgenstein*, in *Wittgenstein-Studien*, Band 4/2013: “I’d like to mention here the views held by authors such as Anscombe, Murdoch, Cavell and Diamond. They criticize, in very different ways, a significant line in modern and contemporary moral philosophy, which has depicted moral thought as the operation of a few concepts cut off from the life form where they have place”, p. 224. Moreover, see Donatelli P. in the introduction
MacIntyre and other philosophers, who belong to Virtue Ethics and look at Aristotle as a model, then on the other there are some, such as S. Cavell and C. Diamond, who present themselves as a reliable alternative to Virtue Ethics. Although they share the same need to rethink ethics (and, above all the category of the self) beyond the limits of the moral ought conception and they do it by moving from the same claim of the form of life, the similar aim is pursued from different perspectives. In fact, the shared importance of the conceptual context as backdrop of moral life leaves room to a considerable distance, caused by the idea of “constitutive nature”. According to Anscombe’s account, virtues have to be inferred through the discovery of “natural” characteristics, that define what is truly human. This renewed naturalistic hypothesis supports then the assumption that the moral concepts of the Virtue Ethics are still valid, since they are stably based on human nature, while the morality of ought is temporarily attached to a particular form of life. The reference to a constitutive human nature is also the reason why Anscombe looks for a paradigm in ancient philosophy (e.g. in Aristotelian Practical Philosophy\textsuperscript{10}), while Diamond and Cavell refute this option. Indeed, they have attempted to overcome the boundaries of a morality reduced to the ought not by referring to some proper human functions, but by focusing on the historical and dynamic density of moral concepts.

As a result, given this general framework, the real problem is not whether moral philosophy should be rethought or not, but how to perform this operation, in what direction and with which basis. It is for this reason that testing the efficacy of Virtue Ethics becomes even more important as well as remembering that the same philosophical necessity and root have produced at least two alternatives.

To recall the cardinal points, Virtue Ethics is a moral approach that can be interpreted as a laboratory for rethinking the praxis and it is qualified by two traits, inherited from the Aristotelian perspective and integrated with some contemporary philosophy contributions: (1) the problem of the self and of the virtuous character and (2) the inalienable contextual dimension. These two elements are inextricably shaped together. So, what matters above all, even more than their single exams, is always their relationship.

But now, after having rebuilt the contemporary philosophical scenario and having outlined the main features of Virtue Ethics, it is time to inquire directly practical philosophy and to analyze Aristotle’s account of the relationship between the agent and the context.

\textsuperscript{10} The issue of what kind of “naturalism” we can talk about in Aristotelian Practical Philosophy will be problematized in the following pages.
2. Aristotle and the different meanings of ethos

a. The character preformation: from the collective ethos to the individual one

As stated above, the scope of this second section is the study of the interactions between the different meanings of ethos in Aristotle. The supposition is that through a path in practical philosophy their necessary co-implication will be demonstrated.

However, preliminarily it is essential to make some methodological remarks. In fact, though the main book reference is surely Nicomachean Ethics, if we consider Aristotelian practical philosophy as a prism whose faces show different perspective of the same figure, it becomes evident that we also have to include the analysis developed in other practical works, such as Eudemian Ethics and Politics. As a matter of fact, all these works have a common theme, which is the praxis, or better the eu-praxia, approached from different points of view and priorities. Moreover, it is always because of the continuity of contents, that it is impossible to elude the centrality of another treatise, De Anima, whose topic is the cardinal concept of the soul (psyche). In fact, a proper study of the eudaimonistic goal needs to broaden the horizon to the Aristotelian psychology and also to the metaphysics, since they both provide the coordinates for the analysis\(^\text{11}\). Always following the figurative representation, we could imagine a Cartesian plane whose axes are respectively Aristotelian metaphysics and psychology and whose point of intersection coincides with the prism of practical philosophy with all its sides. Thanks to this synoptic perspective, it will become clear that when we talk about a single concept we are at the same time engaging all the connected concepts (e.g. the concept of ethos is necessarily linked with those of psyche, of physis, of polis and of phronesis). Every concept is completely integrated in a network, giving a meaning to it. This principle is essential because it will play the role of evaluative parameter in the last section of the paper, being aimed at testing the pertinence of this point of view in contemporary ethics. Indeed, the result of this coordination of concepts is that they only make sense when kept together, therefore they may

---

\(^{11}\) Irwin T. H. proceeds exactly in this direction, in The metaphysical and psychological basis of Aristotle’s ethics, in Essays on Aristotle’s Ethics, edited by Rorty A. O., University of California Press, Berkeley and Los Angeles, 1980. Irwin wants to prove the point that sectorial and self-sufficient interpretation of the Aristotelian’s writings is unsatisfactory, while it is more productive to imagine them as intersections of a more complex texture. Moreover, it is important to remind that Aristotelian psychology is far from an anthropological reduction, since the concept of psyche embraces all the living things. About the metaphysical coordinates of practical philosophy, the last Book is paradigmatic, the chapters from VI to VIII and the problem of the “human condition”.

258
be unacceptable singularly. If there is a significant critical point of the present analysis, it is probably caused by this necessary co-implication of concepts.

But the above-outlined synoptic principle is the reason why if we want to engage in the issue of the reciprocal influences among the two *ethos* forms, we have to also address our interest to the virtue theory. This expression refers to the process of learning, training and practice of the *political*12 virtue. This clarification of the *political* dimension is necessary since Aristotle makes a distinction between the “natural” virtue (*physike arete*) and the virtue “in the full sense” (*kyria*)13. Following the philosopher’s remarks, we shall turn our attention to the second meaning, which from now on will simply be called virtue. The reason behind the *minimization* of the natural roots of the virtue is that Aristotle refuses to qualify his ethics as a simple form of naturalism14. That means that his ethical perspective does not imply a natural exercise of *dynameis*, capacities, but a *canalization* of them towards *ton agathon* (the good). Nevertheless, although the agent cannot be naturally moral, since ethics needs a detachment from simple naturalness, the interpretation of character as a “second nature” is sustainable. This is only one of the paradoxes of practical philosophy that will be dissolved in the next pages, like the one concerning the new form of automatism generated by the acquisition of this second nature. Barely outlined now, these themes will be widened below.

---

12 Here and in all the paper *political* will be used in an etymological sense; so, it is directly connected with the dimension of the *polis*, not “simply” with the political horizon, as we are used to believe now.

13 The unique definition of character we have is in *Eudemian Ethics*, Inwood B. and Woolf Raphael (eds.), Cambridge University Press, 2013 (Book II, 1220 b 1) and is imbalanced towards the natural virtue. However, in *Nicomachean Ethics* we witness an understatement of the natural virtue importance and a strengthening of the habituation process. The reconstruction of the confutation of the natural virtue in *Nicomachean Ethics* and its reasons are exposed in *La catena delle cause, determinismo e antideterminismo nel pensiero antico e contemporaneo*, edited by Natali C. and Maso S., Hakkert, Amsterdam, 2005. Another fundamental reference is *Abitudine e saggezza. Aristotele dall’Etica Eudemia all’Etica Nicomachea*, Donini P., Edizioni dell’Orso, Alessandria, 2014. In this book Donini explains the differences between the *Eudemian Ethics* and the *Nicomachean Ethics* analyzing their chronology, and proving the anteriority of the *Eudemian*. In fact, while in the *Eudemian Ethics* Aristotle emphasizes the natural foundation of virtue, in the *Nicomachean Ethics* he focuses on the metabolization of habits. Donini then retraces the cause of this turn and of this deeper analysis of the habituation process on the influence played by Plato’s *Laws* on Aristotle.

14 *Nicomachean Ethics*, translated with Commentaries and Glossary by Hippocrates G. Apostle, D. Reidel Publishing Company, 1975, Book II, 1103 a 15 ff.: “an ethical virtue is acquired by habituation (*ethos*), as it indicated by the name ‘ethical’, which varies slightly from the name ‘ethos’. From this fact it is also clear that none of the ethical virtues arises in us by nature [at birth], for no thing which exists by nature can be changes into something else by habituation (…) Hence virtues arise in us neither by nature nor contrary to nature; but by our nature we can receive them and perfect them by habituation”.
However, what exactly is the thread that connects the relationship between individual ethos and collective ethos with virtue in practical philosophy? To understand it, let me turn to the first cornerstone idea: the preformation of character or the habit-forming process. With these two equivalent expressions I mean the way collective ethos shape and educate the practical agent. This is synthesizable by the concept of induced virtue, which is functional to this analysis for two reasons: (a) it creates an initial hierarchy between the forms of ethos in favor of the collective one and (b) it demands us to give reasons on how the virtue is induced.

First, we assumed that from Aristotle’s account the polis is responsible for character habituation through paideia, education. This point of view is clearly supported by the continuity existing between Nicomachean Ethics and Politics. As it emerges from Nicomachen Ethics, Book I, the goal of politics is the highest good and politics takes the greatest care in making the citizens of a certain quality, i.e., good and disposed to noble actions (praktikoi ton kalon)¹⁵. This institutes a direct correspondence between the polis and the possibilities for the agent of practicing the human end, which is the eudaimonia¹⁷. The programmatic declaration of intent recurs not only at the beginning of the logoi¹⁸, in Book I and II, but also at the end, in Book X. In this sense, the Aristotelian logoi create circularity, where the end of the course confirms the beginning. What matters is that in these passages Aristotle explains why virtuous attitude can grow only at the political level and introduce the problem of how the polis generates this metabolization. Let us try to summarize the Aristotelian argumentation, with references to the aforementioned books:

(1) The objective of practical philosophy is the highest good, the human flourishing, which is defined as a virtuous activity¹⁹. This definition shifts the problem from the highest good to the assimilation of virtues;

¹⁶ The purpose of these logoi is as well the action: “since the end of such discussions is not knowledge but action”, Nicomachean Ethics, Book I, 1095 a 5-6.
¹⁹ “Then the good for a man turns out to be an activity of the soul according to virtue, and if the virtues are many, then according to the best and most complete virtue” The Nicomachean Ethics, Book I, 1098 a 16 – 18.
The virtue is a habit (hexis) and it is acquired by acting in a virtuous way\textsuperscript{20}, so it is crucial to acquire the right habit as soon as possible\textsuperscript{21}. This is the scope of the polis;

the polis promotes the virtuous habits with the good forms of government, constitutions and the coercive force of laws\textsuperscript{22}. In order to do that, the politician has to study the human soul, because by ‘human virtue’ we mean not that of the body but that of the soul, for it is of the soul, too, that happiness is stated by us to be an activity. If such be the case, it is clear that a statesman should understand in some way the attributes of the soul\textsuperscript{23}.

