

MCL
B1
P46

Pensées

Canadian Undergraduate Journal of Philosophy
Revue étudiante canadienne de philosophie

McGILL UNIVERSITY LIBRARIES
SEP 7 2004
McGILL UNIVERSITY LIBRARIES

Volume 3, 2002

McGill University

Pensées

Canadian Undergraduate Journal of Philosophy
Revue étudiante canadienne de philosophie

Volume 3, 2002

McGill University

2134031
mcl
v.3
2002

Pensées

We would like to thank the following contributors for their generous support:

Canadian Philosophical Association – Association canadienne de philosophie
Trent University Department of Philosophy
McGill University Department of Philosophy
Students' Society of McGill University
Principal Bernard Shapiro
Arts Undergraduate Society
Carman Miller, Dean of Arts
Lucille and Aubrey Groll

Pensées is published annually by the Philosophy Students' Association of McGill University, Montreal, Quebec. Undergraduates from Canadian universities are encouraged to submit papers in either English or French. Please mail submissions, comments and questions to:

Pensées: Canadian Undergraduate Journal of Philosophy
c/o Department of Philosophy
McGill University
Room 908, Leacock Building
855 Sherbrooke St. W.
Montreal, Quebec, H3A 2T7

E-mail: cujp@hotmail.com

For information on submitting or editing or to read past copies of the journal, please visit our website:
www.pensees.homestead.com

Copyright of the papers included in this journal remains the property of the authors.

ISSN: 1492-2797

Editorial Board

Editors-in-Chief

Avi Cramer
Daniel Groll

French Translation

Gabriel Hébert

Copy Editor

Zach Dubinsky

Editors

Megan Bradley (University of St. Andrews, Scotland)
Mike Caie (McGill University)
Anna Carasthatis (University of Alberta)
Bethany Fisher (McGill University)
Guillaume Gingembre (McGill University)
Jay Gordon (University of Guelph)
Christian Lavoie (McGill University)
Marc Legacy (University of British Columbia)
Benji Leszcz (University of King's College, Halifax)
Brian Monteiro (University of Toronto)
Ernest Morrow (Dalhousie University)
Jenny Pelletier (McGill University)
Jeremy Proulx (Trent University)
Adrien Viens (University of Toronto)
Sasha Viminitz (University of Alberta)
Vicki Weafer (Mount Allison University)

Graduate Readers

Karen Bardsley (McGill University)
Kimberley Brownlee, co-founder (University of Cambridge)
Ryan Doran (Queen's University)
Thomas Garman (McGill University)
Matthew Grist (McGill University)
Alice MacLachlan (Queen's University)
Steve McKay (McGill University)
Tracy Nicholls (McGill University)
Jakeet Singh (University of Victoria)
Brian Van der Broek (McGill University)

Contents

Introduction	vii
Philosophical Passports: An Interview with Charles Taylor	11
Passeports philosophiques: Une entrevue avec Charles Taylor	23
G.A. Cohen's 'A Brief Look at Socialism': A Response by Jan Narveson	26
Alcibiades Redeemed: The Ascent to Divine Madness <i>Tara Tonja Mitchell, Trent University</i>	43
McDowell and Kim on Functionalism <i>Michael Brent, University of Toronto</i>	70
The Value of 'Authentic': A Defence of Authenticity in Music <i>Carl Joakim Gagnon, McGill University</i>	84
Amor Fati: Consciousness in Nietzsche's <i>Gay Science</i> <i>Mark Sentesy, Carleton University</i>	97
Decision-theoretic Consequentialism and the Nearest and Dearest Objection: A Critique <i>Jared Will, McGill University</i>	116
Contributors	132

Introduction

Since its inception three years ago, *Pensées* has provided a forum for Canadian undergraduate philosophy students both to publish their work and to take part in the assembly of a national journal through our national editorial board. In many regards, this year has been our most successful yet. In addition to receiving the most varied selection of papers that we have seen thus far, this year's editorial board was the largest and most diverse we have had. Essays were submitted from universities across Canada, and our editorial board comprised not only students from across the country, but also Canadian students studying abroad in Cambridge and St. Andrews, Scotland. We are fortunate this year to be able to include an interview with Canadian philosopher Charles Taylor (McGill University). Furthermore, Jan Narveson (University of Waterloo) has provided a provocative response to last year's interview with G.A. Cohen (Oxford). We are sure the reply will stir up heated debate in undergraduate circles. Finally, at long last, *Pensées* has established its own website at www.pensees.homestead.com, where there are articles from past volumes as well as information about how to submit to the journal or join its editorial board. We encourage all interested students to visit our new site.

This year marks a turning point in the life of *Pensées*. First, we saw the departure of co-founder and editor-in-chief Kimberley Brownlee, who has moved on to pursue graduate studies at the University of Cambridge. Kim's presence has been sorely missed, but was ably filled by Avi Cramer, who took over as co-editor-in-chief alongside Daniel Groll. Second, this year marks co-founder Daniel Groll's last as editor-in-chief, as he, too, will be pursuing graduate studies in parts unknown (at the time of writing) next year. This is, then, a year of transition that will continue into next, when the journal will be in the hands of Avi, and an as-yet undetermined colleague.

Dès le départ, l'un des buts de Pensées était d'être vraiment bilingue au moyen de documents français ainsi que par la participation de membres francophones à l'équipe de la

rédaction de McGill et d'autres universités. Malheureusement, malgré nos efforts continus, nous n'avons reçu qu'une seule soumission francophone et aucune requête concernant l'équipe de rédaction venant d'étudiants francophones avant la date buttoir du 15 janvier. Nous sommes encouragés, cependant, par l'intérêt de plusieurs étudiants d'universités françaises à soumettre et à réviser des documents, bien après la date buttoir. Nous espérons que cet intérêt va porter fruit et entraîner une collaboration pour le journal de l'an prochain. Malgré le fait qu'il n'y ait pas de documents français dans notre journal de cette année, nous avons inclus une traduction de l'entrevue avec Charles. Nous encourageons fortement tous les étudiants francophones des universités au travers du Canada à s'impliquer dans le journal, soit en soumettant un document ou en devenant un rédacteur. Pour plus d'information, veuillez visiter notre site internet au www.pensees.homestead.com.

Once again, the papers we've chosen reflect the high quality and diversity of undergraduate work in philosophy across Canada. In selecting essays, each submission was reviewed by three undergraduate editors, each of whom made extensive comments. On the basis of these assessments, the best papers were further reviewed by graduate students and professors. The five papers selected for publication were recognised for their clear exposition of ideas and coherent argument. Tara Tonja Mitchell's (Trent University) paper, 'Alcibiades Redeemed: The Ascent to Divine Madness,' offers an alternative approach to the character Alcibiades in Plato's *Symposium*, an approach that argues that Alcibiades is much more complex than standard interpretations suggest. In 'McDowell and Kim on Functionalism,' Michael Brent (University of Toronto) argues that, based on John McDowell's critique of functionalism, Jaegwon Kim's particular brand of functional reduction of the mind is untenable in the face of our intuitions about what it is to be human. Carl Joakim Gagnon (McGill University) argues in his paper, 'The Value of "Authentic": A Defense of Authenticity in Music,' that there is both a coherent conception of an authentic musical performance and that this kind of performance is something with distinct aesthetic value. The paper 'Amor Fati: Consciousness in Nietzsche's *Gay Science*,' by Mark Sentesy (Carleton University), offers an analysis of the role played by language, communication, and consciousness in Nietzsche's epistemology. Finally, in 'Decision-theoretic Consequentialism and the Nearest and Dearest Objection,' Jared Will (McGill University) argues that Frank Jackson's decision-theoretic approach to ethics fails to answer the "nearest and dearest objection."

We would like to thank everyone who helped make the third volume of *Pensées* a success. First, Charles Taylor, who took the time to sit down with us for an interview. We thoroughly enjoyed having the opportunity to discuss his ideas with him and are very grateful for his contribution to the journal. We would also like to thank Jan Narveson for offering to respond to some of G.A. Cohen's remarks from last year's interview. As always, we are very grateful for the support of our faculty supervisor, Professor Philip Buckley, as well as that of Professor Margarite Deslauriers, who pro-

vided helpful comments on one of the papers. There are many others whose help has been invaluable in preparing the journal for publication: Jenny Pelletier and Jeremy Proulx for editing two of the selected papers; Zach Dubinsky for copy-editing the final version; Dom Michaud for his great layout job; Serge Schiller at AGMV printing; and, of course, all of our graduate and undergraduate editors.

We invite students and faculty alike to send us any comments and suggestions you might have about the journal. We hope you find *Pensées*, Volume 3, a thought-provoking read.

— Avi Craimer,
Daniel Groll,
Gabriel Hébert
Montreal, 2002

Philosophical Passports: An Interview with Charles Taylor

Charles Taylor was for many years Chichele Professor of Moral Philosophy at the University of Oxford and a professor of political science and philosophy at McGill University. He is the author of numerous articles and many books, including *Hegel* and *Sources of the Self*.

The historical origins of ideas plays a prominent role in much of your work. This stands in distinction from the ahistorical, problem-based approach associated with the analytic tradition. What do you see as the current role of the "history of ideas" in philosophical scholarship and education and what role should it play?

I think that you can't really do certain kinds of philosophy without history. I want to say certain kinds, because philosophy doesn't really exist as a single discipline; it is really the study of the most fundamental problematic questions that arise in any study. So philosophy really consists of 'philosophy of this,' 'philosophy of that': philosophy of logic, philosophy of history and so on. What I'm saying doesn't necessarily apply to certain areas, like philosophy of logic or philosophy of mathematics, where, for all I know, you could do the contemporary work very well while forgetting about the history. But you can't do that in these more fundamental issues that have anything to do with human life, ethics, and understanding human beings. Because human beings are historical beings that change over time, you can only understand them in this kind of development. On the one hand, we have to correct someone like Aristotle; on the other hand, he's correcting us, because he reminds us that our current concerns were understood totally differently earlier on and *that* gets us unstuck from our *very* parochial modern Western European understanding. This whole swatch of areas, which are most of the 'philosophy ofs,' are of a kind that you can't really do without understanding, even in a broad sense, the history of how human beings developed and *certainly* the history of how thought about a particular 'philosophy of' developed.

I can imagine someone playing the devil's advocate might say: Suppose we have this problem, let's say free will, which has been around for thousands of years and today it's the same problem. Why do I need to know the development of the 'philosophy of' in this case? Why can't I just choose to engage with the problem on its own terms?

There isn't an ahistorical perspective on *that*, you see, because translate 'freedom of the will' into Greek for me — what's 'the will' in Greek? Maybe you can massage out some Aristotelian contribution to this, but to do this is to look at the problem in such a different way. And why do we have to look at it in that way? Well, because our way of looking at it isn't necessarily the final, last, irrevocable word in human history.

There is a distinction in your work between the contingent things that change over time and what's more basic in human psychology and intellect. How do we distinguish the parts of our thought that are real contributions to philosophy and what is simply a will o' the wisp?

It's a very pessimistic answer, you might think, but I don't think we will ever get a definitive answer to that, because there's no way we could ever be at that final point. There are two ways in which we could be at that point. If, contrary to what I believe, we had the kind of science of human nature that could be grounded in something like natural science, a socio-biological account — well, not even socio-biological, because that involves just-so stories about the evolutionary period, but let's say a straight biological account — then you would have a way of getting at what the basic human instincts are by reading them off your chart of the neurophysiology of the human being and applying a reductive explanation. You could then get the perennial human nature that is buried. But I don't think we will ever get to that. The other thing would involve being at the end of history, where we could survey the whole history behind us and say, "these are all the variations," and then find what the constant is that explains why there are these variations. It's almost incoherent to think that we could be at either of those points, so we're going to be constantly revising what our ancestors said about this. I don't think that that means that we can't make progress, because I think there is a kind of progress you make where you correct errors, even though this corrected error is not the end of the story. It can in turn be corrected. We mentioned Aristotle earlier and it's clear that we can correct the kind of ahistorical picture he has of human nature. We can get a sort of idea today of what is perennial and what changes, but there is no way we can be sure that that's not going to be heavily revised by our successors. So, in a way, it's a little bit pessimistic if you're looking for a definitive answer.

You mentioned the two-pronged approach you might take in coming up with these perennial issues. The first was some sort of pure biological account, which you said was "almost incoherent." I take it what you mean by that is that it's not simply a question of the gross impracticality of trying to unpack the biology of it, but that there is some conceptual reason why we'll never be able to reduce one to the other.

I can't, as it were, give the argument with confidence now, but I can imagine being at a point where we could see what the form of the argument against a reductionism of this sort could be. Take the attempts to explain consciousness in books like

Consciousness Explained: I think Dennett doesn't succeed in getting in contact with the real phenomenon. Think of how completely at sea we are when we ask the interesting question, "Why is it that consciousness, feeling, and so on arises in or out of organic beings like this [slapping hand]?" which I think people like Dennett can't ask because they're in another conceptual space. Of course, this immediately has to start off with the realisation that these various things made of transistors (such as computers) aren't even in the same league. We just have no idea of why this supervenes on that, in the way that we have when we're dealing with the mechanic energy of molecules and felt heat. And it may be that we'll find that there's a real dilemma here, that is, that in order to get to the point of explaining what people do consciously and intentionally you have to enrich whatever biology you're doing by introducing a historical element, and then you're right back into it. We'll see, that's just my opinion.

You mentioned that we should enrich the vocabulary of science, as it were, and this presumably involves striking a balance between reductive physicalism and the richer but implausible stories told by modern-day Cartesians. Owen Flanagan offers one way of doing this when he claims that the only thing the physicalist really needs is some purely physical account for why our sensations can't be accessed by other people and so can't be reduced and 'objectified.' He thinks that a simple account of how our brain hooks up with the rest of our body can accomplish this quite easily.ⁱⁱ What's your response to that approach?

But that's a ludicrously tiny bit of the phenomenon, the bit of it that involves my having access to my thoughts in a way that you don't. There's also everything else in it: feeling, pain, desire, intention — there are all these things and there is also the phenomenology of them. This first-person access is one tiny bit...and I challenge someone to be able to show that even that can be exhaustively explained in the current vocabulary of natural science. Without being able to show the whole phenomenon, the privileged access of *what* is still unclear. So these attempts, such as Flanagan's, are just stabs in the dark. I mustn't put too much weight on this [because] they're all promissory notes. But where I think I have a point against these people is that they don't seem to have taken aboard what an awesome, difficult project this is. It comes back to what Aristotle said, and Husserl after Aristotle and so on: philosophy starts in *thaumazein*, 'in wonder.' There's something really remarkable about the background of the world understood in terms of modern physics and there's something very remarkable about that whole panoply of phenomena, such as consciousness, feeling, and so on. There just is not the very beginning of an intimation of how that supervenes on this kind of stuff [slapping hand]. And the proof of it is that almost all these people are tremendously excited about artificial intelligence, without seeing that it's not really the same phenomenon at all. They're trained systematically to miss how extraordinary a phenomenon consciousness is, and therefore never to face the issue of how far we are from explaining it.

If I can play the devil's advocate again: Suppose I'm an eliminative materialist or something of the sort. It seems to me that your position commits you to some sort of occultism. If it is not a materialist account, what could it possibly be? What are these terms — intentionality, belief, desire — that we can't explain further?

I think the word materialism really does spread immense confusion here, because if it means some account of how these material things are also thinking things, that's exactly what I'm talking about getting at. But if it means, which it usually does, some account of this with the following restrictions: that it will fit *exactly* the categories of natural science as we now know them, closed brackets; well, if *that's* materialism, then there's no reason to believe in it. Materialists are equivocating on the word — so the word can really be tossed out the window, and we'll be back in the real debate.

You make that point in your essay 'How is Mechanism Conceivableⁱⁱⁱ': we shouldn't assume that the language and the categories of science today are the ones we're going to be working with forever. It may well be that these categories will expand — some will be dropped, others will be added — and the categories that will accommodate things like intentionality will be added, I take it?

Yes, something like that. Imagine a sort of Gerald Edelman-type theory pushed farther: how do you account for the fact that a certain kind of ontogenesis in a Japanese family produces something different in many respects from a similar ontogenesis in a Canadian or Norwegian family? So how are you actually going to relate these cultural differences to ontogenesis characterised in terms of what Edelman's doing, and so on. I don't see how, in the end, you're going to be able to do it, without beginning to introduce some of the terms describing these differences into it. I must stress, though, that what I'm doing now is making certain conjectures. But what the conjectures are meant to point out is how little we really know. That these conjectures can be made shows that the gamut of possibilities is much wider than people now think.

Your non-reductive philosophy of mind also seems to be mirrored in a non-reductive ethics, connected by the rich notion of 'strong evaluation,' which is understood as the tendency to have qualitative desires about other desires, condemning or approving of them from the point of view of worth. If human feeling can't be reduced in philosophy of mind, our 'fundamental' preferences can't be reduced in ethics either. Could you briefly explain how strong evaluation functions in your ethics?

Yes, before functioning in ethics, it's something general about human beings, it goes beyond the realm of ethics — maybe this is not a totally coherent category in the end, I'm never quite sure [laughs]. But let me try this: there are certain things that we value

more than others. I like strawberry ice cream better than chocolate, or something like that. And then, there are valuations where in a certain sense we value the valuation, and this can be shown by asking, "With what kind of feeling would you confront the possibility that you might lose this valuation?" And here, the kind of feeling with which you would confront it would also, if it were a strong evaluation, relate to the area of contempt, admiration, sense of greatness, sense of littleness, and so on. So you see that if somebody says, "Imagine that you cease valuing strawberry ice cream and you suddenly like chocolate more," unless I have certain sentimental reasons why I like strawberry ice cream more, I would say, "Okay, that's just the way it is." But if you say, "Imagine that you became a follower of al-Qaida tomorrow and thought that killing a lot of civilians was a fine thing," I would be horrified. I would be morally horrified at myself; I would feel I was going downhill. All of the things that recruit these senses of admiration, contempt, awe, etc., enter here in making these valuations and that is what distinguishes strong evaluation from mere taste.

It goes beyond this, however. What is the boundary of the moral or the ethical? It is notorious that people have the kinds of feelings like awe, contempt, etc., about things that we might call aesthetic. Let's say that I'm Baudelaire's Dandy and that it's predicted that I'll become completely indifferent to dressing and standing smartly and will walk around in a slovenly fashion in ten year's time. In this case, I'm equally horrified. Now, following some definitions of 'ethical,' this is the same reaction as above, following others it isn't. So this is an extra problem here; how we draw the actual boundary of the moral is hotly contested. But the phenomenon of strong evaluation goes potentially beyond the boundary. That is, it ranges over a whole range of things, some of which are inside some drawings of the boundary of the ethical, some of which are outside of it.

So within the boundary of the ethical, we think that the question, "Do you like strawberry or chocolate ice cream?" isn't really worth discussion. However, if you were to join al-Qaida we would be horrified. Were you to do this, though, there would be a community somewhere that would celebrate and say, "Finally! You got it right." We want to condemn that attitude, but how? Clearly, we want to say that some desires or values are more desirable than others in an ethical sense, but what mechanisms do we use to adjudicate between them when we've got these conflicting communities?

We've shifted, rightly so, but shifted nonetheless, in the sense that the thesis that the language of strong evaluation has bite is quite independent of whether or not one would adopt a totally relativistic view of values. Given that, I think that we find ourselves capable — not in all cases of conflict, but in some cases — of arguing them out. I've tried to describe this kind of thing in a paper like 'Explanation and Practical Reason.'^{iv} Again, we cannot ever move to the position where we can confidently

believe that this is the ultimate, final, unbridgeable position. But we can move in the sense of showing that *this* is superior to *that*, even if this is not to say that we are guaranteed that this is always going to be the case: we can also come across horrifying differences, yawning gaps where we couldn't do anything like this. These differences could be *de facto* or *de jure* unbridgeable; *de facto* in the sense that some people could be so far gone that they don't begin to consider your position as in any way valid.... Imagine sitting down long enough with Bin Laden to talk it out. It's not possible, he'd just shout you down. Which doesn't mean to say that with a more reflective, tractable follower you couldn't sit down. So we may never know whether some differences are in principle unbridgeable, because the *de facto* unbridgeability takes a form that cuts off the possibility of really trying it out.

But we do know cases where you can find such *de facto* unbridgeable arguments, so what attitude ought we to adopt? And here I think that the current philosophical theories are all in the wrong place, because the attempt is to show somehow *a priori* that you can't bridge the differences or *a priori* that you can. But how could this ever be a valid argument? It's only by actually trying it out that you could know. So my point — my disagreement with just about everybody else in analytic philosophy at the moment — is that I think that the issue is being misplaced somewhere. And I think the issue is being misplaced because of the baleful power of a certain epistemological view that I call mediational epistemology, where people are supposed to have a set of final assumptions over which they screen whatever comes in, and these assumptions include the area of value. That being so, when you hit on someone with a different assumption, nothing can be said. This kind of *a priori* proof of the unbridgeable nature of moral or cultural differences doesn't work, but this doesn't mean that we have a guarantee that the opposite is the case, that all differences are *a priori* bridgeable. We may come across gaps that are so great that there's no bridging them.

If there was some way, as you suggested, of talking out the differences, is this because we are members of a universal community, a human race? That is, are there some basic values that could be read off our common nature in some sense, as Aquinas tried?

In some sense, but what would that have to mean to us today? This is why I talk about correcting our predecessors to a certain degree. It's clear that the boundaries of nature are not as absolutely or neatly marked as Aquinas, following Aristotle, thought they were. So it's certainly a fact that there is some human nature, one of the features of which is the capacity and the necessity to give some kind of construal to human life through culture. But you can see already from that that it's not going to get us all the way. It gets us some part of the way, because we know that this necessity is connected to us as language-beings, which enables us to have interesting arguments about how you ought to treat human beings in virtue of their being language-beings. And this is

how I reconstrue writers like Habermas: I think that they're actually operating in a way that ought to be understood in a much more traditional fashion. But this doesn't cover the whole gamut: it *doesn't* cover the issues of what a good or meaningful life is. And there we also recognise, since Herder, that while there can be a wide gamut of different approaches that don't have to be ranked, some do, and in order to work that out we have to look at the whole development and variation of human beings. So it's not like the kind of thinking that you have with Aquinas, where the whole issue of the different cultural workings of human life doesn't come on the horizon at *all*.

In order to work out that sort of ranking, some commentators like Sir Isaiah Berlin, who have otherwise been very sympathetic to your philosophy, think you might be committing yourself to an untenable teleological account of human values. What's your comment on that?

I think that's another bad take on me, because of this pervasive philosophical culture of *a priorism* around us. Everybody is read as having an *a priori*, and so that's one of the possible *a prioris* you could attribute to me: I'm coming at human life with this heavily teleological account of where it's all heading. This philosophical culture can't somehow allow for the fact that, no, we don't know all of these things. I like Isaiah Berlin's business about conflict between goods, unless it gets hardened into an *a priori*, when we say that we can know today that we'll *never* be able to reconcile them...just a minute, just a minute. If the point is the very plausible and heavily backed empirical conjecture that there's always going to be some set of very-difficult-to-combine values or goods in human life wherever you are in human history, yes, that's as likely as that the sun will rise tomorrow. But we never know whether we can no longer struggle towards a much better combination of A and B. We're now struggling with combinations of radical free agency and fully meaningful lives, and this is very hard to combine. That's one of the big oppositions between possible goods in our day. You wouldn't say that about the Greeks, but you would say that about us. Who can say how much headway can be made on this? So what I don't like about the way Berlin's position often is played out, including by Berlin, is that it's hardened into an *a priori*. And my position of saying "Just a minute" is interpreted as saying, "Of course they can be combined. How can I believe that? Well, I must have a teleology." This is always happening, it's very, very funny to see. It's a philosophical culture that doesn't allow for another possibility and therefore always wants to commit people to some *a priori*. You have to have your *a priori*. It's like you have to have your passport: "Where's your passport? Your Aristotle passport, your Hume passport!" No, no, no! I'm stateless in that world!