The importance assigned by Aristotle to the polis and to its legislative process, connected with the educational purpose, is exemplified by the rôle éminemment éducatif qui la fin de l’Ethique à Nicomaque fait jouer à la loi politique instituée par le nomothète\textsuperscript{24}. Indeed, as R. Bodéüs demonstrated\textsuperscript{25}, Nicomachean Ethics is a

\textsuperscript{20} Ivi, Book II, 1105 b 20 ff., paragraph IV. Here Aristotle qualifies virtue as a hexis and shows the difference between habits, powers and feelings.

\textsuperscript{21} Ivi, Book II, “For it is by making citizens acquire certain habits that legislators make them good, and this is what every legislator wishes, but legislators who do not do this well are making a mistake; and good government differs from bad government in this respect. Again, it is from the same actions and because of the same actions that every virtue comes into being or is destroyed, and similarly with every art (...) In short, it is by similar activities that habits are developed in men; and in view of this, the activities in which men are engaged should be of [the right] quality, for the kinds of habits which develop follow the corresponding differences in those activities. So in acquiring a habit it makes no small difference whether we are acting in one way or in the contrary way right from our early youth; it makes a great difference, or rather all the difference” 1103 b 3 ff.

\textsuperscript{22} Ivi, Book II, “For it is by making citizens acquire certain habits that legislators make them good, and this is what every legislator wishes, but legislators who do not do this well are making a mistake; and good government differs from bad government in this respect. Again, it is from the same actions and because of the same actions that every virtue comes into being or is destroyed, and similarly with every art (...) In short, it is by similar activities that habits are developed in men; and in view of this, the activities in which men are engaged should be of [the right] quality, for the kinds of habits which develop follow the corresponding differences in those activities. So in acquiring a habit it makes no small difference whether we are acting in one way or in the contrary way right from our early youth; it makes a great difference, or rather all the difference” 1103 b 3 ff.

\textsuperscript{23} Nicomachean Ethics, Book I, 1102 a 9 ff.

\textsuperscript{24} Rodrigo P., Aristote et les choses humaines, avec une appendice sur la politique stoïcienne, préface Aubenque P., Ousia, Bruxelles, 1998, p. 37. The last lines of the Nicomachean Ethics completely support this interpretation and directly open to the Politics: “since our predecessors left the subject of lawing without scrutiny, perhaps it is better if we make a greater effort to examine it, and especially the subject concerning constitution in general, so that we may complete as best as we can the philosophy concerning human affairs. First, then, let us try to go over those parts which have been stated well by our predecessors, then form the constitutions we have collected let us investigate what kinds of things tend to preserve or destroy the states or each of

261
course whose conceived audience is mainly made by scholars and politicians, namely the nomothetês. Thus, the emphasis Aristotle puts on these figures and on their particular functions allows us to distinguish two specific correlated aims of the *Nicomachean Ethics*: the political purpose, or the *education of the educators*\(^{26}\), and the spiritual one, or the *transformation of the self*.

Now, after having summarized the Aristotelian argument, we should attempt to grasp more precisely how this metabolization happens. Even though we said that the political education is the preferential channel in order to acquire virtues, we did not specify how it works, while Aristotle reconstructs this passage. The philosopher’s account is that we acquire habits mainly thanks to an emotional education. This implies that above all in the earlier stages\(^{27}\) the good man’s development seems to devalue the cognitive and rational dimension and to appeal mainly to the emotional one. As a consequence, this draws the attention to the role of emotions (or passions, which both translate the Greek word *pathos*) in the habituation process and to the problem of being properly affected. Aristotle’s claim is that:

Thus ethical virtue is concerned with pleasure and pains: for we do what is bad for the sake pleasure, and we abstain from doing what is noble because of the pain. In view of this, we should be brought up from our early youth in such a way as to enjoy and be pained by the things we should, as Plato says, for this is the right education. Again, since virtues are concerned with *actions* and passions, and since every *action* and every passion is accompanied by pleasure or pain, then for this *reason*, too, virtues would be concerned with pleasures and pains. (...) We assume, then, that such virtue is concerned with pleasures and pains and disposes us to do what is best, while vice disposes us to do the contrary.\(^{28}\)

But what is the connection between virtues/vices, desire, passions and pleasure/pain? In fact, what actually characterizes and qualifies the Aristotelian

---


\(^{27}\) About the earlier stages of education, it is paradigmatic the study of Burnyeat M. *Ein Aristotle on Learning to Be Good*, in *Essays on Aristotle’s Ethics*, pp. 69-92. “A wide range of desires and feelings are shaping patterns of motivation and response in a person well before he comes to a reasoned outlook on his life as a whole, and certainly before he integrates the reflective consciousness with his actual behavior”, p. 70.

\(^{28}\) *Nicomachean Ethics*, Book II, 1104 b ff.
account is the effort made to give reason of this dynamic rigorously and to justify it coherently with the other parts of his work. So, first, I will briefly demonstrate the connection between virtues/vices, passions and pleasure/pain, omitting for the moment how desire takes part in this process. From this point of view, we know that in the Aristotelian theory virtue is a *hexis* (habit or disposition). We also know that the philosopher describes habits as *those qualities in virtue of which we are well or badly disposed with reference to the corresponding feelings*\(^{29}\). Thanks to this definition it is possible to connect and, at the same time, maintain distinct virtues and feelings. Moreover, this leads us to shift the focus from virtue to feelings (or passions), identified as *whatever is accompanied by pleasure or pain*\(^{30}\). So, if the *polis* wants to promote an effective educational aim, it has to recognize pleasure and pain as the load-bearing axes in the early habituation stage.

However, we are still missing the conjunction among feelings and behavior, which I will now elucidate thanks to the concept of desire. For this reason, in order to completely understand this dynamic we have to also include Aristotelian psychology, especially the passage exposed in *De Anima*, Book III, chapter 7. Here the philosopher institutes a correspondence between feelings (or better, pleasure and pain) and desire (*orexis*)\(^{31}\). The sensitive faculty and the appetitive one entail each other, since when we feel pleasure we pursue and when we feel pain we avoid. So, though the sensitive function and the appetitive one are logically and physically different, they strictly cooperate together.

This rapid reconstruction of the intersections between all these concepts says something more about the Aristotelian action theory\(^{32}\), especially about the role of desire. As Aristotle asserts, appetency (*orexis*) is the major cause of locomotion:

> Both these, then, are causes of locomotion, intelligence and appetency. By intelligence we mean that which calculates the mean to an end, that is, the practical intellect, which differs from the speculative intellect by the end at which it aims. Appetency, too, is directed to some end in every case: for that which is the end of desire is the starting point of the practical intellect, and the last stage in this process of thought is the starting point of action. Hence there is good reason for the view that these two are the causes of motion, appetency and practical thought. For it is the object of appetency, which causes motion; and the reason why, thought causes motion is that the object of appetency is the starting point of thought. (...) But, as a matter of fact, intellect is not

---


\(^{30}\) *Ivi*, Book II, 1105 b 23.


\(^{32}\) Because of the lack of space it will not be possible here to analyze the criterions of responsibility (or better, the voluntary actions/virtues), that is one of the more important contributions of the Aristotelian action theory. However, as is known, the text reference is Book III, *Nicomachean Ethics*. 

263
found to cause motion apart from appetency. For rational wish is appetency; and, when anyone is moved in accordance with reason, he is also moved according to rational wish. But appetency may move a man in opposition to reason, for concupiscence is a species of appetency.\footnote{De Anima, III, 10, 433 a 13 ff.}

As this quotation underlines, appetency has a greater influence than practical thought. Consequently if desire is the cause of actions (praxeis), it is as well the cause of virtues, because of the connection existing between actions and virtue. To cast some light on this circularity, we acquire virtues through the reiteration of actions, and the acquired virtue guarantees some stability to the agent’s behavior and some reliability to his character (that is to say, to his future actions).

Nonetheless, we should make two considerations about the habituation process: (1) the rediscovery of the role of desire as a revolution, not as a tyranny and (2) the “natural” boundaries of the induced virtue. Focusing on these two considerations is also the turning point from the first part to the second part of the section dedicated to Aristotelian practical philosophy.

First, because of the extraordinary attention directed to the key role of appetency in action we could talk of a desire revolution. In contrast with an intellectualistic prejudice, Aristotle valorizes the primacy of desire over practical reason. A further evidence of this inversion of importance between practical reason and desire is the case of akrasia, known also as the Medea principle\footnote{Definition given by Davidson D. in Paradoxes of irrationality, in Philosophical essays on Freud, Cambridge Editions Wollheim D. R. et Hopkins J., 1982, p. 294.}. Indeed, the incontinence is described as a conflict engaging what we know we should do (which means the cognitive perception of a specific situation) and what we would like to do (which means the influence of passions in that context) resolved by the priority of desire\footnote{This characterization of akrasia corresponds to the strict incontinent action, described by Mele A. R. in Irrationality: An essay on Akrasia, self-deception and self-control, Oxford University Press, New York, 1987. The strict incontinent action is defined as an “incontinent action against a consciously held better judgment about something to be done here and now”, p.7. According to Mele, this is the typical kind of incontinence studied and its features are that it is free and intentional.}. In contrast to Plato’s account\footnote{The reference is to the Platonic writing, Protagoras, and, in particular, to the passages 351 a – 358 d.}, Aristotle conceives the hiatus between reason and desire as plausible and believes that the strength of the first is not sufficient to assert itself. The reason is that even if the intellect issues the order and the understanding bids us avoid or pursue something, still we are not thereby moved to act: on the contrary, action is determined by desire; in the case, for instance, of the incontinent man.\footnote{De Anima, Book III, 433 a.} In a hypothetical scale showing the weight of the action components, the emotional side would clearly turn out to be

\begin{itemize}
  \item \textbf{Desire:} 70%
  \item \textbf{Reason:} 30%
\end{itemize}
predominant. However, it may be objected that the question concerning the action and the hierarchy between the rational sphere and the emotional one is actually far more complicated than that claimed by the “desire revolution”. The main risk is that the deconstruction and the invalidation of a picture of human nature based on a rational prejudice gives room to a new tyranny, one of feelings and desire. The phenomenon of weakness of willing always testifies the danger involved in this perspective, since it shows the consequences of being guided by appetency in contrast with reason or, generally speaking, the danger of a conflict between the plural components of actions. Indeed, when the akратic man is faced with a choice between two chains of actions, he experiences an internal conflict caused by these different tensions, and because of the lack of harmony he can not behave virtuously. But this remark leads us to the analysis of the practical reasons.

b. The double meaning of practical reason: from individual ethos to the collective one

Therefore, the importance given to the akратic phenomenon is a proof of the Aristotelian awareness of the risks hidden behind this revolution. Consequently, the hypothesis is that in practical philosophy we can talk of a morally virtuous action only in presence of the conjunction between intellect and desire. In absence of a reevaluation of the role of intellect, virtue would be an exclusive prerogative of desire and Aristotle could be assimilated to Hume and the emotivists. For this reason, it is possible to read many passages from Nicomachean Ethics to try to throw light on the dangers of this new form of extremism and to mitigate it. In particular, Aristotle makes some interesting remarks about the necessity of harmonizing the distance between reason and desire when he introduces the crucial problem of the principle of action, decision (proairesis) in central Book VI. On this occasion, Aristotle affirms that:

now virtue of character is a state that decides; and decision is a deliberative desire. If, then, the decision is excellent, the reason must be true and the desire correct, so that what reason asserts is that desire pursues. (...) But the function of what thinks about action is truth agreeing with correct desire.


This definition ties the acquisition and the practice of virtue to the agreement and the proportion between reason and desire: *what reason asserts is that desire pursues*. But Aristotle goes further. Not only is decision the source of motion and the above-outlined dyad the source of choice, but also *decision is either understanding combined with desire or desire combined with thought, and this is the sort of principle that a human being is*[^10]. As a result, the necessity of an agreement is directly required by the human composite soul and all the other states are above or below the human condition. Concerning the conditions beyond the truly human one, the beginning of Book VII is paradigmatic, since it offers a complete overview of the human possibilities[^11]. Indeed, here Aristotle talks about six kinds of states. If we exclude virtues and vices, which are both a *hexis* requiring the explained agreement, four different states are left. Our attention is drawn to two of them, *brutality* and the *divine or heroic virtue*[^42], because they represent the negative and the positive extremes. Otherwise stated, *brutality* and the *divine virtue* identify the possible developmental directions beyond *ta anthropina*, the human affairs. On the other hand, the remaining two are the already-seen *akrasia*, incontinence, and its opposite, continence. Whereas the first couple synthesizes the extreme options, the last one embodies the condition of *oi polloi*, the many. In fact, Aristotle states that the virtuous man is one *who has harmonious thoughts, who desires the same things with respect to every part of the soul*[^43], while the majority of men are in conflict with themselves, as the incontinent case proves[^44]. So, not only is Aristotle well aware of the complexity of the agreement between reason and desire, but he is also conscious that the virtuous condition concretely concerns only a minority in the *polis*.

The intrinsic difficulty of acting virtuously legitimates the controversial characterization of *semi-utopian*. On the one hand, it is a form of *utopia*, because it requires the correct education and the correct complete moral development; in this sense, only a few people can actually promote this kind of *care of self*. On the other hand, it is *semi-utopian* since Aristotle still leaves room for the *perfectionism* of the self, both in political and philosophical dimensions. In fact, the recognition of an unattained higher condition and the consequent work on oneself to realize it (or, at least, to get close) are the warranty of a continuous improvement towards the moral development. Thus, the self-realization is a concrete possibility, even though extremely difficult. Moreover, if until now the doubts affected mostly the

[^42]: If we want to clearly identify divine life or virtue, we should think of Book X of *Nicomachean Ethics*, chapter VI – VIII, where Aristotle introduces the *bios theoretikos* and the human tension toward divine life.
[^44]: *Ivi*, 1166 b 7 – 8.
agent and the practicability of the virtuous attitude, because of the complexity of intertwined elements, it is important to note that the doubts also touch the correlated concept, the *polis*. As a matter of fact, Aristotle is well aware that there could be *polis* whose constitutions (*politeia*) do not incentivize the moral development, or, what is worse, they concretely do not allow it. From this point of view, the Aristotelian account rather than being descriptive of a concrete condition, it is indicative of a tension towards the realization of the intrinsically connected aims, as well as the political and the individual. In this sense, Aristotelian practical philosophy is not naively optimistic or anchored in an irenic imagination of reality. But the confutation of a simplified interpretation of practical philosophy could be clarified through a deeper analysis of the concept of practical intellect and of its role.

The renewed importance of reason clearly unveils itself if we consider the formalization of the chain of reasoning, the practical syllogism\(^{45}\). First of all, in order to best capitalize this tool, we should briefly recall the structure of a practical syllogism, that could be schematized like that:

\[
\text{MaP} X - Y \\
\text{MiP} Z - X \\
\text{Con} Z - Y^{46}
\]

Moreover, an example will make this schema more intelligible:

MaP: To not get wet when it rains (X) it is important to be in good health (Y); MiP: To carry an umbrella (Z) is a way not to get wet when it rains (X); Con: You should carry an umbrella when it rains (Z) to be in good health (Y).

What is relevant in our analysis is the exam of the different functions played by each concept and their interactions with each other. Concerning the major term (Y), thanks to the analysis in *Eudemian Ethics*, we know that *no one decides on an end, but rather on what contributes to the end* (...) *for example no one decides to be healthy* (...) *nor to be happy*\(^{47}\). So, according to Aristotle, the goal is not the result

\[^{45}\text{A careful examination of Aristotle’s concept of practical syllogism is offered by Natali C., *La saggezza di Aristotele*, Bibliopolis, Napoli, 1989, especially chapter IV.}\]

\[^{46}\text{Where MaP identifies the major premise, MiP indicates minor premise and Con identifies the conclusion, the syllogism’s outcome, that is an action. Moreover, we have to distinguish:}\]

a) The *major term* (Y), which indicates the end (*telos*) that in the practical syllogism is *ton agathon/eudaimonia*;

b) the *minor term* (Z), which is the efficient cause whose end is the major term;

c) the *middle term* (X), which is a specification of the general goal relative to a particular *kairos* and connects the other two terms.

\[^{47}\text{*Eudemian Ethics*, Book II, 1226 a 6 – 10.}\]
of reasoning or choice. But the philosopher goes further and shows that the end of the MaP is composed by the couple of opinion (doxa) and desire (orexis), which means the recognition of a good habit linked to the above-seen dynamic of pleasure and pain. Thus, the end of the practical syllogism is the result of metabolization and ethismos, as induced virtue explained. However, because of its indeterminacy the MaP is inactive (e.g. human flourishing and to be in good health are still abstract concepts). Due to this passiveness and inactivity, the middle term (not to get wet) plays a key role in concretely re-determining the end (to be in good health) in the particular situation (when it rains) and, by doing that, in connecting the MaP with the conclusion. But this passage is still not enough, since we are missing what the agent actually does in order to realize the specification of the general end. And that is where the minor term (to carry an umbrella) intervenes. As it emerges from the scheme, the minor term is actually the course of action deliberated for achieving the general goal, after having specified it. However, the cooperation between middle and minor term is still unclear. That is why we have to focus on the MiP, since it is the place where they are joint together. Again, it is only through the balance of reason and desire that the agent can make choices, since the decision is a deliberational desire for things that depend on us (in fact, e.g. the weather does not depend on us). The latter implies the transmission of desire, conveyed from Y to X, the specification of the general agathon, and the role of practical wisdom, which allows us to identify the best action model. Finally, we arrive at the study of the concept of practical reason, phronesis, the function of which has always engaged Aristotelian scholars. They are mainly polarized into two perspectives: one that considers phronesis as the search (zetesis) for the ways to reach the goal, whereas the other one sees phronesis as directly involved with the definition of the goal. However, E. Berti has offered a different interpretation, which goes beyond the polarization described above. According to Berti we should read practical wisdom as being differently involved in both. As a matter of fact, phronesis firstly has to recognize the goal.

48 Ivi, Book II, paragraph X.
49 In the Eudemian Ethics, Aristotle directly recognizes the polis as the responsible of this habituation: “hence the good itself would be this: the goal of all that is achievable by human action. This is what falls under the science that has authority over all sciences; this science is politics and household management and wisdom”, Book I, 1218 b 13 ff. By ethismos we mean the process of social transmission of behavioral patterns, through rewards and punishments, as Natali defined it in La saggezza di Aristotele, p. 163.
50 Nicomachean Ethics, Book III, 1113 a 12.
51 In fact, Bodéüs has coined the expression l'irritante querelle des exégèts to talk about the debate that has gone through the XX century about the role of phronesis in practical philosophy.
52 The two interpretative lines respectively are well represented by Aubenque P. and Gauthier R.
and only after the acknowledgement of *ton agathon* (which is the result of the habituation process) is it possible for one to decide on how to achieve it. Thus, this necessary recognition is what distinguishes *phronesis* from cleverness (*deinotes*)\(^{54}\), that is conceived as a form of instrumental reason that calculates the ways to achieve any goal (no matter if right or wrong). In this sense, even though practical wisdom does not deliberate about the *telos*\(^{55}\) nor does it call it into question, its function is not simply the choice of the best chain of actions addressed to any end. Its fundamental premise is *ton agathon*, the good end (to recall the previous example, the first step is the acknowledgement of *health* (Y) as a good to be pursued).

These initial considerations about *phronesis* lead us to the second point, called the “natural” boundaries of the induced virtue. Indeed, it is important to remind ourselves that the *induced virtue* defines an “embryonic moment” that should be overcome. If it is true that political education is fundamental, since it represents the first stage in moral development, it is also true that habituation is not sufficient to make men virtuous, when this also implies autonomy, in the sense of being able to self-exercise practical reason. In other words, two elements work in synergy for *eupraxia*: intellect and character\(^{56}\). If character at the beginning is the outcome of the *induced virtue* and of the *polis’* work, intellect, on the other hand, is always particular and requires the active contribution of the agent. And that is exactly the role of *phronesis*, since it is the intellectual virtue letting the agent deliberate about the possible course of actions.

But, allow me to make some remarks about *phronesis* and its personification, the *phronimos*. We could start from the incarnation of practical reason. The *phronimos* is an agent who shares the common opinion of the highest good, thanks to habituation, and who knows how to achieve it in the situation, thanks to the exercise of *phronesis*. The constant combination of these two features makes the *phronimos* a behavioral model, since *the measure of each thing is virtue or a good man as such*\(^{57}\). But this figure is problematic. Or better, it is, as we remarked for the aims of Aristotelian practical philosophy, a *semi-utopia*. In this case, its main problem is the unity of virtue: as J. Annas said, if we acquired a *hexis*, we should

---

\(^{54}\) *Nicomachean Ethics*, Book VI, 1144 a 24 -29.

\(^{55}\) And it would not be possible, since we cannot deliberate about things that could not be otherwise, as Aristotle states in *Nicomachean Ethics*, Book VI, 1141 b 8 ff.: “prudence, on the other hand, is concerned with things which are human and objects of deliberation; for we maintain that the function of a prudent man is especially this, to deliberate well, and no one deliberates about invariable things or about things not having an end which is a good attainable by action”.

\(^{56}\) *Nicomachean Ethics*, Book VI, 1139 a 31 – 1139 b 6.