In reference to today's values, which you just hit upon, one of the major symptoms that you identify in the modern era, most pointedly I think in Sources of the Self⁶, is this idea of

the emergence of both atomism and what you call facile relativism (exhibited in popular philosophies like 'I can do what I want'). With developments of the past decade that include the impending formation of an International Criminal Court, foreign intervention in places like East Timor, the unprecedented cooperation in the wake of September 11, as well as the protest movements against globalisation, the International Monetary Fund, the World Bank, etc., would you say that there has been a move away from the atomistic and relativistic attitudes that you described several years ago?

No, I think they've always coexisted with these other types of movements, so to the extent that there is oscillation, it's oscillation in the life of one individual, or it can be oscillation, perhaps, in the life of cohorts in a given society. People say, "The students are all into this now" — but that's a certain amount of journalistic exaggeration even in those oscillations. You don't need to read any books, you just need to live as long as I have [laughter]. When you see the fashions move like that, you get suspicious of just how much credence to put into any of these.

These developments relate to the increasing talk of human rights in particular. You have suggested that talk of rights only makes sense within a social setting in which they are coupled with obligations. In light of the movements discussed above, do you think there has been a development towards the obligation-side of the rights-obligations couplet?

Well, in one sense, rights and obligations are logically tied together, so that's not the issue: if you have a right to life, I have an obligation not to take it. But when people like Amitai Etzioni in the States talk about how we have too much emphasis on rights and not enough on obligations, then they're thinking something like, "We have too much emphasis on what I can demand from society, and not enough emphasis on what I have to give to society." As a general proposition, I'm not against this, but I hate remaining on this level of generality because there are lots of things people could put forward as obligations that I wouldn't really accept. And I prefer to put it in terms of understanding how the very nature of democracy requires a certain degree of commitment from the people who constitute a people in the democratic sense. That is, they're self-governing together, which requires a certain degree of commitment to the process, to the project that they have together and, therefore, to other people who are in the project. When that weakens you can see the terrible consequences. For instance, when the sense of solidarity weakens, you get very deep resentments and a sense of alienation and so on. So, I prefer to follow up that line of thinking wherever it leads, and it can lead, among other things, to a sense of obligation, too.

If, philosophically speaking, rights and obligations are in a sense logically tied together, the challenge then is making philosophers and people in general realise that it doesn't make sense to give priority to rights talk, as we see in a lot of thought that posits the right as the sort of

fundamental unit of the human, and only then have talk about society follow. I take it your challenge is to bring to light the notion that rights and obligations can only take place within a particular setting.

That's right, and so the weight of the complementarity — which, I think, is interesting to bring out — is the complementarity between various kinds of affirmation of individual rights, or individualism in various different modes — expressive individualism, instrumental individualism, and so on, on one hand — and the modern forms of social bonding that have grown up along with them, one of the most important being the self-governance of the people. Without the latter, simply affirming the forms of individualism would be tremendously unstable and even self-destructive, self-undercutting. Once you realise this, belonging to those forms of social bonding becomes a very important *desideratum* of life in our society, because without them you cannot be an individual. I'm just saying what Tocqueville and others have said. Again, I think the way the discussion has been framed in the current philosophical environment is very highly *a priori*. At some point you want to say, "Stop talking about who has a right to what," and start talking about minimal social prudence, like what will allow us to have a healthy democratic society in which people's rights are respected.

Along the same lines, a lot of your work emphasises the relationship between the individual and the community and how the latter largely informs the identity of the former. Recent events have brought to the fore the complex web of allegiances constituting people's identities, highlighting the tensions that exist not only within communities, but within a single individual, who often has a multi-faceted set of allegiances upon which her identity is based. Given these differences, on what basis is a community a community? Is it in virtue of living together, having common struggles? How should we understand these allegiances both on an individual and communal level? Do some take priority over others? Should some take priority over others?

I don't think you can give a general answer to the last question because situations vary. Even regarding the first, the generality of my answer is going to be very historically restricted. I think if we want to answer this in the context of the modern democracies we know in the North Atlantic area today, the nature of the bond is clear. You can start off with something like Habermas's 'constitutional patriotism' idea, but give it a certain twist: they're all cases in which the bonds people have must be, and usually are, a mixture of certain principles (democracy, etc.) on the one hand, and a strong commitment to a certain historical project of the realisation of these principles on the other. I say 'historical' because the particular project is only identified really by its coming from somewhere, by its having a history. It's not identified only by having a certain geographical location, because there are countries, like Poland, that have been totally displaced, or almost totally displaced. So it's not the area. Rather, it's the histo-

ry, the people and the constantly changing population. This is what must be generated. If people don't have that then there isn't the kind of bonding that can allow for a healthy democracy. Of course, how you develop that is where the whole issue of multiculturalism arises. In all these cases there has to be a consciousness of the multicultural dimension of these societies. If you're going to bond all these people, the historical project, of let's say Canadian democracy, has to have that dimension as part of its identity, of what it understands itself to be.

We've fought this out in Quebec. We had a big problem, and I think the problem is now receding, thank God. The famous remarks by Parizeau about "l'argent et le vote ethnique" show that for some Quebecois nationalists, the project was defined almost purely in these *pure laine* terms. Almost everybody now recognises that the type of population we have now in Quebec — this was the case before, but it's even more so now — makes that an impossible basis for bonding around this historic project. If people are bonded to this project, they of course understand, and accept, that it had *pure laine* origins. But that's not what it is today, and it's not where it's going. More or less all the democratic societies currently face the task of redefining their identities, which some are doing well, others badly.

I take it that it's a sign of a healthy democracy that you have, on the one hand, these competing allegiances and, on the other, people who recognise that there are principles and a framework within which to sort them out. That is to say that we understand that there's a limit to the way you can express your opinions: non-violence, peaceful protest, or even referenda, which are a very peaceful way of sorting out what some people view as cultural or even ethnic conflict, if we can put it that strongly.

Yes, absolutely, yes.

These allegiances are of course very fundamental to your moral thought in the same sense that they are fundamental to your philosophy of mind, because they make up a lot of what it is to be a person. And that raises the question we touched on before with renewed strength: what do we actually do when fundamental ways of being clash? You've mentioned "talking it out," but because this seems to be such a fundamental issue, people have, as you've said, wanted a stronger set of guidelines within which to answer such questions.

Well, there's not an answer to this question, certainly not a philosophical answer and maybe not an answer *überhaupt*. That is, there can be a situation where you fail, where you can't have an ordered democracy. Supposing some great international force declared tomorrow that all the inhabitants of Palestine, Israeli and Arab, must form a single democracy by tomorrow. There's a condition of virtual civil war. You don't change people like that overnight. Democracy breaks down, because there's too much

violence... We hope this won't happen to our 'well-ordered' North-Atlantic democracies, but it could at any point. As you said earlier, there has to be this agreement on certain limits, which belongs to that first part of the constitutional patriotism and the set of principles. If there is really a deep rejection of some of those, then there isn't a solution. So it's not just that philosophy can't produce an answer. Sometimes there's *nothing* that can produce an answer. We know what a solution would look like, and we can struggle as hard as we can to bring people back to that consensus, but we could fail. History is full of agonizing failures.

To bring the interview full circle, we started out talking about your very historical approach to problems. One of the currents through this interview has been what you see as the strong a priori and ahistorical approach philosophy has to many problems today. I take it those two are very strongly related. What I've gathered from what you have said is that when it comes to moral theory, or talking about communities, or the possibility for democracy and peace, it's very much a question of reading what the possibilities for the future are, in a particular situation, off of history, so to speak, and not simply taking these problems in abstract and coming up with principles that say, "This is the case. This isn't the case."

I think that the Canadian imbroglio has been worsened by this second kind of thinking. What we have is something like a binational or multinational state, but some people have been blocking the constitutional recognition of this, because it seems logically wrong to them. They think that there has to be uniformity of citizens' relation to the state, which by definition is not the case in a multinational society. And so we've had all this blocking on Quebec as a distinct society in other parts of the country, very often based on heavily philosophical views, such as a very uniformitarian interpretation of the *Charter of Rights*. The Meech Lake Accord supposedly wouldn't work, because there might be different interpretations to certain clauses of the charter in different regions. And so they reject out of hand as being impossible or unconscionable the only solution that would work in our country; because if you have a really multinational society you've got to have some kind of recognition of that fact. And so I think bad philosophy has done a lot of political harm, and I feel very strongly about that because it's my country that's been badly damaged here.

Notes

- ⁱ Daniel C. Dennett, *Consciousness Explained* (Toronto: Little, Brown, 1991).
ⁱⁱ Owen Flanagan, *Consciousness Reconsidered* (Cambridge: MIT Press, 1992), 87-107.
ⁱⁱⁱ Charles Taylor, *Human Agency and Language: Philosophical Papers 1* (New York: Cambridge University Press, 1986), 164-186.
^{iv} Charles Taylor, *Philosophical Arguments* (Cambridge: Harvard University Press, 1997), 34-60.
^v Charles Taylor, *Sources of the Self: The Making of Modern Identity* (Cambridge: Harvard University Press, 1989), 111-207.

Passeports philosophiques: une entrevue avec Charles Taylor

Charles Taylor a été pendant plusieurs années Professeur de Philosophie Morale à Oxford et professeur de science politique et de philosophie à l'Université McGill. Il est l'auteur d'innombrables articles et de nombreux livres incluant *Hegel* ainsi que *Sources of the Self*.

Les origines historiques des idées jouent un rôle prédominant dans la plupart de votre œuvre. Ceci se distingue de l'approche ahistorique basée sur le problème associée à la tradition analytique. Que voyez-vous comme le présent rôle de 'l'histoire des idées' en rapport au savoir de la philosophie et à l'éducation et quel rôle doit-elle jouer?

Je pense que vous ne pouvez pas faire certains types de philosophie sans histoire. Je veux dire certains genres, car la philosophie n'existe pas sous la forme d'une seule discipline; c'est vraiment l'étude des questions problématiques les plus fondamentales qui surgissent dans toute étude. Donc, la philosophie consiste vraiment en 'la philosophie de ceci,' 'la philosophie de cela,' la philosophie de la logique, la philosophie de l'histoire et ainsi de suite. Ce que je dis ne s'applique pas nécessairement à certains domaines, comme la philosophie de la logique ou la philosophie des mathématiques, où, d'après ce que je sais, vous pouvez très bien faire le travail d'aujourd'hui tout en oubliant l'histoire. Mais vous ne pouvez pas faire cela avec les enjeux plus fondamentaux qui ont quelque chose en rapport à la vie humaine, à l'éthique et à la compréhension des êtres humains. Parce que les êtres humains sont des êtres possédant une dimension historique qui change avec le temps, vous ne pouvez les comprendre qu'en considérant ce genre de développement. D'un côté, nous devons corriger quelqu'un comme Aristote; de l'autre côté, c'est lui qui nous corrige en nous rappelant que nos préoccupations actuelles étaient comprises de façon complètement différente autrefois et *cela* nous détache de notre très paroissiale compréhension moderne occidentale. Cet ensemble d'échantillons de domaines, qui sont la plupart des 'philosophie de' sont d'un type que vous ne pouvez pas vraiment faire sans la compréhension, même au sens large, de l'histoire du développement des êtres humains et *certainement* pas sans l'histoire de la façon dont la pensée de la philosophie d'un sujet particulier s'est développée.

Je peux imaginer quelqu'un jouant l'avocat du diable dire: Supposons que nous avons ce problème, disons le libre arbitre, qui a été ressassé pendant des milliers d'années et aujourd'hui, le problème est toujours le même. Pourquoi dois-je connaître le développement de la philosophie de dans ce cas? Pourquoi ne puis-je pas simplement choisir d'engager le problème sous ses propres termes?

Il n'y a pas de perspective ahistorique sur cela, car traduisez 'libre arbitre' en grec pour moi — que signifie 'arbitre' en grec? Peut-être vous pouvez en ressortir une contribution aristotélique, mais faire ainsi revient à regarder le problème sous un angle si différent. Et pourquoi devons-nous regarder le problème sous cet angle? Bien, parce que notre façon de regarder le problème ne correspond pas nécessairement à la dernière définition irrévocable du mot dans l'histoire de l'humanité.

Il y a une distinction dans votre œuvre entre les choses contingentes qui changent avec le temps et, ce qui est plus fondamental dans la psychologie et l'intellect humain. Comment distinguons-nous les parties de nos pensées qui sont de réelles contributions à la philosophie et celles qui ne le sont pas?

C'est une réponse très pessimiste, vous vous dites, mais je ne crois pas que nous allons un jour avoir une réponse définitive à cette question, car il n'y a aucun moyen pour nous d'atteindre ce point final. Il y a deux moyens par lesquels nous pouvons atteindre ce point. Si, contrairement à ce que je crois, nous possédons le type de science de la nature humaine qui pourrait être déterminée comme une science de la nature, un compte socio-biologique — en fait, même pas socio-biologique, car cela implique des histoires à propos de la période évolutionnaire, mais disons un compte seulement biologique — alors vous auriez un moyen d'arriver à savoir ce en quoi consiste les instincts humains fondamentaux en les lisants sur votre charte de la neurophysiologie de l'être humain et en appliquant une explication réductionniste. Vous pourriez ensuite obtenir la nature humaine perpétuelle qui est ensevelie. Mais je ne pense pas que nous allons nous rendre à ce point. L'autre chose impliquerait de se retrouver à la fin de l'histoire, où nous pourrions analyser toute l'histoire en arrière de nous et dire, "Voici toutes les variations," et ensuite trouver la constante qui explique pourquoi il y a ces variations. C'est presque incohérent de penser que nous pourrions être à l'un de ces deux points, alors nous allons constamment réviser ce que nos ancêtres ont dit à propos de ceci ou de cela. Je ne pense pas que cela veuille dire que nous ne pouvons pas faire de progrès, car je crois qu'il y a un type de progrès que vous faites lorsque vous corrigez les erreurs, même si l'erreur corrigée n'est pas la fin de l'histoire. Elle peut à son tour être corrigée. Nous avons mentionné Aristote plus tôt et il est clair que nous pouvons corriger le type de dessin ahistorique qu'il a de la nature humaine. Nous pouvons avoir une idée aujourd'hui de ce qui est perpétuel et de ce qui change, mais il n'y

a aucun moyen pour nous de s'assurer que ce ne sera pas lourdement révisé par nos successeurs. Donc, d'une façon, c'est un peu pessimiste si vous cherchez une réponse définitive.

Vous avez mentionné la double approche que vous allez possiblement inclure avec les enjeux perpétuels. La première était un genre de compte purement biologique dont vous avez dit qu'il était "presque incohérent". Je prends en considération que vous signifiez par ceci que ce n'est pas simplement une question de manque de réalisme d'essayer de déchiffrer toute la biologie impliquée, mais qu'il y a une raison conceptuelle pour laquelle nous ne serons jamais capable de réduire l'un à l'autre.

Je ne peux, comme il était, argumenter avec confiance maintenant, mais je peux imaginer être à un point où nous pourrions voir comment serait la forme de l'argumentation contre le réductionnisme de ce type. Essayer d'expliquer la conscience dans des livres tel que *Consciousness Explained*: Je pense que Dennett ne réussit pas à entrer en contact avec le vrai phénomène. Pensez à quel point nous sommes perdus quand nous posons l'intéressante question: «Pourquoi est-ce que la conscience, les émotions et ainsi de suite émanent des êtres organiques comme ceux-ci [claquement de mains]?» à laquelle je pense que des personnes comme Dennett ne peuvent poser car ils sont dans un autre espace conceptuel. Évidemment, ceci doit immédiatement débiter avec la réalisation que ces diverses choses faites de transistors, (comme les ordinateurs) ne sont même pas de la même ligue. Nous n'avons simplement pas idée pourquoi ceci intervient sur cela, de la façon que nous avons lorsque nous avons à faire à l'énergie mécanique des molécules et à la chaleur ressentie. Et il se peut que nous trouvions qu'il y ait là un vrai dilemme, ceci dit, c'est dans le but d'arriver au point d'expliquer ce que les gens font consciemment et intentionnellement, vous devez enrichir toute biologie utilisée en introduisant un élément historique, et ensuite, vous y revenez directement. Nous allons voir, c'est seulement mon opinion.

Vous avez mentionné que nous devrions enrichir le vocabulaire de la science, et ceci implique vraisemblablement de troubler l'équilibre entre le physicalisme réductionniste et les histoires plus riches mais peu plausibles racontées par les Cartésiens des temps modernes. Owen Flanagan offre une façon de faire cela lorsqu'il clame que la seule chose dont les physicalistes ont vraiment de besoin est un compte purement physique, ce qui expliquerait pourquoi nos sensations ne peuvent être accédées par d'autres personnes et alors ne peuvent être réduites et 'objectivées'. Il pense qu'un simple compte expliquant la façon dont notre cerveau est connecté avec le reste de notre corps peut accomplir ceci très facilement⁶¹. Quelle est votre réponse à cette approche?

Mais c'est une partie ridiculement minuscule du phénomène, la partie qui implique

que je puisse accéder à mes pensées d'une façon que vous ne pouvez pas. Il y a tout le reste dans cela : sensation, douleur, désir, intention – il y a toutes ces choses et il y a aussi la phénoménologie de ces derniers. L'accès de cette première personne est une minuscule partie... et je défie quelqu'un d'être capable de démontrer que même cela peut être expliqué de façon exhaustive avec le vocabulaire contemporain des sciences de la nature! Sans être capable de montrer le phénomène en entier, l'accès privilégié aux pensées demeure ambiguë. Alors ces tentatives, comme celle de Flanagan, sont seulement des coups de poignards dans le noir. Je ne dois pas mettre trop de poids sur cela car ils sont tous des billets à ordre. Mais où je crois avoir un point contre ces gens repose sur le fait qu'ils ne semblent pas avoir réalisé à quel point ce projet est difficile et imposant. Cela revient à ce que disait Aristote, et Husserl après Aristote et ainsi de suite : la philosophie commence en *thaumazein*, 'en merveille'. Il y a quelque chose de vraiment remarquable à propos du contexte de la compréhension du monde en termes de physique moderne et il y a quelque chose de très remarquable à propos de toute la panoplie de phénomènes, tel que la conscience, les sensations et ainsi de suite. Il n'y a pas le commencement d'une annonce de comment cela interviendrait sur ce genre de choses [claquement de mains]. Et la preuve est que presque toutes ces personnes sont extrêmement excitées par l'intelligence artificielle, sans voir que ce n'est pas du tout le même phénomène. Ils sont entraînés systématiquement à ne pas s'apercevoir du caractère extraordinaire du phénomène de la conscience et de plus, à ne jamais faire face à l'enjeu qui est à quel point nous sommes loin de pouvoir l'expliquer.

Si je peux encore jouer l'avocat du diable: supposez que je sois un matérialiste éliminatif ou quelque chose du genre. Il me semble que votre position vous engage en une sorte d'occultisme. Si ce n'est pas un compte matérialiste, qu'est-ce que cela peut bien être? Quel sont ces termes – intentionnellement, croyance, désir – que nous ne pouvons pas expliquer plus en détail?

Je pense que le mot matérialisme répand ici une confusion immense, car si il signifie un compte de comment ces choses matérielles sont aussi des choses pensantes, c'est exactement ce à quoi je veux en venir. Mais si cela signifie, ce qui est habituellement le cas, le même compte avec les restrictions suivantes : qu'il va correspondre *exactement* aux catégories des sciences de la nature tel que nous les connaissons, fermer la parenthèse. Donc, si *cela* est du matérialisme, alors il n'y a pas de raison d'y croire. Les matérialistes sont équivoques sur le mot – alors le mot peut vraiment être jeté par la fenêtre, et nous serons de retour dans le vrai débat.

Vous faites cette remarque dans votre essai : 'How is Mechanism Conceivable¹¹': nous ne devrions pas assumer que le langage et les catégories de la science d'aujourd'hui sont ceux avec lesquels nous allons travailler éternellement. Il se peut bien que ces catégories prennent

de l'expansion – certaines seront laissées de côté, d'autres vont être ajoutées – et les catégories qui vont accommoder les choses tel que l'intentionnalité vont être ajoutés, est-ce que j'ai saisi?

Oui, quelque chose du genre. Imagine un genre de théorie de type Gerald-Edelman poussée plus loin: comment tenir compte du fait que certains genres d'ontogenèse dans une famille japonaise produise quelque chose de différent sous plusieurs aspects à un ontogenèse similaire dans une famille canadienne ou norvégienne. Alors comment allez-vous relier ces différences culturelles à l'ontogenèse caractérisée en terme de ce que fait Edelman et ainsi de suite. Je ne vois pas comment, à la fin, vous allez être capable de réaliser cela, sans commencer à y introduire certains des termes décrivant les différences. Je dois cependant préciser que je fais présentement certaines conjectures. Mais le but de ces conjectures est d'indiquer à quel point nous en savons peu. Que ces conjectures puissent être faites montre que la gamme de possibilités est beaucoup plus vaste que les gens pensent de nos jours.

Votre philosophie non réductionniste de l'esprit semble aussi être miroitée dans une éthique non réductionniste, connectée par la riche notion de 'forte évaluation', qui est comprise comme la tendance à avoir des désirs qualitatifs à propos d'autres désirs, les condamnant ou les approuvant du point de vue de la valeur. Si les émotions humaines ne peuvent pas être réduites en une philosophie de l'esprit, nos préférences fondamentales ne peuvent l'être en éthique. Pouvez-vous brièvement expliquer quel est l'importance des fonctions d'évaluation dans votre éthique?

Oui, avant de fonctionner en éthique, c'est quelque chose de général à propos des êtres humains, cela surpasse la portée de l'éthique – peut-être que ce n'est pas une catégorie totalement cohérente à la fin, je ne suis jamais certain [rires]. Laissez-moi essayer ceci : il y a certaines choses auxquelles nous accordons plus de valeur qu'à d'autres. Je préfère la crème glacée aux fraises à la crème glacée au chocolat ou quelque chose du genre. Et ensuite, il y a des évaluations que, dans un certain sens, nous évaluons l'évaluation, et ceci peut être démontré en se demandant, «Quel genre de sensation confronteriez-vous dans l'éventualité que vous puissiez être à tort dans cette évaluation?» Et ici, le genre de sensation avec laquelle vous seriez confronté serait aussi relié, si c'était une forte évaluation, au domaine du mépris, de l'admiration, du sentiment de grandeur, du sentiment de petitesse et ainsi de suite. Donc vous voyez que si quelqu'un dit, «Imaginez que vous cessiez d'accorder de la valeur à la crème glacée aux fraises et que vous préféreriez subitement le chocolat,» à moins que j'aie certaines raisons sentimentales pour lesquelles je préfère la crème glacée aux fraises, je dirais, «Ok, les choses sont ainsi faites.» Mais si vous dites, «Imaginez que demain vous deveniez un suivant de al-Qaida et que vous pensiez que tuer beaucoup de civils est une bonne chose,» je serais horri-

fié. Je serais moralement horrifié envers moi-même; je me sentirais descendre en enfer. De toutes les choses qui génèrent ces sentiments d'admiration, de mépris, de crainte etc. entrent ici en ligne de compte pour faire ces évaluations et ceci est ce qui distingue les fortes évaluations des simples préférences.