\(^{57}\) *Ivi*, Book X, 1176 a 16 – 17. The distance between the virtuous man, the excellent, and *oi polloi* is stated in Book III, 1113 a 25 – 1113 b 2.
be able to acquire all, since the link is the exercise of phronesis\textsuperscript{58}. Thus, according to Annas’ analysis, the complete wise-man, in possession of all the virtues, is a normative ideal, while we could talk of the concrete phronimos as a continuous “approximation”, getting closer and closer to the ideal paradigm.

The second consideration directly aims at the characteristics of practical intellect. We could explain it by introducing two metaphorical pictures, often used among Aristotelian scholars in order to better represent the moral developmental dynamic: the athlete and the student\textsuperscript{59}. In both cases, there is a fundamental initial training, where the athlete and the student are taught the skills that they will be able to practice, and then a moment of autonomous application. For instance, when you learn to play the piano, first you have to learn the fundamentals and the technique. Only at a later date you will be able to play more difficult compositions or to extemporize. In this case, assimilation goes through the repetition and recognition of the developed skills. That is why you will be able to use it and you will be persuaded of their goodness. The combination of repetition and recognition moreover generates a form of automatism, which allows the definition of virtues, and more generally of character, as “second nature”. To understand this statement, the example of the athlete is perfect: as a matter of fact, the spontaneous and graceful movement made by athletes looks natural, and one could almost think that no effort is required. So, the result of the training is that the athlete or the student does not have to concentrate on what to do or how to do it all the time, thanks to the long repetition of the movement and to the acknowledgement of its reasons. As a consequence, the metabolization makes the thought and the resulting action immediate and natural. But the automatism, induced by lifelong repetition, generates the action naturally, so only a new metabolization can modify it. The concept of second nature should be now clearer: a developed character which is spontaneous and intuitive, but not impulsive. If we think of the practical syllogism, it means that the cardinal passage of the transmission of desire and of the practice of phronesis (MiP) becomes an automatism.

All these remarks about the chain of practical reasoning and the role of phronesis provoke two related questions:

1) How many chains of actions does phronesis really consider and practice?

\textsuperscript{59} See Annas J., \textit{The intelligent virtue}, Oxford University Press, Oxford, 2011. Moreover, it has to be recognized that one of these figures is primarily Aristotelian, since the philosopher talks about the athlete in order to explain the existing connection between activities and character in \textit{Nicomachean Ethics}, Book III, 1114 a: “for it is particular activities which produce men of a certain kind. This is clear in the case of those who train themselves for any contest or action; for they are constantly active”. In addition, the student, as it has been said at the beginning, is the ideal audience of these logoi.
Does it really leave room for different chains of actions or do we follow some already fixed behavioral patterns?

2) Does Aristotle imagine any form of criticism of the politically accepted end? If yes, what is it and how does it work, since we know that by the end practical wisdom, although concerned, does not deliberate upon it?

The first question could be posed in two different ways: on the one hand it concerns the reliability of character, on the other hand it implies a form of character determinism. The two interpretations differ from the weight of the automatism seen above. While the first reads the mechanism as an assurance of character stability, the second one interprets it like a “prison of the character”60. The supporters of the first interpretation, like Annas and Burnyeat61, affirm that character positively conditions future actions, since ethos is a compound of virtues, being stable habits acquired through repetition. That is the previously announced circularity between hexeis and praxeis. The problem is to understand whether training and metabolization can enrich our possibilities or if they produce some standardized and predictable patterns. In other words, what is at stake is the chance of acting differently from how we are supposed to do according to our ethos. Otherwise, there is a risk one might fall into what has been called the “prison of character”, since the habit would become too strong and inevitably influence praxeis. This problem has been perfectly analyzed by P. Donini, who states that

il possesso di un carattere formato, qualificato da abiti stabili, consente dunque a un agente di rispondere correttamente nel modo richiesto dalla situazioni che gli si presentano, anche senza calcolare e deliberare intorno ad ogni parola e a ogni azione: ma gli abiti e il carattere comportano una sorta di automatismo nelle risposte anche in un altro senso: escludono, cioè, di per sé la possibilità di dar corsa a un’integra serie di azioni.62

This concern for character crystallization finds some coherent remarks in the Aristotelian text, especially when Aristotle recognizes that it is difficult, if not impossible, to change an ethos once it has been acquired. The main reference is the debate about the alcoholic, who could have avoided drunkenness but did not, or

60 Donini P., Ethos, Aristotele e il determinismo, Edizioni dell’Orso, Torino, 1989.
61 Annas J., The morality of happiness, and Burnyeat M. F., Aristotle on learning to be good in Essays on Aristotle’s ethics, pp. 69-72.
62 Donini P., Ethos, Aristotele e il determinismo, p.79 and Abitudine e saggezza. Aristotele dall’Etica Eudemia all’Etica Nicomachea, chapter IV, Unidirezionalità degli abiti e posizione del fine, pp. 91–137.
the one, who became ill because of his way of life\textsuperscript{63}. In both cases, their activities produced the stable negative state and the consequence is that a sick person will not become healthy by merely wishing to become healthy\textsuperscript{64}. Again in Book V, Aristotle states that a habit does not leave room for opposite actions due to its tendency, such as from healthy only healthy things are done and not both contraries\textsuperscript{65}. However, it is true that Aristotle contemplates also the reversibility of ethos when he talks of friendship, or better when he faces equally the corruption and the improvement of a virtuous character. In both cases, the shared question is if the character involution\textsuperscript{66} or the superiority in virtue\textsuperscript{67} justifies or not the rupture of friendship.

As it is evident, the problem generated by character automatism cannot be determined easily and Aristotle offers no decisive help in answering this question. Anyhow, it is needless to say: in all the considered situations representing cases of practical wisdom, the general end is never discussed. For instance, who can bring the choice of playing the piano or of learning a language into question and how? According to Aristotle’s account, the same student cannot do that for two reasons. First, he is persuaded of its activity, which means that he has absorbed the right motivations\textsuperscript{68}. Second, the practical reason, as it has been described till now, does not consider such a possibility. Although simplified, this example directly leads to the second point, namely the eventual critic of the telos, the goal. We could now explain this problem by going back to the formal presentation of the practical syllogism. In the MaP, we saw that the highest good is never actually brought into question and it is the outcome of orexis plus doxa. As already remarked, it is accepted due to the habituation process and practical reason (MiP) is active in all its concrete specifications, in the kairos (given by the middle term). For this reason, phronesis is fully immanent or, in other words, never separated from the shared telos. Like that, all solutions and courses of action are always within the possibilities offered by the collective ethos. This sort of finalistic dogmatism could be highly dangerous, in particular because it inhibits any form of criticism. The intrinsically negative potentialities of an immanent practical reason becomes clear if we consider some paroxysmal examples: just imagine living in a context or form

\textsuperscript{63} Both the examples come from Nicomachean Ethics Book III, in the discussion about ignorance and responsible actions.
\textsuperscript{64} Ivi, Book III, 1114 a 15 -16.
\textsuperscript{65} Ivi, Book V, 1129 a 15.
\textsuperscript{66} Ivi, Book IX, 1165 a 13 – 14: “again, if one accepts another ad a friend, taking him as a good man, but the latter turns out to be evil and is regarded to be such, should he still be kept as a friend?”
\textsuperscript{67} Ibidem, 1165 b 23 - 25: “but if one friend were to remain the same while the other were to become better and far superior in virtue, should the latter treat the former as a friend, or should not?”
\textsuperscript{68} If you talk to an athlete or to a student of music or in general to someone who has for long time practiced an activity, he would certainly be able to give you many good reasons for doing that action, such as, for instance, it is healthy or it makes you feel good.
of life where wealth or physical wellness respectively are considered the highest good and everything (such as the educational system or the welfare state) always tends to it. In both cases, the simple exercise of phronesis does not provide any kind of instruments to discuss those ends, and the phronimos would be the agent able of better performing them.

This is potentially the aporetic climax of practical philosophy. And exactly here politike episteme intervenes, with its task of contesting and determining the telos. In fact, this form of rationality aims at discussing the shared ends and at testing their contradictions through the dialectical method. To better understand how practical philosophy works, the beginning of Nicomachean Ethics is paradigmatic, where Aristotle discusses the different conceptions of eudaimonia and he shows the incongruity of the ways of life based only on pleasure, wealth and honor (edone, ploutos and time) with respect to the human soul. In fact, the concepts of the human psyche and its excellence, perfection are the basis of the confutation of these common opinions. Again, another opportunity to witness how politike episteme works are Books VIII and X, where Aristotle talks of friendship (philia). Also in this case, the philosopher proceeds dialectically and demonstrates that the central type of philia rather than being based on pleasure (dia to edu) and advantage (are dia ton chresimon) is dia ten areten (according to virtue, so to character).

Moreover, thanks to these examples it is also possible to verify how far Aristotelian practical philosophy differs from the common sense. Indeed, on the one side, the recognition of bios theoretikos as the highest and most preferable way of life does not match the sensibility of the polis, that demands the political involvement typical of the bios praktikos. On the other side, although we have seen how uncommon moral excellence is, the perfect form of friendship is characterized by being the one among good men, who have developed the virtuous ethos. In both cases, politike episteme radically puts common sense into question and proposes alternative ways. These do not exclude shared opinions (such as

---

69 One of the clearest studies about the different meanings of the practical reason is the already quoted work of Berti E. Nuovi studi aristotelici vol. III, Filosofia pratica; especially chapters II, Ragione pratica e normatività in Aristotele, p.25-38 and chapter III, Phronesis et science politique, pp. 39-59.
70 Nicomachean Ethics, Book I, chapter 2.
71 Which in the Aristotelian vocabulary means the skill of accomplishing its most important part. In the case of the human soul, the highest function is the rational one, or better the bios theoretikos.
72 The definition of the highest kind of friendship is in Nicomachean Ethics, Book VIII.
73 Ivi, Book I and Book X.
74 Which is, after all, the proper human way of life, Nicomachean Ethics, Book X. However, the conciliation of the practical and contemplative lives is one of the biggest issue of the Nicomachean Ethics, played between an inclusive or exclusive interpretation.
pleasure, honor and wealth, or pleasure and advantage), but reallocate their importance in a wider perspective.

In the light of this, the distinction between the two co-existing plans in practical philosophy can be summarized by this picture: *phronesis* focuses on how to exercise skills properly, while *politike episteme* pays attention to the purpose on its whole and it is guided by the understanding of basic human properties. In this sense, the practical philosopher is a cardinal figure in Aristotelian reflection, since it guarantees the existence of a critical circularity between the *individual ethos* and the *collective* one. Whereas the *phronimos* and *phronesis* show the immanent practice of reason, practical philosophy demonstrates that there is a reflective and transcendent movement towards the *polis* and its constitution anyway. Again, while the immanence of *phronesis* is typically conservative, since it finds its categories only inside the *polis*, the transcendence of *politike episteme* assures a form of transformation and of dynamism. This also enlightens a new figurative representation of their relationship: *politike episteme* is the compass whose variable and dynamic openness delimits the space of *phronesis*, all statically \(^75\) contained within the circumference. Here all the different variations of *phronesis* take place \(^76\), while *politike episteme* remains external. This graphical representation also gives reason of another fundamental Aristotelian idea, especially if compared with Plato’s point of view: the distinction between the roles of the philosopher and the politician. The first one influences politics at the base (in terms of drawing boundaries by defining the goal, *telos*), whereas the second one practices the political exercise concretely. These two figures are never overlapping, as regards the different functions.

The importance of *politike episteme* emerges also by considering the contemporary interpretation of the Aristotelian practical philosophy. As a matter of fact, one shortcoming of many scholars, especially from the analytic scenario, has been to underestimate the role of practical philosophy, emphasizing only one kind of practical reason, *phronesis*. As Berti has shown, this has led to a misunderstanding of the different plans involved in the *praxis*, overestimating the

\(^{75}\) In this case, statically is not negatively qualified, since it is actually the right practice of a specific function. All the concepts of dynamism, transformation and conservatism rather than being the expression of values’ judgment want to be a neutral description of functions.

\(^{76}\) As a matter of fact, in the *Nicomachean Ethics* Aristotle clearly distinguishes different practices of *phronesis*: “both politics and prudence are the same disposition, but in essence they are not the same. Of prudence concerned with the stat, the one which is architectonic is legislative, while the other which is concerned with particulars has the common name ‘political prudence’; and the latter is concerned with particular *actions* and deliberations, for a particular measure voted on is like an individual thing to be acted upon. (...) Prudence is thought to be concerned most of all with matters relating to the person in whom it exists and with him only; and this disposition has the common name ‘prudence’. Of the other kinds, one is financial management, another is law-giving, and a third is political, of which one part is deliberative and the other judicial”, Book VI, 1141 b 24 – 35.

274
importance of *phronesis*. But the consideration of only a one-dimensional practical wisdom, merely endogenous, has had as a consequence a flattening of the agent on the context and an agent’s inability to dissociate himself from the values of life form. This problem becomes clear if we consider two examples, coming from authors who are influenced by this interpretation, J. Annas and A. MacIntyre. In *Intelligent Virtue*, Annas analyzes the case of someone who grew up in a slum. Annas' account is that the considered context radically compromises his possibilities of developing and practicing virtue. At the same time, in *After Virtue*, MacIntyre seems to connect virtues mainly to the common morality, moving the goal of criticizing contemporary ethics to the background. Even if it is true that when MacIntyre values the role of tradition he is thinking of the problem of individualism, his emphasis on the community and on its values apparently do not consider the possibility of refusing or redirecting that tradition. Although in MacIntyre’s account the importance of the interpretative framework does not entail neither a relativistic conception of values and truth, nor an exclusive interpretation of traditions (as if they could exist closed on themselves), his voice is functional to emphasize one of the possible risks caused by the understatement of the role of *politike episteme*.

This rapid account of a misunderstanding of the Aristotelian view leads us to the last part of this paper, whose content refers directly to contemporary philosophy. Here I will approach the problem of the relationship between the agent and the context by trying to understand if the Aristotelian answer is still suitable for ethical experience.

3. The pertinence of an inactual thought for the contemporary debate

To address this final topic, I would like to start from the above-mentioned problem of the imbalance in the relationship between the agent and the context,

---

77 In Annas’ thought it is important to remember that she distinguishes the difference between *not being able of virtue* and *not expecting virtue here*, where the first case if an error while the second is reasonable. That is why “we do not expect people raised on a garbage dumps outside a Third World megalopolis to be kind and generous in their everyday behavior, but this, I suggest, for the same kind of reason that we do not expect them to play the piano or to do the crossword. Their environment has obviously lacked the opportunities to learn and to do these things, and because this is so obvious we do not assume that they are naturally unable to do them”, p. 31.

78 MacIntyre A., *After Virtue*, University of Notre Dame Press, Notre Dame, Indiana, third edition, 2007: “I inherit from the past of my family, my city, my tribe, my nation, a variety of debts, inheritances, rightful expectations and obligations. These constitute the given of my life, my moral starting point. (...) This thought is likely to appear alien and even surprising from the standpoint of modern individualism”, p. 220.

79 Which is surely far more complicated than this brief reconstruction.
as it has come to light thanks to previous references to Annas and MacIntyre. My purpose is to recall this *aporia* and to show that, if seen through the analysis made in the first part, an asset has paradoxically emerged.

As already remarked, the *impasse* is created by the omission of the function of *politike episteme*. Once the task of practical philosophy is properly sharpened and the importance of transcendence has arisen, the role played by tradition and community still remains fundamental, but it is attenuated and specified. In these terms, *phronesis* and *politike episteme* co-operate and they can guarantee a bidirectional circularity between *individual ethos* and *collective one*. While *phronesis* proves the efficacy of the habituation process and of the individual responsibility, both within the *polis’* perspective, *politike episteme* assures the reflective movement upon the *polis* itself. At the exact point where Virtue Ethics seems to produce a deadlock, it is possible to rediscover the complexity of Aristotelian practical philosophy, the main strength of which is the ability of weaving together complementary tensions. As just seen, the harmony between the constitutive belonging to an ethical substance, or the immanence of *phronesis*, and its necessary passing, or the transcendence of *politike episteme*, are all perfect examples of this peculiarity.

The clarification of how all the different levels in practical philosophy imply each other revokes the evaluative principle introduced at the beginning, that is to say the coordination of concepts. As explained, each concept finds its sense in the network built by the relationship with others. The latter, meaning that if any concept is abstracted from the network, it loses its sense or its efficacy. For instance, the misunderstanding about the role of practical reason is the result of the overestimation of one concept, *phronesis*. But, allow me to establish another critical example, which has already been introduced many times: the desire revolution. Although it represents a fascinating acquisition in comparison with the rational prejudice that demonizes the importance of desire in the decision-making process, as I have tried to explain we can talk of it only by accepting the premises of Aristotelian ethics and psychology. Otherwise, the concept of desire is excluded from the context where it becomes productive and it loses its demonstrative strength.

From this point of view, we cannot look for an immediately implementable model in Aristotelian practical philosophy. In fact, a break has occurred between the modern conception of *praxis* and the ancient one; because of it, the cardinal concepts and their relationship have deeply changed\(^{80}\). An impracticable actualization of the whole train of thought, which is radically different from the contemporary one\(^{81}\), would be required to try to actualize one concept (e.g. desire).

---

\(^{80}\) This is the strong theory exposed by Chignola S. et Duso G. in *Storia dei concetti e filosofia politica*, FrancoAngeli, Milano, 2008.

\(^{81}\) A paradigmatic example is also the Aristotelian category of *agent*. As a matter of fact the
These remarks lead to the *inactuality* of Aristotelian practical philosophy and the necessity to respect its difference. This issue occupies the final part of the paper, which once again calls into question the relationship between *individual ethos* and *collective ethos* and introduces to the last and, probably, biggest problem. In fact, the concluding topic refers to the existence of one context or the *explosion* of the forms of life. This issue radically challenges the structure of the essay, since up to now I have focused on the connection between one agent and one life form (namely, the *polis*). The pluralization of the forms of life could jeopardize the thesis of the dynamic circularity between the two meanings of *ethos*, or could have to modify and extend it to this new input. In fact, the eventuality of considering many forms of life would confirm the *inactuality* of Aristotelian practical philosophy, which is based on a *one-to-one* relationship (the agent and the *polis*), and the impossibility to actualize it. To pursue this itinerary, first I will briefly draw a picture of what I mean by the pluralization of the contexts in contrast with the *one-to-one* connection, and then I will try to sum up some conclusive considerations.

First of all, at the beginning of the paper I defined the *collective ethos*, with all its meanings, as the context or the *polis*. In this sense, the *polis* has been thought of as a singular and coherent reality, where different parts interact with each other. This interpretation is suggested and confirmed also by Aristotle, when in *Politics* he talks about the relationship between the whole and the parts and the primacy of the whole. As a consequence, the relationship between the *individual ethos* and the *collective* one is bidirectional, initially imbalanced towards the *polis* but never vicious thanks to the roles of practical reason. But what is relevant at this point is that Aristotle rigidly establishes both practical agent and *polis*. Indeed, Aristotle is thinking of a precise *polis*, Athens, in contrast with the other *polis* and the Barbarians. Moreover, the philosopher is considering a specific agent, the free man, in contrast with all the other excluded categories (e.g. women philosopher in his *logoi*, and generally in his philosophy, is thinking of specific figure: the adult and free man in the *polis*. The evident consequence is that he is leaving out from this category many other figures that we are used to consider in it, such as women or children. I will come back soon on the issue of *polis* dimension.

---

82 See the *Politics*, Second Edition, University of Chicago Press, Chicago, 2013, Book I, 1253 a 18 ff.: “Further, the state is by nature clearly prior to the family and to the individual, since the whole is of necessity prior to the part; for example, if the whole body be destroyed, there will be no foot or hand, except in an equivocal sense, as we might speak of a stone hand; for when destroyed the hand will be no better than that. But things are defined by their working and power; and we ought not to say that they are the same when they no longer have their proper quality, but only that they have the same name. The proof that the state is a creation of nature and prior to the individual is that the individual, when isolated, is not self-sufficing; and therefore he is like a part in relation to the whole. But he who is unable to live in society, or who has no need because he is sufficient for himself, must be either a beast or a god: he is no part of a state.”
or slaves). As a consequence, the match between the possible combinations is all played inside these two specific terms and it does not consider different kinds of agents or forms of life.

To schematize it, we could imagine this kind of representation $X \leftrightarrow y$, where $X$ stands for the context (or the whole, or the polis), $y$ for the agent, and the majuscule or minuscule indicates the balance of power between them. But this rigid bidirectional equivalence no longer fits our ethical experience, due to one main reason: the expansions of the contexts, or what C. Taylor has defined as the "nova effect". With this expression the Canadian philosopher indicates the pluralization of perspectives and the openness of different forms of life. Rethinking the figurative representation seen above, it becomes clear that the two parts of the equivalence no longer correspond to the ethical scenario. To begin with the first term of the relationship ($X$), instead of living within one context we now have to face the co-existence of many different ways of life. If in Aristotelian Practical Philosophy was acceptable to recognize a specific collective ethos (namely, the polis) as the reference frame, in the contemporary ethical horizon this strict identification is reductive. Not only the polis is not the current political dimension, but it has also been "substituted" by a multiplicity of co-existing contexts. In this sense, the texture composed by the forms of life seems to be thicker and more dynamic if compared to the case of the polis. To pass from one graphic representation to another, the supernova effect could be drawn as a stratification and intersection of many contextual levels, the directions and intensities of which are different.