Cela va beaucoup plus loin, cependant. Quelle est la frontière de la morale ou de l'éthique? Il est notoire que les gens aient des sensations comme la crainte, le mépris, etc. à propos des choses que nous pouvons appeler esthétique. Disons que je sois un Dandy de Baudelaire et qu'il est prévisible que je devienne complètement indifférent à ma posture et à ma tenue vestimentaire et que je promènerai de façon très sale et négligée d'ici dix ans. Dans ce cas, je suis également horrifié. Maintenant, en suivant quelques définitions de 'éthique', c'est la même réaction que la précédente, en suivant d'autres, ce ne l'est pas. Donc, nous avons ici un problème supplémentaire; la façon de tracer la frontière de la morale est fortement contestée. Mais le phénomène de forte évaluation peut potentiellement dépasser la frontière. Ceci dit, elle varie selon une grande sélection de choses, dont certaines sont à l'intérieur de la frontière de l'éthique, d'autres sont à l'extérieur.

Donc à l'intérieur de la frontière de l'éthique, nous pensons que la question "Aimez-vous la crème glacée au chocolat ou à la vanille?" ne vaut pas vraiment la peine d'être débattue. Cependant, si vous alliez vous joindre à al-Qaida, nous serions horrifiés. Si vous feriez cela cependant, il y aurait une communauté quelque part qui célébrerait et qui dirait "Finalement! Vous avez compris." Nous voulons condamner cette attitude, mais comment? En clair, nous voulons dire que certains désirs ou valeurs sont plus désirables que d'autres au sens éthique, mais quels mécanismes utilisons-nous pour juger entre eux lorsque nous avons ces communautés en conflits?

Nous avons changé de position, correctement, mais néanmoins changé, dans le sens que la thèse reposant sur les langages de forte évaluation ait de l'emprise est nettement indépendante de si oui ou non l'un adopterait une vue complètement relativiste des valeurs. Ceci étant donné, je pense que nous sommes capables – pas dans tous les cas de conflits, mais dans certains cas – d'argumenter sur le sujet. J'ai essayé de décrire ce genre de choses dans un document comme 'Explanation and Practical Reason'.^{iv} Encore, nous ne pouvons jamais nous trouver dans une position où nous pouvons avec confiance croire que ceci est la position ultime, finale et irréversible. Mais nous pouvons aller dans le sens de démontrer que ceci est supérieur à cela, même si ce n'est pas pour dire que nous sommes certains que ce sera toujours le cas : nous pouvons aussi rencontrer d'horribles différences, des écarts béants où nous ne pourrions rien faire du genre. Ces différences peuvent être *de facto* ou *de jure* impossibles à rejoindre; *de facto* dans le sens que certaines personnes pourraient être parties si loin qu'elles ne com-

menceront même pas à considérer votre position comme étant valide dans quelque sens que ce soit... imaginez vous asseoir assez longtemps avec Bin Laden pour mettre les choses au clair. C'est impossible, il ne voudrait simplement rien savoir de votre position. Ce qui ne veut pas dire que vous ne pourriez pas vous asseoir avec l'un de ses suivants moins extrémistes. Donc, il est possible que nous ne sachions jamais si certaines différences sont impossibles à rejoindre, car l'impossibilité de rejoindre *de facto* prend une forme qui coupe la possibilité de vraiment l'essayer.

Mais nous connaissons des cas où vous pouvez trouver des arguments impossibles à rejoindre *de facto*, alors quelle attitude devons-nous adopter? Et ici je pense que les théories philosophiques présentes sont toutes à la mauvaise place, car la tentative est de montrer d'une certaine façon *a priori* que vous ne pouvez pas rejoindre les différences ou que *a priori* vous le pouvez. Mais comment cela puisse être un argument valide? C'est seulement en essayant que vous pouvez le savoir. Donc mon point – mon désaccord avec environ tout le monde en philosophie analytique pour le moment – est que l'enjeu n'a pas été bien cerné à cause de la puissance maléfique d'une certaine vue épistémologique que je nomme épistémologie médiatrice, où les gens sont supposés avoir un ensemble de suppositions finales par lesquelles ils regardent tout ce qui se présente, et ces suppositions incluent le domaine de la valeur. Cela dit, lorsque vous tombez sur quelqu'un avec une supposition différente, rien ne peut être dit. Ce genre de preuve *a priori* de la nature impossible à rejoindre des différences morales ou culturelles ne fonctionne pas, mais cela ne veut pas dire que nous avons la garantie que dans le cas contraire, toutes les différences sont possibles à rejoindre *a priori*. Nous pouvons retrouver des écarts qui sont si grands qu'ils sont impossibles à rejoindre.

Si il y avait une façon, comme vous l'avez suggéré, de mettre au clair les différences, est-ce que cela est parce que nous faisons partie d'une communauté universelle, la race humaine? Dans ce cas, y a-t-il des valeurs fondamentales qui pourraient être relevées de notre nature commune dans un certain sens, comme l'a tenté Aquinas?

Dans un sens, mais qu'est-ce que cela signifierait pour nous aujourd'hui? C'est pourquoi je parle de corriger nos prédécesseurs à un certain niveau; il est clair que les frontières de la nature ne sont pas aussi définitivement ou nettement tracées que Aquinas, suivant Aristote, pensait qu'elles l'étaient. Donc, c'est certainement un fait qu'il y ait une nature humaine, l'une des caractéristiques de cette dernière est l'aptitude et la nécessité de donner un genre d'interprétation à la vie humaine au moyen de la culture. Mais, à partir de cela, vous pouvez déjà voir que cela ne va pas tous nous rejoindre sur toute la ligne. Cela nous fait parcourir une partie du chemin, car nous savons que cette nécessité nous est connectée comme étant des êtres de langage, ce qui nous permet d'avoir des arguments intéressants à propos de la façon dont on doit

traiter les êtres humains en vertu du fait qu'il sont des êtres de langage. Et voilà la façon que j'interprète les écrivains tel que Habermas: je pense qu'ils opèrent d'une manière qui doit être comprise de façon beaucoup plus traditionnelle. Mais ceci ne couvre pas toute la gamme: ceci ne couvre pas les enjeux concernant qu'est-ce qu'une bonne vie ou une vie qui a un sens. Et ensuite nous reconnaissons d'ailleurs, depuis Herder, que même s'il peut y avoir une vaste gamme de différentes approches qui n'ont pas à être classées, certaines ont à être classées, et afin de résoudre le problème, nous devons regarder au développement et à la variation des êtres humains en entier. Donc ce n'est pas comme le type de raisonnement que vous avez avec Aquinas, où tout l'enjeu des différents mécanismes culturels de la vie humaine ne se pointe pas du tout à l'horizon.

Afin de réaliser ce genre de classement, certains commentateurs comme Sir Isaiah Berlin, qui a autrement été très sympathisant avec votre philosophie, pense que vous vous compromettez à un compte téléologique insoutenable des valeurs humaines. Quel est votre commentaire là-dessus?

Je pense que c'est une autre mauvaise vision de moi, à cause de cette culture philosophique pénétrante de *a priori*s autour de nous. Tout le monde est lu en ayant un *a priori*, et donc c'est l'un des possibles *a priori*s que vous pouvez m'attribuer: je me présente à la vie humaine avec ce lourd compte téléologique où tout se dirige. La culture philosophique ne peut pas d'une certaine façon permettre le fait que, non, nous ne savons pas toutes ces choses. J'aime les affaires de Isaiah Berlin à propos des conflits entre les vertus, à moins que cela soit endurci dans un *a priori*, lorsque nous disons que nous pouvons savoir aujourd'hui que nous ne serons jamais capable de les réconcilier... seulement une minute, seulement une minute. Si le point est la conjecture lourdement soutenue et très plausible qu'il y aura toujours un ensemble de valeurs ou de vertus très difficiles à combiner de la vie humaine peut importe où vous vous trouver dans l'histoire de l'humanité, oui, c'est aussi probable que le levé du soleil demain. Mais nous ne savons jamais si oui ou non nous pouvons lutter dans le sens d'une meilleure association de A et B. Nous luttons présentement avec l'association des organisations sans extrémisme et qui donnent beaucoup de sens à la vie, et c'est très difficile à combiner. C'est l'une des majeures oppositions entre deux éventuelles vertus de nos jours. Nous ne dirions pas cela à propos des Grecs, mais nous le dirions à propos de nous. Qui peut dire quel progrès peut être fait sur cela? Donc, ce que je n'aime pas à propos de la façon dont la position de Berlin est utilisée, incluant par Berlin, est qu'elle est endurcie dans un *a priori*. Et ma position de dire «Juste une minute,» est interprétée comme disant «Bien sûr qu'elles peuvent être combinées. Comment puis-je le croire? Bien, je dois avoir une téléologie.» Ceci arrive tout le temps, et c'est très, très drôle à voir. C'est une culture philosophique qui ne permet pas une autre possi-

bilité, et de plus veut toujours remettre les gens à un *a priori*. Vous vous devez d'avoir votre *a priori*, c'est comme votre passeport: «Où est ton passeport? Ton passeport Aristote, ton passeport Hume!» Non, non, non! Je suis sans position dans ce monde!

En référence aux valeurs d'aujourd'hui, sur lesquelles vous vous êtes retrouvé, l'un des symptômes majeurs que vous identifiez dans l'ère moderne, plus précisément, je pense à Sources of the Self, est-ce que l'idée de l'émergence à la fois de l'atomisme et de ce que vous appelez le relativisme facile (exposé dans les philosophies populaires comme 'Je peux faire ce que je veux'). Avec les développements de la dernière décennie, qui incluent l'imminente formation d'une Cour Criminelle Internationale, l'intervention étrangère dans des lieux comme le Timor Oriental, la coopération sans précédent dans le sillage du 11 septembre, ainsi que les mouvements de protestation contre la globalisation, le FMI, la Banque Mondiale etc., diriez-vous qu'il y eu un éloignement des attitudes atomistes et relativistes que vous avez décrites il y a plusieurs années?

Non, je pense qu'ils ont toujours coexisté avec ces autres types de mouvements, alors dans la mesure où il y a une oscillation, c'est une oscillation dans la vie d'un individu, ou il peut y avoir oscillation cependant, dans la vie de cohortes dans une société donnée. Les gens disent, «Les étudiants sont dans ceci maintenant» - mais c'est un certain niveau d'exagération journalistique, même dans ces oscillations. Vous n'avez pas besoin de lire aucun livre, vous avez seulement besoin de vivre aussi longtemps que j'ai vécu [rires]. Quand vous voyez les modes changer comme cela, vous devenez méfiant à propos de la foi dont vous devez mettre dans l'un de ces mouvements.

Ces développements sont reliés en particulier à l'augmentation des discussions sur les droits de l'homme. Vous avez suggéré que parler des droits n'a de sens que dans un cadre social dans lequel ils sont associés avec des obligations. À la lumière des mouvements discutés précédemment, pensez-vous qu'il y a eu un développement dans le sens des obligations du duo droits-obligations?

Bien, dans un sens, droits et obligations sont logiquement soudés, attachés ensemble, donc ce n'est pas l'enjeu: si vous avez le droit à la vie, j'ai l'obligation de ne pas la prendre. Mais lorsque des gens comme Amitai Etzioni aux États-Unis parlent de à quel point nous avons trop d'emphase sur les droits et pas assez sur les obligations, alors ils pensent quelque chose du genre, «Nous avons trop d'emphase sur ce que nous pouvons demander de la société, et pas assez d'emphase sur ce que je dois donner à la société.» Comme proposition générale, je ne suis pas contre cela, mais je déteste demeurer à ce niveau de généralité car il y a beaucoup de choses que les gens pourraient mettre de l'avant comme obligations et que je n'accepterais pas vraiment. Je préfère la mettre en termes de comprendre comment la vraie nature de la démocratie requiert un

certain degré de dévouement de la part des gens qui constituent un peuple au sens démocratique. Ainsi, ils s'auto-gouvernent ensemble, ce qui requiert un certain degré de dévouement au procédé, au projet qu'ils ont ensemble et, de plus, aux personnes qui sont dans le projet. Quand cela faiblit, vous pouvez voir les terribles conséquences. Par exemple, lorsque le sentiment de solidarité s'affaiblit, vous avez de très profonds ressentiments et un sens d'aliénation et ainsi de suite. Donc, je préfère suivre cette ligne de pensée peu importe où elle mène, et elle peut aussi mener, entre autres, à un sentiment d'obligation.

Si, philosophiquement parlant, les droits et les obligations sont dans un sens logiquement liés ensemble, le défi est alors de faire réaliser aux philosophes et aux gens en général ce qu'il est insensé de donner priorité à la discussion sur les droits (comme nous le voyons dans beaucoup de pensées qui posent le droit comme une sorte d'unité fondamentale de l'être humain), et seulement par la suite vient la discussion sur la société. Je prends pour acquis que votre défi est de mettre à jour que la notion de droits et d'obligations peut seulement prendre place à l'intérieur d'un certain arrangement.

C'est correct, et donc le poids de la complémentarité, qu'il est, je pense, intéressant de ressortir la complémentarité entre divers types d'affirmations des droits individuels, ou l'individualisme en divers modes différents, l'individualisme expressif, l'individualisme instrumental et ainsi de suite, d'un côté, et les formes d'alliances sociales qui se sont développées avec eux, l'un des plus importants étant l'auto-gouvernement du peuple. Sans ce dernier, simplement affirmer les formes d'individualisme serait extrêmement instable et même autodestructeur, auto-undercutting. Une fois que vous avez réalisé cela, appartenant à ces formes de liens sociaux deviennent un *desideratum* très important de la vie dans notre société, car sans eux vous ne pouvez pas être un individu. Je dis simplement ce que Toqueville et les autres ont dit. Encore, je pense que la façon dont la discussion a été restreinte dans l'environnement philosophique présent est extrêmement *a priori*. Au même point, vous voulez dire "Arrêtez de parler de qui à le droit de quoi," et commencez à parler de la prudence sociale minimale, comme ce qui va nous permettre d'avoir une société démocratique saine dans laquelle les droits du peuple sont respectés.

En suivant les mêmes lignes, une bonne partie de votre œuvre met l'accent sur la relation entre l'individu et la communauté, et comment cette dernière informe largement l'identité de l'autre. Les événements récents ont mis en évidence le réseau complexe d'allégeances qui compromettent les identités des peuples, soulignant les tensions qui n'existent pas seulement à l'intérieur des communautés, mais aussi à l'intérieur d'un seul individu, qui a souvent un ensemble d'allégeances à plusieurs facettes sur lequel son identité est basée. Étant donné ces différences, sur quelle base une communauté est une communauté? Est-ce en vertu de

vivre ensemble, de livrer les mêmes batailles? Comment devons-nous comprendre ces allégeances sur un niveau à la fois individuel et communautaire? Est-ce que certains ont la priorité sur d'autres? Est-ce que certains devraient avoir la priorité sur d'autres?

Je ne pense pas que vous puissiez donner une réponse générale à la dernière question parce que les situations varient. Même en regard de la première, la généralité de ma réponse va être très historiquement restreinte. Je pense que si nous voulons répondre à cette question dans le contexte des démocraties modernes que nous connaissons dans la région du nord de l'Atlantique de nos jours, la nature du lien est clair. Vous pouvez commencer par quelque chose comme l'idée du 'patriotisme constitutionnel' de Habermas, mais donnez-y une certaine tournure : elles sont toutes des cas dans lesquels les liens entre les personnes doivent avoir été, et le sont habituellement, un mélange de certains principes (démocratie, etc.) d'un côté, et un fort dévouement à un certain projet historique de la réalisation de ces principes de l'autre côté. Je dis 'historique', le projet particulier est seulement réellement identifié par le fait qu'il vient de quelque part, par le fait qu'il a une histoire. Ce n'est pas identifié seulement en ayant un certain emplacement géographique, parce qu'elles sont des pays, comme la Pologne, qui a été totalement déportée, ou presque totalement déportée, en 1945. Donc, ce n'est pas la région. C'est cependant l'histoire, le peuple et la population continuellement changeante. C'est ce qui doit être généré. Si les gens n'ont pas cela, alors il n'y a pas le genre de lien qui peut permettre une saine démocratie. Évidemment, la façon selon laquelle vous développer cela fait ressortir tout l'enjeu du multiculturalisme. Dans tous ces cas, il doit y avoir une conscience de la dimension multiculturelle de ces sociétés. Si vous voulez rejoindre tous ces gens, le projet historique, disons de la démocratie du Canada, se doit d'avoir cette dimension en tant que partie de son identité, de ce quoi il comprend être.

Nous en avons débattu au Québec, nous avons un gros problème, et je crois que le problème est maintenant en train de s'estomper, dieu merci. Les fameuses remarques par Parizeau à propos de 'l'argent et le vote ethnique' démontrent que pour certains Québécois nationalistes, le projet était défini presque totalement en termes *pure laine*. Presque tout le monde reconnaît maintenant que le genre de population que nous avons présentement au Québec – c'était le cas avant, mais ce l'est encore plus aujourd'hui – rend impossible pour une telle base de se rejoindre autour de ce projet historique. Si les gens sont liés à ce projet, ils comprennent évidemment, et acceptent, qu'il a des origines *pure laine*. Mais ce n'est pas ainsi aujourd'hui, et ce n'est pas là où cela se dirige. Plus ou moins toutes les sociétés démocratiques font présentement face à la tâche de redéfinir leur identité, certaines le font bien, d'autres mal.

Je prends cela comme un signe d'une saine démocratie quand vous avez, d'un côté ces

allégeances qui s'affrontent et, de l'autre, les gens qui reconnaissent qu'il y a des principes ainsi qu'un cadre à l'intérieur duquel les classer. Ceci pour dire que nous comprenons qu'il y a une limite dans la façon dont vous pouvez exprimer vos opinions: protestations pacifiques et non-violentes ou même un référendum qui sont une façon très pacifique de régler ce que certains voient comme conflit culturel ou même ethnique si nous pouvons aller aussi loint.

Oui, absolument, oui.

Ces allégeances sont évidemment très fondamentales à votre pensée morale dans le même sens qu'elles sont fondamentales à votre philosophie de l'esprit, car elles éveillent beaucoup de ce que c'est d'être une personne. Et ceci soulève avec une force renouvelée la question que nous avons touché précédemment: Que faisons-nous réellement lorsque les moyens fondamentaux s'entrechoquent? Vous avez mentionné, "Mettre les choses au clair," mais parce que cela semble être un enjeu si fondamental, les gens ont voulu, comme vous l'avez dit, un ensemble de lignes directrices plus fortes à l'intérieur desquelles répondre à de telles questions.

Bien, il n'y a pas de réponse à cette question; certainement pas de réponse philosophique et peut-être pas de réponse *überhaupt*. Ceci dit, il peut y avoir une situation où vous échouez, où vous ne pouvez avoir une démocratie ordonnée. Supposant qu'une grande force internationale déclare demain que tous les habitants de la Palestine, Israélites et Arabes, doivent former une démocratie unique avant le lendemain: il y a une condition de guerre civile virtuelle, vous ne changez pas les gens de cette façon en une nuit. La démocratie se rupture, car il y a trop de violence... nous espérons que cela n'arrive pas à nos démocraties occidentales 'bien ordonnées', mais c'est possible à tout moment. Comme vous l'avez dit plus tôt, il doit y avoir cet accord sur certaines limites, qui appartiennent à cette première partie du patriotisme constitutionnel et l'ensemble des principes. Si il y a vraiment un profond rejet de certaines, alors, il n'y a pas de solution. Donc ce n'est pas seulement que la philosophie ne peut produire une réponse, parfois *rien* ne peut produire une réponse. Nous savons à quoi ressemblerait la solution, et nous pouvons lutter aussi fort que possible pour ramener les gens à ce consensus, mais nous pourrions échouer. L'histoire est pleine d'échecs agissants.

Pour fermer le cercle de cette entrevue, nous avons commencé en parlant de l'approche très historique des problèmes. L'un des courants de cette entrevue a été ce que vous voyez que la philosophie de l'approche ahistorique et de fort a priori cause trop de problèmes de nos jours. Je considère que ces deux sont fortement liés. Ce que j'ai recueilli de ce que vous avez dit est que lorsqu'il est question de théorie morale, ou de parler des communautés, ou de la possibilité de la démocratie et de la paix, c'est beaucoup plus une question de lire ce que sont les

possibilités pour le futur, dans une situation particulière, hors de l'histoire, donc de parler, et pas simplement prendre ces problèmes de façon abstraite, et en venir à des principes, qui disent "C'est le cas. Ce n'est pas le cas."

Je pense que l'imbroglia au Canada a été empiré par ce second type de pensée. Ce que nous avons est quelque chose comme un état binational ou multinational, mais certaines personnes en ont empêché la reconnaissance, parce que cela semble logiquement faux à leurs yeux. Ils pensent qu'il doit y avoir une uniformité de la relation des citoyens avec l'état, qui par définition n'est pas le cas dans une société multinationale. Et donc nous avons eu tout ce blocage envers le Québec comme étant une société distincte des autres parties du pays, souvent basé sur des points de vue hautement philosophiques, comme une interprétation très uniforme de la Charte des Droits. L'accord du Lac Meech ne fonctionnerait peut-être pas, car il peut y avoir différentes interprétations de certaines clauses de la charte à différents endroits. Alors ils rejettent hors de la main comme étant impossible ou inconcevable, la seule solution qui fonctionnerait dans notre pays; car si vous avez une société très multinationale, vous devez avoir un certain genre de reconnaissance de ce fait. Et donc je pense que la mauvaise philosophie a fait beaucoup de dommages politiques, et je suis très ferme là-dessus car c'est mon pays et il y a eu beaucoup de dégâts ici.

Notes

- ⁱ Daniel C. Dennett, *Consciousness Explained* (Toronto: Little, Brown, 1991).
ⁱⁱ Owen Flanagan, *Consciousness Reconsidered* (Cambridge: MIT Press, 1992), 87-107.
ⁱⁱⁱ Charles Taylor, *Human Agency and Language: Philosophical Papers 1* (New York: Cambridge University Press, 1986), 164-186.
^{iv} Charles Taylor, *Philosophical Arguments* (Cambridge: Harvard University Press, 1997), 34-60.
^v Charles Taylor, *Sources of the Self: The Making of Modern Identity* (Cambridge: Harvard University Press, 1989), 111-207.

A Response to G.A. Cohen's 'A Brief Look at Socialism'

Jan Narveson, University of Waterloo

Jan Narveson has been at Waterloo since 1963 and is the author of *The Libertarian Idea* (recently published by Broadview Press, Canada), *Moral Matters* (Broadview, 2nd ed. 1999), and *Respecting People — In Theory and Practice* (Rowman & Littlefield, USA). He has known Professor Cohen "for a long, long time" and counts him as a friend.

Cohen does not agree that left-wing *thought* has collapsed in the wake of the events of a decade or so back, and depending on what you count as "thought" and mean by "collapsing," I suppose he is right: people will keep on making the sort of assertions of which "left-wing" thought is said to consist. But I think we can make a more robust claim: that there simply isn't anything left to defend, that all the old arguments, most especially those of Marx and Marxists, have been shown to be worthless. Those attempting to regroup have a problem that might just be insoluble. In the very least it will certainly require something radically different than the tired stuff we get, even from people as acute as Cohen.