But in this perspective, the focal point of which is the bidirectional relationship between the two means of ethos, the pluralization of the frames of reference problematizes two main questions:

1. Given that the basis of virtue are political, not natural, which grid of intelligibility is prevalent in the character preformation?
2. What happens to the correlative minor term ($Y$) of the equivalence? That is to say, what changes in the way we approach the issue of the agent?

Regarding the grid of intelligibility, the question could be posed in this way: in Aristotelian Practical Philosophy the emphasis on the context corresponds to the early stage of the character habituation, when virtue is induced and the agent educated. In this sense, if the polis was responsible for the character habituation, who plays this role now and how? Let us consider an example: what mainly

---

83 Taylor C., *A Secular Age*, Harvard University Press, 2009. In particular, the chapter *The Nova Effect*, pp. 299-419. In that case, Taylor talks of the nova effect in the context of the secularization problem. From this paper point of view, we will use that expression to identify the pluralization process.
The Concept of Ethos: Aristotle and the Contemporary Ethical Debate

characterizes our self-perception and identity? The sense of belonging to our city, to our country, to the European Union, to a certain conception of the European Union or to any other structure of society? What level actually prevails on the others and in which case, given that each one seems to be reasonable and acceptable? Consequently, if the concept of the form of life is modified or extended, also the agent has to be put into question (not to mention that the concept itself of an agent enlarged from the Aristotelian free adult man). From this point of view, the issue of the pluralization of collective ethos (X) radically involves the correlative term: the agent (y). Indeed, recalling the graphic representation of a stratification and an intersection of different contextual levels, the agent seems to discover himself as a changing segment that experiences and crosses some of them. Moreover, the pluralization of forms of life concerns not only the "passive" phase (so, the character preformation), but also the "active" one, since all the practices of practical reason are displayed through the constitutive reference to these complex stratification of contexts. As a matter of fact, a shared and recognized pattern of virtues is the premise for phronesis, for the consequent recognition of the phronimos and for the critical task of politike episteme.

It is unlikely that this is the right place to analyze and discuss in greater detail the depth of the problem of the pluralization of forms of life, that would deserve an independent and much wider analysis. However, what is interesting here is that the current ethical scenario seems to be more complex and articulated than the one described by Aristotle and this density is the cause of a new impasse. In fact, this stalemate is produced by the change in the definitions of the terms of agent and context, and, as a result, of their relationship itself.

Now let me conclude by making some remarks, that will summarize the hot-spots of the paper and will recall the above-seen issue.

Conclusion

As previously stated, I would like to conclude by briefly summing up some of the most relevant points emerged in the paper. They are principally two: (1) the intimate contiguity and bidirectionality between individual ethos and collective one and (2) the impasse of the distance and the solution of the contingence.

The demonstration of the first theme has occupied all the initial section and has been encouraged by the reconstruction of the contemporary debate. First, thanks to Anscombe and Slote’s analysis, a window into Virtue Ethics and its reasons has been given. Then, the clarified necessity of rethinking the moral agent has supported the examination of practical philosophy, guided by the inclusive concept of ethos. An itinerary through a composite complex of Aristotelian writings (Nicomachean Ethics, Eudemian Ethics, Politics and De Anima) has led to
the conclusion that the relationship between the two meanings of ethos is both circular and dynamic. These two features are guaranteed by the distinct steps of the moral development. From one side, we have (a) the metabolization or habituation process, denoted with the induced virtue. From the other side, we have (b) the co-existence of different levels of freedom, identified with the two practices of practical reason, prudence and political knowledge. Moreover, another fundamental characteristic of Aristotelian practical philosophy has come to light: the constant balance between desire and reason. Its examination is unavoidable in order to fully understand the circularity just mentioned, because it gives reason of all the dimensions involved in the inner workings behind the above-seen expression.

An intermediate remark about the inactuality of Aristotelian Practical Philosophy has then conducted to the last section of the essay, whose problem is the legitimacy of Aristotelian practical philosophy in contemporary thought. Thus, I tried to draw the attention to an aporetic moment, referring to the problem of the collective ethos and its pluralization. From this point of view, the distance and the inactuality of Aristotelian reflection are at the same time an advantage and a limit. But this is precisely this issue I would like to recall here. Indeed, I would like to point out that one of the main contributions of this analysis is having demonstrated that next to the importance of the distance there is always the need for the contingency and for the specificity of the situation. In this sense, ancient philosophy does not directly provide an abstract and fixed model to be put into practice, since the same practice would contradict it. After all, the immanence of the praxis is one of the main Aristotelian lessons, as it is proved by his care for endoxa (shared opinions) in the dialectical method and for the role of prudence in the decision-making process. Thus, if the rediscovery of Aristotelian practical philosophy wants to be included in the contemporary ethical horizon, it must always be accompanied by an immanent point of view, that has to take the peculiarities of each specific context and historical situation into consideration.

Bibliography

The Concept of Ethos: Aristotle and the Contemporary Ethical Debate

Donini P., Ethos, Aristotele e il determinismo, Edizioni dell’Orso, Torino, 1989.
Natali C., La saggezza di Aristotele, Bibliopolis, Napoli, 1989.
Natali C. e Maso S.(edited by), La catena delle cause, determinismo e antideterminismo nel pensiero antico e in quello contemporaneo, Hakkert, Amsterdam, 2005.
Salis R., La polis tra filosofia e storia nel primo libro della Politica di Aristotele, in Aristotele e la storia, a cura di in Rossitto C., Coppola A. e Biasutti F., Cleup, Padova 2013.
Condividere relazioni: riflessioni leibniziane a partire dall’esperienza dei social

Pierpaolo Marrone
Università di Trieste
Dipartimento di Studi Umanistici
marrone@units.it

ABSTRACT
In this paper, I offer a short reflection on the problem of sharing relations with other subjects. I think that this problem is exacerbated by the experience of social networks. I show that some suggestions from Leibniz’s philosophy can be helpful in understanding this important social experience.

KEYWORDS
Leibniz, monadology, social network

Se è vero, come sosteneva Hegel, che la filosofia è il pensiero del presente, allora nulla di quanto accade ci può essere indifferente in linea di principio e tutto può essere oggetto di discussione.1 Tutto quanto pertiene alle relazioni umane in particolare è oggetto potenziale della nostra considerazione e riflessione etica. Deve essere così, perché noi non sappiamo a priori che cosa è realmente rilevante nelle relazioni umane. Del resto, una delle lezioni della filosofia di Hegel credo sia proprio questa, ossia che tutto è relazione e nulla può essere compreso nella sua individualità a meno che non comprendiamo che questa sua individualità è dialetticamente negata nella nostra stessa esperienza.

Questo brevissimo cappello introduttivo per proporre un’analisi, del tutto parziale e forse anche aneddotica, di alcuni problemi che mi paiono sollevati dalla nostra esperienza all’interno dei social network, o meglio: dell’esperienza che accade nei social.

Sembrerebbe che l’enorme successo dei social sia dovuto a una necessità di condivisione delle proprie esperienze tra gli esser umani che può legittimamente essere vista come un aspetto della naturale socialità umana e della generale propensione, che naturalmente non esclude affatto

comportamenti egoistici, alla cooperazione tra gli attori sociali. Condividiamo foto, ricordi, eventi ai quali parteciperemo, chattiamo con persone che mai conosceremo di persona, discutiamo dei libri che abbiamo letto, dei film e delle serie che abbiamo visto, dei viaggi che abbiamo fatto, dei piatti che abbiamo mangiato. Facciamo vedere, e spesso esibiamo, le foto dei nostri figli, dei nostri partner occasionali, dei nostri matrimoni. C’è ovviamente molto esibizionismo in tutto questo, ma si potrebbe sempre dire che non c’è gusto a vivere qualcosa se non sei in grado di condividerlo con qualcuno. Naturalmente, il problema dei social è che questo ‘qualcuno’ è un indeterminato, per cui il sospetto di un esibizionismo e di un narcisismo di massa è ampiamente giustificato. Tutto questo non significa affatto avere un atteggiamento moralistico di condanna, ma semplicemente cercare di comprendere le motivazioni di un’esposizione di parti della propria vita. In fin dei conti, tutto questo sembra, ancora una volta, dare ragione ad Hegel: anche questa esposizione pubblica non è altro che una declinazione dell’idea che noi siamo un fascio di relazioni e fuori da questo intreccio non esistiamo effettivamente. Per questo, si potrebbe dire, i social sono filosoficamente interessanti. In definitiva, non fanno altro che intensificare quella che è una nostra caratteristica ontologica, ossia il bisogno di condivisione.

Proverò in queste pagine a mettere alla prova questo bisogno con l’ausilio di alcune teorie di Leibniz. La cosa potrebbe suonare strana, ma mi auguro di mostrare che esistono delle buone ragioni per ricercare in alcune teorie di Leibniz un ausilio per comprendere questi importanti fenomeni sociali ed etici. Non pretendendo di procedere in maniera sistematica, perché l’esperienza sui social, dalla quale queste mie riflessioni si originano sistematica non è. Si tratta dell’approccio di un utente interessato e non di quello di uno scienziato sociale. Tuttavia, mi auguro che alla fine qualcosa di sensato, nella medesima problematicità di questa esperienza, ne risulterà.

Il problema che mi interessa è ovviamente la condivisione delle esperienze, ma non tanto la condivisione che vediamo esibita effettivamente sui social. Questa è sempre empiricamente parziale e non potrebbe essere altrimenti. Quello che mi interessa è la promessa di condivisione reale e intima, se non totale, che l’esperienza dei social ci promette. Sappiamo che non è così, sappiamo che si tratta molto spesso di una condivisione che non si realizza affatto. Tuttavia, questa condivisione ci è promessa. Ed allora ecco la domanda: che cosa significa condividere in un senso profondo? È una domanda che ho provato ad analizzare proprio a seguito di un post letto su un social, elaborato in termini scarni e assoluti, sarei tentato di dire, da una
Leggevo sulla sua bacheca questa domanda disarmante e radicale: “pochi minuti fa guardando un trailer ho sentito una domanda disarmante, a cui non ho saputo dare alcuna risposta: come si condivide la vita con qualcuno?”. Poco dopo, vedevo il post di una conoscente che metteva sulla propria bacheca una citazione che vi riporto: “the secret of change is to focus all of your energy, not on fighting the old, but on building the new”. L'autore di questa è presentato come Socrates (più probabilmente il protagonista delle opere di Platone, che non il calciatore brasiliano, immagino).

L'autrice della prima riflessione a me sembra onesta sino alla scarnificazione, perché inquadra in modo assolutamente semplice e diretto un problema enorme. Al medesimo tempo, sembra dare una risposta a questo problema. E la risposta che indica implicitamente è o il suggerimento di una semplice constatazione, relativa alla propria incapacità di rispondere alla domanda, oppure, cosa che io ovviamente sono propenso a credere, una risposta completamente pessimistica sulla impossibilità di un commercio intimo e fondamentale tra due individui (diverso dal commercio intimo e superficiale, offerto dal sesso ad esempio), i quali fanno esperienze diverse anche quando credono di condividere uno stesso contesto (magari quello dell'innamoramento e della fusione).

Questa ultima credenza è infondata: non c’è nessun contesto comune, ma solo il proprio punto di vista che non riflette certo quello di un’altra persona, destinata a rimanerci inaccessibile. In definitiva, se entriamo in contatto con qualcuno è perché pensiamo si incastri in qualche nostro bisogno. Questo nostro bisogno non sarà mai sovrapponibile al suo in maniera completa, ma, oltre a questo, noi non sapremo mai nemmeno le ragioni di questa impossibilità. Noi non viviamo se non nella nostra mente. Questa stessa mente ci è largamente inaccessibile. Come potremmo allora mai credere di avere un accesso privilegiato alla mente di qualcun altro? Eppure, tanto le esperienze fusionali quanto le esperienze sui social ci fanno intravedere un accesso alle altri menti, e non un accesso qualsiasi, bensì un accesso significativo.

Vediamo al secondo avvenimento. La citazione postata dalla mia conoscente (“the secret of change is to focus all of your energy, not on fighting the old, but on building the new”), se attribuita al Socrate che noi conosciamo attraverso le opere di Platone e di Senofonte, è naturalmente

---

2 La scrittrice è Mary Barbara Tolusso.
falsa. Facendo una ricerca si scopre che la fonte è tale Dan Millman, autore di manuali di self-help, tra i quali un *La via del guerriero di pace*, che sia nell’edizione italiana sia in quella americana ha come modesto sottotitolo questo: *Un libro che cambia la vita* (probabilmente quella dell’autore, che ne ha vendute parecchie copie). Non dirò nulla sui manuali di self-help, né sul fatto che talvolta sono purtroppo scritti da praticanti di arti marziali (come sembra essere questo caso), che pensano di essere in contatto con abissali profondità metafisiche e morali, perché hanno raggiunto una certa capacità tecnica e atletica. Quello che mi interessa è invece un fenomeno diverso, ossia quello delle citazioni apocriffa e la cosa riveste un suo notevole interesse.

Tempo fa un mio amico, di formazione filosofica, ma con un’ampia esperienza professionale nel mondo di internet, scrisse nel suo blog che attribuiva la massima quantità possibile delle cose che diceva a Shakespeare. Nessuno tra i suoi interlocutori era un lettore assiduo del teatro elisabettiano, ma tutti si sentivano in qualche modo obbligati ad essere intimoriti. In questo senso, l’attribuzione fasulla è l’analogo dell’utilizzo di formalismi matematici inutili nelle scienze sociali. È un utilizzo terroristico in entrambi i casi. ‘Lo ha detto Shakespeare!’ ‘Lo dice la matematica!’: cose fatte apposta per esonerarsi da uno sforzo critico. Ma non si tratta solo di questo; io credo; piuttosto: attribuire a qualcuno di importante dei pensieri semplici, che ognuno avrebbe potuto incontrare scartando qualche Bacio Perugina, ci fa sentire più interessanti e più profondi ai nostri stessi occhi. E ci fa pensare che queste siano parole piene di significato, se non di vita e di speranza: change, focus, energy, building, new. Anche se, a pensarci bene, il change può essere quello di tua moglie che ti butta fuori di casa o quello di chi perde il suo impiego e non riesce a rientrare nel mercato del lavoro. L’energy può essere quella del Tir che ti investe in autostrada o quella delle cellule del tuo corpo che si riproducono incontrollatamente nella malattia che ti porterà alla morte. Il focus può essere il tuo sordo rancore per un torto che hai subito e per il quale aspetti il momento opportuno per esercitare quella che tu ritieni essere la tua giusta vendetta. Lasciamo invece perdere e allora pensiamo tutti all’unisono davanti al nostro schermo: change, focus, energy, building, new, quasi fossimo a una convention aziendale per piazzisti. E in effetti qualcosa piazziamo l’un l’altro: l’energia – anzi: l’energy – di una ripetizione.
automatica, che deve pur avere le sue ragioni. La banalità, infatti, in fin dei conti ha le sue ragioni ed anche con una qualche verità, sebbene, sempre lo crediamo, in un senso ovvio e poco interessante, per quanto potrebbe non essere così per un motivo sul quale ritornò. 

Eccoci allora a Leibniz. Gli elementi di interesse in Leibniz sono innumerevoli, ma penso che almeno tre sue celebri teorie possono gettare una luce all’interrogativo della scrittrice, che ha posto realmente una questione metafisica: possiamo condividere qualcosa? La cosa interessante è che queste tre teorie sembrano andare in due direzioni radicalmente differenti.

La prima è la sua dottrina delle monadi. “1. La monade, di cui intendiamo parlare, è una sostanza semplice, che entra nei composti: semplice, cioè senza parti.”

Si tratta di sostanze spirituali che replicano all’infinito la definizione che di sostanza aveva dato Descartes: la sostanza è una cosa che esiste in maniera tale da non avere bisogno di null’altro per esistere. Nel senso di Descartes non è una sostanza né la penna con cui scrivo, né il mio tablet, né i pensieri che hanno costituito la progettazione sia della mia penna sia del mio tablet. Non sono una sostanza io stesso, dal momento che dipendo per esistere sia dai miei genitori, sia dal pane che mi guadagno (e da mille altre cose, tra le quali essere stato creato da Dio). È una definizione che sembrerebbe non lasciare altro spazio se non all’affermazione di un dio, increato e onnipotente e onnisciente, come unica sostanza esistente. Al di fuori di lui, tutto il resto ha bisogno di qualcos’altro per esistere.

Molti di noi sanno come va a finire: Descartes aggiunge che dal momento che ci sono cose che hanno bisogno solo di Dio per esistere, anche di queste si può, in fin dei conti, dire che si tratta pur sempre di sostanze. Così, anche del pensiero (il mondo delle menti) e dell’estensione (il mondo delle cose materiali) possiamo con buon diritto dire che sono sostanze. Descartes scopre addirittura che noi siamo in grado di dimostrare di essere soprattutto ed essenzialmente delle sostanze mentali (è il suo famoso ragionamento: ‘penso, dunque sono, dunque sono una sostanza pensante’, ritenuto un paralogismo, ossia un ragionamento solo apparentemente corretto). Fin qui Descartes.

Leibniz moltiplica all’infinito l’autonomia che Descartes aveva pensato come attributo della sostanza-dio e la attribuisce a tutte le infinite sostanze spirituali che secondo lui costituiscono la realtà, e che lui chiama appunto

4 G.W. Leibniz, La monadologia, Bari-Roma, Laterza, 1975, p. 117.
5 R. Descartes, Opere filosofiche, Torino, Utet, 2013, p. 750.
monadi. Facendo questo, radicalizza l’autonomia di Descartes e da questa radicalizzazione deriva la sua dottrina delle monadi che non hanno né porte né finestre, ossia non comunicano realmente tra di loro, sebbene riflettano, ognuno dal proprio punto di vista, l’intero universo e quindi anche tutte le altre monadi con ognuno dei loro attributi, anche i più infimi. “7. Neppure v’è un mezzo per spiegare come una monade possa venire alterata o mutata nel suo interno per opera di qualche altra creatura, poiché non è concepibile in essa alcuna trasposizione di parti, e alcun mutamento interno che vi possa essere eccitato, diretto, accresciuto, come avviene nei composti, dove vi sono mutamenti tra le parti. Le monadi non hanno finestre, attraverso le quali qualcosa possa entrare od uscire. Gli accidenti non potrebbero distaccarsi dalle sostanze, né passeggiare fuori di esse, come una volta facevano le specie sensibili degli scolastici. Così, né sostanza né accidente può penetrare dal di fuori nell’interno d’una monade”.6

Capite, allora quale enorme problema si pone: le monadi non condividono nulla. “10. Inoltre considero come ammesso che ogni natura creata, e quindi anche la monade, è soggetta al mutamento, e che questo è continuo in tutte. 11. Di qui segue che i mutamenti naturali delle monadi derivano da un principio interno, perché una causa esteriore non potrebbe influire sul loro interno.” Mai e dall’eternità ci sarà una comunicazione reale tra le monadi, anche se ognuna di loro riflette l’intero universo dal suo punto di vista. “56. Ora questa connessione o questo adattamento reciproco di tutte le cose create fa che ciascuna sostanza semplice abbia dei rapporti esprimenti tutte le altre, e che essa sia per conseguenza uno specchio vivente, perpetuo, dell’universo”8 Gli attributi di una monade non potranno, mai e poi mai, essere quelli di un’altra monade, con la quale del resto non potrebbe essere in contatto reale, poiché se vogliamo prendere sul serio l’autonomia non possiamo farne questione di gradi, ma di tutto o niente. Se condividessero un mondo sarebbero la stessa monade: è questa la dottrina leibniziana dell’identità degli indiscernibili. “9. Occorre ancora che ciascuna monade sia differente da ogni altra. Infatti non vi sono mai in natura due esseri perfettamente eguali, tali cioè che non si possa scorgere in essi alcuna differenza interna, o fondata su una denominazione intrinseca.”9 Se due monadi condividono uno stesso mondo, ossia hanno gli stessi attributi, allora

---

6 G.W. Leibniz, La monadologia, cit, pp. 118-119.
7 G.W. Leibniz, La monadologia, cit, p. 120.
8 G.W. Leibniz, La monadologia, cit, p. 139.
9 G.W. Leibniz, La monadologia, cit, p. 120.
Condividere relazioni. Riflessioni leibniziane a partire dall’esperienza dei social

non sono due, bensì una. Ovvero: non si condivide nulla realmente e radicalmente, se non nell’identità.