Here's an example. Cohen remarks that the various Communist countries that collapsed at the end of the 1980s were

far too tyrannical to represent an inspiring realisation of the socialist ideal. But those countries nevertheless constituted a non-capitalist space upon which leftists could project their hopes for a non-tyrannical socialism, for a society neither driven by the market, and therefore dominated by the rich, nor controlled by an anti-democratic elite.¹

It's the "therefore," as well as the implied assertion following it, that I object to. In what sense is capitalist society "dominated by the rich"? Is it meant that particular people with high incomes more often get into the papers than particular people of low incomes? Or what? The thesis of "domination" sounds ominous, but if it is, it ought to refer to something, and what it refers to should be bad. So, let's hear about it. What is it? Not, I trust, that people in general prefer higher incomes to lower ones. Nor, I

hope, that people will do anything for another dollar: they usually won't, of course, and capitalism, far from "driving" them to do this, is neutral on the subject. Its genius is to allow people to do more of what they want rather than to require them to do what someone else wants. If some people spend time keeping up with the Joneses, well, that's their call: in no way do they "have to," and, moreover, there is no conceivable society in which it would be impossible to do that sort of thing.

For that matter, if we look at the newspapers in capitalist countries, especially America, they are dominated not by capitalists but by baseball stars, movie stars, country-western singers, and on and on. Most of these people make a lot of money, to be sure — but they are not "capitalists," nor is their income the main reason they get into the papers. On the contrary, they are proletarians, in Marxist lingo: they make their money mainly from the sale of their services, not from investments (and most of us lowly proletarians also make some money from investments, as Cohen well knows). And the people who *do* make money from investments, with the occasional exception such as Bill Gates, who managed to get himself investigated by an American congressional committee, aren't seen much in the newspapers, unless you count the financial sections. And in those sections, of course, what else could you expect? "Domination"? Fiddlesticks!

Cohen has, in his writings, spelled out further his thesis about an "irreversible transformation in the structure of capitalist class society." Well, what "structure" and what "class"? I am tempted to ask. This Marxist category of a capitalist "class" is, after all, nothing but a tautological figment of Marx's analysis. The capitalist class is the class of capitalists, yes. But so what? Marx, and apparently still Cohen, after all these years, wants to tell us that this nasty class "oppresses" the rest of us, but the theoretical structure erected to support this conclusion lies in ruins, and the sheer claim itself is pretty close to meaningless, unless the meaning is that the rich have — hey, what do you know!? — more money than the poor. This would be bad, no doubt, if there were no good reason for it, but there generally is: investment is important and does all sorts of good things for the rest of us, in the way of producing new goods, lower prices, more variety, and so on. Cohen used to push an idea that capitalists don't *do* anything, but in no useful sense of the word "do" is this true. If he asked me to give, or lend, him a million dollars and I did it, the claim that I hadn't "done anything" for him would be silly. In like manner, the people who invest huge amounts in risky industries — most of which fail, by the way — and the successful ones among them who do all these good things for us have certainly earned their incomes on any reasonable view of the matter. Socialists, indeed, are in the business of taking an unreasonable view of the matter, but we shouldn't let them get away with it.

Anyway, the four characteristics of the "working class" that were supposed to change forever the prospect of a socialist revolution are these:

"it was the section of society on whose productive activity everyone depended"

"it constituted the great majority of society"

"it was composed of the exploited people in society"

"its members were in serious need; they experienced a substantial lack of the essentials of a fulfilling life"

Cohen's claim is that there has been this big change — that the proletariat no longer lives up to those specifications. If it did, he thinks, then "it would produce a socialist revolution." Against this I want to say that (a) it wouldn't do anything of the sort, and the idea that it would is a piece of Marxist romanticism (all the actual socialist revolutions were pushed by arrogant, gun-toting intellectuals and assorted thugs — not by working-class people as such); moreover (b) it in any case never did exemplify those supposed properties, in the sense in which Marx (and Cohen) meant (and mean) them. The whole Marxist story is misguided from the start, resting on faulty analysis, biased reporting, and romantic enthusiasm rather than hard-nosed social science.

To get back to the four characteristics:

(1) As to the "workers" being those "on whose productive activity everyone depended," that, of course, is true, but equally true is that those productive people depended on the others: the visionaries, the technically and entrepreneurially venturesome, the managers — something both Marx and, much more culpably, Cohen tend to forget. You wouldn't have hundreds of thousands working in factories if there were no factories for them to work in, and you wouldn't have the factories if capital had not been intelligently invested, or without ingenious and hard-working engineers, management-design people, and so on, getting it all together. In capitalist society, everybody depends on everybody, and the idea that it's split between two unequal segments, one of whom is the Bad Guy, is, in a word, bosh.

(2) The claim that proletarians constituted the "great majority of society" brings up, most pressingly, the need for a more precise and also relevant explanation of what is meant by "proletariat" or "working class." For example, during most of the 19th century there was still a huge agricultural "class" in England as well as in America. And in America, the main component of that country's meteoric rise to the status of the pre-eminent industrial country in the world was the improvements in agriculture — not in factory-type work, except for the factories that produced agricultural implements — that enabled a smaller and smaller number of people to produce more and more of the food. And in America especially, those farmers were all capitalists while they were at it: they invested in the equipment, etc., that made this a true industrial revolution, while at the same time providing most of the manpower. The Marxist story, in other words, is ludicrously under-informed, just as all leftists who talk that way are talking, by and large, through their hats. They don't know what's coming off, and their schemes don't

have all that much to do with reality.

(3) Next we are told that those proletarians are "the exploited people in society" — again, a claim that tends not to be explicated with care and relevance. As we know, the Marxist thesis of "immiseration" was altered from an original version, which seemed to claim that working-class people in England were actually getting poorer as a direct result of capitalism, to the vague claim that the gap between rich and poor was rising or some such thing. Now, the first version is a claim that was the very reverse of the truth, as we now know, at the level of facts — and moreover which, as I have pointed out elsewhere, also made no sense whatever in theory. Marx had *no* reason to make the first claim. As to the second, there is the very burning question of, Who gives a damn? If Mr. Proletarian is a great deal better off than he was before, which was in virtually all cases true, why should the fact that some other people are even better off be of any interest to him? Cohen, with his egalitarian proclamations, apparently thinks that there is something important about this, something that the rest of us should worry about. It isn't so, and there's no reason for it to be so, and nobody on the left has bothered to explain what is supposed to be important about it, and there the case stands — with a completely unsupported, purely moral proclamation, or a complete falsehood, or both.

(4) Finally, and related to the preceding, there is the claim that the proletariat's "members were in serious need; they experienced a substantial lack of the essentials of a fulfilling life," a claim which would be nice if it were made in a meaningful way instead of remaining a platitude that all "correct-thinking" leftists get to assume without further analysis. But, in fact, it is a platitude insofar as it's anything. People in the 19th century were poorer than they are today, of course, and the poor today, though immensely wealthier than their ancestors, certainly wouldn't mind having three VCRs instead of two. But, for goodness's sake, what else is new? We all continually revise upward the minimum which those at the "bottom" of society supposedly "need," but this no longer means "need in order to live" — now it's "need in order to meet the Joneses standards." Stirring up sympathy for North American "poverty-level" people with their decent housing, indoor plumbing, two TVs, 1.3 cars, and so forth, would be an uphill go in, say, Cuba. The truth of the matter is that the leftist concerns on this front are largely fraudulent, and insofar as they aren't, we can blame the efforts of the modern welfare state itself at least as much as the workings of bad luck and of capitalism.

The legacy of Marxism remains extremely effective in contemporary front-line countries, and perhaps more so in the British, American, and Canadian academe, even, than the rest of Europe. But it's largely hooey and has coloured the thinking of people in positions of major political power quite enormously and almost entirely for the worse, so far as I can see. It's quite true that Marxism is dead, but it always was — it isn't that something has happened, somehow, that made this originally powerful and

plausible idea less powerful and less plausible. Marx, in fact, was a sophomore, an intellectual brat whose ideas had been refuted even before he got them on paper, and there is no excuse for anyone, and certainly not for very intelligent people like Cohen, to be taking him seriously.

Regarding Analytical Marxism, Cohen says: "While you do not need to read Galileo or Newton to be a good physicist...we have not progressed so far that it is time to stop reading Marx. The study of Marx and Engels remains an indispensable element in a scientific socialist's education." Asked whether he feels that the canon studied by undergrads in philosophy today does justice to "the issues addressed by left-wing thinkers," Cohen replies that there is "much more radical substance *on offer* than there used to be." And he is, alas, entirely right about that. What gets really short shrift is a decent, clear-eyed look at capitalism, which most contemporary intellectuals compulsively dump on, almost always with very little idea what they're talking about, and which few students coming out of high school have much of a clue about. It's not too much to say that on the public-relations and educational fronts, the leftists have things pretty much to themselves nowadays, especially here in Canada, though substantially in the United States as well.

In response to the question of how he sees his own contribution to the "egalitarian debate" playing out on the social and political level, Cohen replies, "I do not look at my particular academic work as a cause, operating on its own, that I hope will have a political and/or social effect. I am a small part of something much larger, a stream of radical thought, a stream that is always in being....If you want to know what the effect of that stream is, try to think of how things would be if it did not flow. There would never have been freedom from feudal bondage, there would never have been universal suffrage, there would certainly never be that society of community and equality for which we continue to struggle." It would be hard to know what to say to that if it didn't so thoroughly confuse the genuinely liberating tendencies in Western society, especially, from the irrelevant and the counterproductive. Cohen's statement brings up, in a major way, how we are to understand this "stream."

It certainly was not socialism that got the feudal serf out of bondage. It was, rather, the draw of the cities with their better incomes and greater opportunities. It was not socialism that got the Industrial Revolution under way — it was, rather, the very bourgeois ideology and looming capitalist opportunities that Marx and his fellow socialists denounce. It is much less than clear what universal suffrage has to do with it. Suffrage in Britain in the early 19th century was far from universal, and yet that was the period rightly described as the Industrial "Revolution." And as to the "society of community and equality," Cohen does not clearly define it in his work, nor do most who identify themselves as being on the left these days. But the rhetoric of the left today, as well as its historic character, as for example in the works of Marx and Engels, tends to reinforce the suggestion that the left wants to drag society kicking and scream-

ing out of the market altogether and back, in effect, to the pre-industrial era. On this version, I see the left being the enemy, not the friend, of human progress. Certainly in the 20th-century versions of real-world socialism that Cohen understandably has his doubts about, that seems a fair verdict. I think it still is as applied to Castro's Cuba as well as North Vietnam — the only ones left.

We need to know from "the left" more precisely what they want. Many of them end up just saying that we should have a big welfare state. But we already have that, and it's hard to see that as being a sufficient instantiation of "egalitarian" intentions. It is, I think, hard to see it as a good example of anything but the golden opportunity for bureaucratic expansion at the public expense afforded by democracy. More to the point, I find it very hard to see what role substantial egalitarianism *could* play, other than a bad one. Taking money from the most productive people in our community and handing it over, no questions asked, to the have-nots does not strike me as a rational idea, and certainly not an example of justice in action on any reasonable view of the matter. Nothing any leftist has said looks to me like any sort of real justification for that. Cohen himself has spoken out against Paretianism — the view that if X is better for some people and no worse for anyone, then it's OK. That should mean that he objects to such changes. And if he does, I want to know where the freedom that he frequently talks as though he is in favour of has gone, and how a state that took seriously that rejection could be anything but the totalitarian monster I have objected to above.

Notes

¹ All citations of G.A. Cohen are from 'A Brief Look at Socialism: An Interview with G.A. Cohen,' *Pensées*, 2, 2001, 11-17.

Alcibiades Redeemed: The Ascent to Divine Madness

Tara Tonja Mitchell, Trent University

Many readers of Plato's *Symposium*, after reading Alcibiades's entrance and speech, conclude that the story of Alcibiades is a "story of waste and loss."¹ Alcibiades is often interpreted as the character who has allowed his vanity, sexual appetite, and worldly ambitions to become more important to him than the cultivation of his soul. He has been accused of forsaking his desire to become "as good a person as possible" (*Symposium*, 218d)² for the distractions of the sensual world. I will argue that these charges against Alcibiades are, for the most part, unfounded. Instead, I will argue that Alcibiades reveals himself as one who has a degree of understanding both of things beyond the sensual world and of philosophy — rare qualities indeed. Like Socrates, Alcibiades is a Silenus statue. He appears to be one thing on the outside (drunk/ridiculous), but on the inside he reveals himself as having an inner understanding of things divine.