È l’ultima parola di Leibniz? Forse no, e potrebbe essere non sia così se pensiamo alla sua teoria della verità. La teoria della verità che noi, direi, spontaneamente abbracciamo è quella aristotelica: un enunciato è vero o falso se il suo contenuto corrisponde o non corrisponde a quanto c’è nella realtà. “L’enunciato ‘la neve è bianca’ è vero se e soltanto se la neve è bianca” ossia se nella esperienza che facciamo fuori dalla nostra mente si verificano delle condizioni causali che fanno sì che la neve sia bianca. Leibniz la pensa diversamente e non può che essere così poiché le monadi non entrano mai in relazioni causali con altre monadi. Per Leibniz un enunciato è vero se siamo in presenza di un enunciato analitico, ossia di un enunciato che è sempre vero in virtù della definizione dei termini che lo compongono, o della sua descrizione definita in maniera esaustiva ed esclusiva. Il problema è che di verità sembra essercene fin troppa, anche senza gli enunciati analitici: ‘Cesare ha varcato il Rubicone’; ‘Putin era un funzionario del KGB sovietico’; ‘Berlusconi è il fondatore di Forza Italia’. Chi penserebbe mai che tutti questi, e infiniti altri, enunciati che dicono qualcosa di Cesare, Putin, Berlusconi non siano anch’essi veri? Per Leibniz infatti lo sono, ma non perché a un certo momento, di punto in bianco, Cesare attraversa il Rubicone, Putin intraprende una carriera nei servizi segreti sovietici, Berlusconi sceglie di scendere in campo, bensì perché le stesse ragioni che rendono gli enunciati ‘il triangolo ha tre lati’ o ‘ogni scapolo è non sposato’ sempre veri: varcare il Rubicone, fare carriera nel KGB, fondare Forza Italia sono predicati contenuti sin da sempre nei loro soggetti ossia nella loro descrizione definita in termini esaustivi ed esclusivi. Se Berlusconi non avesse fondato Forza Italia non sarebbe Berlusconi, così come non sarebbe Putin quello che non facesse carriera nel KGB, e non sarebbe Cesare quello che non avesse detto “Alea iacta est!”.

È una teoria apparentemente semplice che Leibniz porta a conseguenze vertiginose, come gli era capitato di fare con l’idea di autonomia. Se tutte le verità sono analitiche, tutto allora è necessario. Se noi avessimo la mente del Dio di Leibniz, potremmo capire in tutti i suoi passaggi formali che il destino di Berlusconi di scendere in politica è analogo alla necessità degli angoli interni di un triangolo di essere la somma di due angoli retti. Di più: la discesa in campo di Berlusconi fa parte analiticamente del concetto di Berlusconi non meno di me che adesso ne scrivo, proprio in questo momento, e non meno di voi, che ne state leggendo ora o che ne leggerete in futuro, o di

Se le sostanze sono autonome e quindi spirituali, allora non comunicano; se la verità è analitica, allora tutto accade per necessità. Ma se tutto accade per necessità, allora noi non siamo autonomi in nessuno dei significati che il senso comune attribuisce all’espressione autonomia. Nessuna tragedia in tutto questo per Leibniz, ma anzi un senso di sollievo quasi: “così è facile giudicare che l’anima sia un automa spirituale […] che, per effetto della preformazione divina, produca quelle meravigliose idee, in cui la nostra volontà non gioca alcun ruolo, e alle quali la nostra arte non saprebbe arrivare.”

Ciò che dunque Leibniz mi costringe a riconoscere è che sono stato troppo severo e moralista a giudicare delle false citazioni e del plauso che vi viene riservato, perché banali e stupide e poco profonde. Anche loro sono parte della definizione analitica di qualsiasi sostanza, anche loro fanno parte del tutto che ogni sostanza è. In una concezione in cui l’anima è un automa, quello che noi snobisticamente chiamiamo banalità è la parte maggioritaria del nostro comportamento. La maggior parte dei nostri atti sono atti meccanici, che tutti potrebbero compiere e non hanno nulla di speciale, e tanto meno ce l’ha la glorificazione del pensiero critico, come momento di illuminazione in cui viene esaltata la nostra originalità e la nostra individualità. Ma se le cose stanno così, allora sono adesso finalmente in grado di rispondere, attraverso Leibniz, alla domanda iniziale della scrittrice: come si può condividere una vita? La risposta è che non si può. Nulla si condivide sul serio, se non con se stessi e del tutto parzialmente. Nello stesso tempo, devo dire che tutto si condivide, perché ognuno di noi è un automa spirituale che compie dei gesti e dei pensieri che non sono certo suoi (almeno la grandissima parte delle volte, se proprio non desideriamo identificarci con Leibniz). E la conclusione allora qual è? Non saprei dirlo con precisione, ma mi sento invece di dire che è meglio così, è meglio non condividere, ché se davvero condividessimo realmente una vita – in maniera completa e prendendo sul serio questa espressione – andrebbe a finire che moriremmo di noia. Anche se quella vita fosse in un qualsiasi senso la nostra vita che

10 G.W. Leibniz, Saggi di teodicea, Milano, Rizzoli, 1993, pp. 338-339.
emerge completamente dall’oscurità meccanica e senza interesse della maggior parte dei nostri gesti e dei nostri pensieri, nella quale, innegabilmente, è molto meglio che vi rimanga.


**Information on the Journal**

*Etica & Politica/Ethics & Politics* is an open access philosophical journal, being published only in an electronic format. The journal aims at promoting research and reflection, both historically and theoretically, in the field of moral and political philosophy, with no cultural preclusion or adhesion to any cultural current. Contributions should be submitted in one of these languages: Italian, English, French, German, Portuguese, Spanish. All essays should include an English abstract of max. 200 words. The editorial staff especially welcomes interdisciplinary contributions with special attention to the main trends of the world of practice. The journal has an anonymous double peer review referee system. Three issues per year are expected. All products on this site are released with a Creative Commons license (CC BY-NC-SA 2.5 IT) http://creativecommons.org/licenses/by-nc-sa/2.5/it/

*Etica & Politica/Ethics & Politics* position on publishing ethics. The Editors of *E&P* have taken every possible measure to ensure the quality of the material here published and, in particular, they guarantee that peer review at their journal is fair, unbiased and timely, and that all papers have been reviewed by unprejudiced and qualified reviewers. The publication of an article through a peer-review process is intended as an essential feature of any serious scientific community. The decision to accept or reject a paper for publication is based on the paper’s relevance, originality and clarity, the study’s validity and its relevance to the mission of the journal. In order to guarantee the quality of the published papers, the Editors encourage reviewers to provide detailed comments to motivate their decisions. The comments will help the Editorial Board to decide the outcome of the paper, and will help to justify this decision to the author. If the paper is accepted with the request of revision, the comments should guide the author in making the revisions for the final manuscript. All material submitted to the journal remains confidential while under review. Once the author receives a positive answer, he/she should send the final version of the article since proofs will not be sent to him/her. *E&P* will publish the paper within twelve months from the moment of the acceptance, and the author will be informed of the publication. The journal is committed to such standards as originality in research papers, precise references in discussing other scholars’ positions, avoiding plagiarism. *E&P* takes these standards extremely seriously, because we
think that they embody scientific method and are the mark of real scholarly communication. Since *Etica & Politica/Ethics & Politics* is devoted solely to scientific and academic quality, the journal does not have any submission charges. Past issues with download and visitors statistics for each article are provided here: http://www.openstarts.units.it/dspace/handle/10077/4673

**DIREZIONE/EDITOR:**
RICCARDO FANCIULLACCI (Venezia) riccardofanciullacci@libero.it
PIERPAOLO MARRONE (Trieste) marrone@units.it

**REDAZIONE/EDITORIAL BOARD:**
ELVIO BACCARINI (Rijeka) ebaccarini@ffri.hr
ROBERTO FESTA (Trieste) festa@units.it
GIOVANNI GIORGINI (Bologna) giovanni.giorgini@unibo.it
EDOARDO GREBLO (Trieste) grebloe@libero.it
FABIO POLIDORI (Trieste) polidori@units.it

**WEBMASTER:**
ENRICO MARCHETTO (Trieste) enrico.marchetto@gmail.com

**COMITATO SCIENTIFICO NAZIONALE/ITALIAN ADVISORY BOARD:**
A. AGNELLI † (Trieste), G. ALLINEY (Macerata), S. AMATO (Catania), M. ANZALONE (Napoli), D. ARDILLI (Modena), F. ARONADIO (Roma), G. AZZONI (Pavia), F. BACCHINI (Sassari), E. BERTI (Padova), M. BETTETINI (Milano), P. BETTINESCHI (Venezia), R. CAPORALI (Bologna), A.A. CASSI (Bergamo), G. CATAPANO (Padova), L. COVA (Trieste), S. CREMASCHI (Vercelli), G. CEVOLANI (Modena), U. CURI (Padova), G. DE ANNA (Udine), P. DONATELLI (Roma), P. DONINI (Milano), M. FARAGUNA (Trieste), M. FERRARIS (Torino), L. FLORIDI (Oxford), R. FREGA (Bologna), S. FUSELLI (Verona), C. GALLI (Bologna), R. GIOVAGNOLI (Roma), P. KOBAU (Torino), E. IRRERA (Bologna), E. LEGALDANO (Roma), E. MANGANARO (Trieste), G. MANIACI (Palermo), R. MARTINELLI (Trieste), F.G. MENGA (Tübingen), R. MORDACCI (Milano), V. MORFINO (Milano), B. DE MORI (Padova), M. PAGANO (Vercelli), G. PELLEGRINO (Roma), V. RASINI (Modena-Reggio Emilia), M. REICHLIN (Milano), M. RENZO (Stirling), A. RIGOBELLO (Roma), P.A. ROVATTI (Trieste), A. RUSSO (Trieste), M. SBISÀ (Trieste), S. SEMPLICI (Roma), A. SCHIAVELLO (Palermo), A. SCIAMÈ (Bergamo), M. Sgarbi (Venezia), F. TOTO (Roma), F. TRABATTONI (Milano), F. TRIFIRÒ
(London), M.S. Vaccarezza (Genova), C. Vigna (Venezia), S. Zeppi † (Trieste)

Comitato scientifico internazionale/International Advisory Board:
J. Allan (New Zealand), K. Ballestrem (Germany), T. Bedorf (Germany), G. Betz (Germany), W. Block (USA), M. Byron (USA), S. Chambers (Canada), J. Coleman (UK), C. Cowley (Ireland), W. Edelglass (USA), C.L. Geshekter (USA), A. Kalyvas (USA), J. Kelemen (Hungary), F. Klampfer (Slovenia), M. Knoll (Turkey), C. Illies (Germany), D. Innerarity (Spain), A. Lever (Switzerland), H. Lindahl (Netherlands), J. Marti (Spain), M. Matulovic (Croatia), J. McCormick (USA), N. Miscevic (Croatia), A. Moles (Hungary), L. Paulson (France), A. Przybleski (Poland), J. Quong (USA) V. Rakic (Serbia), A. Schaap (UK), N. Tarcov (USA), D. Webb (UK), J.P. Zamora Bonilla (Spain).