The Seduction of Alcibiades

Alcibiades is often accused of being unable or unwilling to allow Socrates to guide him in the cultivation of his soul. Daniel Anderson's argument is that Alcibiades, resentful about Socrates's rejection of his sexual advances, let his vanity get the best of him.³ But Alcibiades's attempts at seducing Socrates should not be a mark against his moral aptitude, nor does it indicate an essential character flaw that forever hinders the cultivation of his soul. If we allow for Socrates's development, despite his own base attraction to boys and gold in his youth,⁴ then we should grant the same consideration to Alcibiades.

Alcibiades's speech suggests that, in his youth, he was ultimately concerned with seducing Socrates, but his intention to do so was more noble than vanity or sexual gratification. Alcibiades was physically rejected by Socrates (repeatedly), but he did not

doubt that Socrates would be of value to him. He tells us that Socrates once cautioned him about "trying to get true beauty in return for its appearance." He also says that Socrates told him to "make sure you're not making a mistake in thinking I am of value to you" (*Symp.* 218e-219a). These are ironic cautions, since Socrates is fully aware of own his value to Athens⁵ and of his personal duty (and thus value) to Alcibiades.⁶ If the lesson is about the necessity of Alcibiades recognising for himself Socrates's true value, he does. Alcibiades says:

Nothing is more important to me than becoming as good a person as possible, and I don't think anyone can help me more effectively than you can in reaching this aim....My plans are exactly as I've said. It is now up to you to consider what you think is best for you and for me. (*Symp.* 218d-219a)

Alcibiades pursues Socrates out of a desire to know, not sexual frustration. The problem is that Alcibiades thought the only way to learn everything he could from Socrates was to seduce him: "I thought he was seriously interested in my looks and that this was a godsend and an amazing piece of good luck, because, if I gratified him, I'd be able to hear everything he knew" (*Symp.* 217a). Although it may appear that Alcibiades was trying to trade "gold for bronze" (or trying to trade absolute Beauty for physical beauty), he was merely following the conventions of Athens at the time, which supposed that a youth was best educated trading "youthful charms" as a form of devotion towards an older, wiser man who would educate him in virtue and affairs of the state.⁷ Moreover, Socrates had a reputation for being "erotically attracted to beautiful boys" (*Symp.* 216d) and yet, according to Alcibiades, Socrates "triumphed over my good looks — and despised, scorned and insulted them — although I placed a very high value on these looks" (*Symp.* 219c). Socrates's rejection of his physical beauty and advances was, no doubt, a shock to a young and beautiful Alcibiades. It went against conventional beliefs and practices on how a youth was to obtain an education, and it went against the conventional belief that Socrates was attracted to beautiful boys and "always hanging around them in a state of excitement" (*Symp.* 216d). Therefore, Alcibiades's initial reaction of humiliation and confusion was appropriate given the context.

Socrates's rejection served the purpose of leveling Alcibiades's vanity as a sense of identity, to be sure. He thought that his physical beauty would guarantee him an education, but instead he "got nowhere" with Socrates (*Symp.* 217c). Socrates's physical rejection does not leave him permanently insulted, nor does he misunderstand this rejection. Alcibiades, although "humiliated," says that he "admired [Socrates's] character, his self-control and courage" (*Symp.* 219d). Humiliated, surely, and yet convinced that Socrates possessed a rare quality: "Here was someone with a degree of under-

standing and tough-mindedness I'd never expected to find" (*Symp.* 219d). This is not a reaction of resentment; it shows that Alcibiades values the qualities of tough-mindedness and self-control more than his own vanity as a sense of identity. It is significant that after recognising "a degree of understanding" in Socrates, Alcibiades no longer mentions the "high value" he placed on his looks nor any other attempts to seduce Socrates for the rest of the speech. Socrates's rejection went deeper than vanity or humiliation; it put Alcibiades's desire for virtue in jeopardy: "What I'd seen as the only means of catching him had proved a failure. I was baffled; and I went around more completely enslaved to this person than anyone else has ever been to anyone" (*Symp.* 219e). Alcibiades is baffled because he does not know how to approach the one person he believes could help him "become as good a person as possible." Therefore, I disagree with Anderson's assertion that Alcibiades ultimately sees Socrates's rejection of him as an "insult to his youthful charms." Recognising a "degree of understanding" in Socrates indicates that Alcibiades has a degree of understanding himself. It suggests that he recognises the virtue in Socrates's rejection: that self-control over physical desires is important and that the importance we attach to appearances needs to be questioned.

Alcibiades: Human or Divine Madness?

To address Anderson's charge that Alcibiades merely tolerates Socratic discourse so that he can ultimately seduce him,⁸ it is important to note that a theme throughout his speech is the significance he places on Socrates's words and the effect they continue to have on him: "the effect his words have had on me — an effect they still have now, whenever I listen to him...my heart pounds...tears flood" (*Symp.* 215e). The spell-binding effect that Socrates has on Alcibiades is experienced only when Socrates speaks; other great orators do not have this kind of effect on him (*Symp.* 215e). Alcibiades claims to have been "struck and bitten by the words of philosophy, which cling on more fiercely than a snake when they take hold of a young and talented mind, and make someone do and say all sorts of things" (*Symp.* 218a). Alcibiades not only feels the impact of Socrates's words, he feels compelled to do what Socrates tells him to do. Indeed, he claims that he *should* do what Socrates says, that he has *agreed* to do what Socrates says (*Symp.* 216b), and that he *must* do what Socrates says (*Symp.* 217a). This is not someone who is acting from the compulsions of vanity; this is someone deeply affected and disturbed by the words of Socrates.

For Rosen, Alcibiades's words and his slavish behaviour towards Socrates show that "the initial enchantment of contact with Socrates produces confusion and self-doubt, not virtue." To Rosen, Alcibiades's slavish behaviour suggests that he suffers from human, not divine, madness.⁹ I disagree that Alcibiades suffers from human madness and will argue that Alcibiades suffers from divine madness. For the purpose

of this paper I will define human madness as attachment to appearances or sensory experiences (such as erotic physical attraction, drunkenness, reputation). Divine madness as an attachment to the abstract, immortal, and divine (the soul and true Beauty).

In his eulogy to Socrates, Alcibiades describes the experience of once seeing something in Socrates "so divine, golden, so utterly beautiful and amazing that – to put it briefly – I had to do whatever Socrates told me to" (*Symp.* 217a). Alcibiades's account of his slavish behaviour towards Socrates and his discourses is strikingly similar to Socrates's account of divine madness and the soul's recollection of divine Beauty in Plato's *Phaedrus*. In that dialogue, all souls catch sight of divine Beauty while disincarnate in the divine realm. However, very few souls are able to recollect and *recognise* a likeness of divine Beauty in the temporal world:

Every human soul has, by reason of her nature, had contemplation of true Being; else would she never have entered into this human creature; but to be put in mind thereof by things here is not easy for every soul; some, when they had the vision, had it but for a moment; some when they had fallen to earth consorted unhappily with such as led them to deeds of unrighteousness, wherefore they forgot the holy objects of their vision. *Few indeed are left that can still remember much: but when these discern some likeness of the things yonder, they are amazed, and no longer masters of themselves, and know not what is come upon them by reason of their perception being dim.* (*Phaedrus* 249e-250b)¹⁰

According to this passage, Alcibiades does not belong to the category of souls who "forgot the holy objects of their vision" because of misdeeds while incarnate. Nor does he seem to belong to the category of souls who "had the vision...but for a moment," only to never perceive the divine while incarnate, since he claims he encountered the divine through Socrates. The passage from the *Phaedrus* shows that a direct consequence of beholding a divine image while in the temporal world is that it leaves one "no longer a master of themselves."¹¹ Therefore, Alcibiades's slavishness is not associated with human madness; it is a direct consequence of an encounter with the divine while existing in the temporal world. In fact, Alcibiades's ability to recollect and discern the divine while incarnate places him among the very few souls "whose memory is good enough" to perceive an image of what they saw while disincarnate in the divine realm (*Phaedrus*, 250a).¹²

To further support Alcibiades's rather privileged ability to perceive the divine while incarnate, it is important to note that Alcibiades is presented in the *Symposium* as someone who is privy to something that few are privy to. He claims to know Socrates and his discourses better than anyone else. Alcibiades declares at the begin-

ning of his speech, "You should realise that none of you really knows him. But I'll show what he's like, now that I've made a start" (*Symp.* 216c). At the end of his speech, Alcibiades claims that Socratic arguments are "most divine" and "contain the most images of virtue. They range over most — or rather all — of the subjects that you must examine if you're going to become a good person" (*Symp.* 222a). These claims are significant because they imply that Alcibiades has a certain insight into Socrates that no one else has at the Symposium ("none of you really knows him"). They show that he is intimately familiar with the *entire* corpus of Socratic discourses and that to study them is to aid the soul's progression towards the good ("They range over most — or rather all — of the subjects you must examine..."). They show that he is aware of an inherent divine nature in Socratic discourse ("they're the most divine and contain the most images of virtue"). These assertions tell us that even those present at the Symposium are not as intimate with Socrates as Alcibiades is, enamoured and experienced as they are with Socrates and "the Bacchic frenzy of philosophy" (*Symp.* 218b). Nor does Socrates dispute these claims, even though Alcibiades says he must, if what Alcibiades is saying is false (*Symp.* 217b). Alcibiades stands alone in the *Symposium* as the character who is most intimate with Socrates's divine nature and discourses. As such, it is unfortunate that Alcibiades's experience of seeing something divine in Socrates has been dismissed as an unclear experience that is of little worth since it does not adequately explain what he saw or what it meant.¹³ Alcibiades's disturbed personality is a consequence of divine recollection. Therefore, he does, in a sense, explain what the vision means; it means disorientation of self in the world, it means divine madness. This disturbance is misunderstood by others as nothing more than peculiar behaviour, but it is much more significant than that. Socrates says that one possessed by divine madness "stands outside human concerns and draws close to the divine; ordinary people think he is disturbed and rebuke him for this, unaware that he is possessed by a god" (*Phaedrus* 249d).¹⁴ What many interpret as the human madness of Alcibiades may, in fact, be divine madness. Often our tendency is to believe that the soul's encounter with the divine must be characterised by expressions of sobriety and self-control. However, the soul's evolution, according to Socrates, "is inevitably a painfully difficult business" (*Phaedrus* 243b).¹⁵ This is the business of Alcibiades; Socrates "disturbed my whole personality," he says (*Symp.* 215e).

"I Think You Are Sober After All, Alcibiades"

Socrates's legendary ability to consume extraordinary amounts of liquor without any ill effect on his body and mind is mentioned more than once by Alcibiades (*Symp.* 214a, 220a). According to him, Socrates's ability to drink without getting drunk puts him above the lot of common men who would be, no doubt, of unsound body and mind. And yet if we consider Socrates's inability to get drunk a superhuman feat, we

should consider that Alcibiades has the same peculiar quality. Contrary to his appearance at the Symposium, he is not particularly drunk, but the way Plato presents him and the way he presents himself may lead us to assume otherwise.

Alcibiades tells us that "the most amazing thing of all is that no one has ever seen Socrates drunk" (*Symp.* 220a). Socrates's inability to get drunk is also mentioned earlier in the dialogue in a rather matter-of-fact manner by Eryximachus. Socrates appears to be the exception to Eryximachus's medical opinion that "drunkenness is harmful to human beings" (*Symp.* 176d). The others at the Symposium seem to share Eryximachus's opinion; they either do not have heads for drinking (*Symp.* 176c) or "the strongest heads for drinking have given up" (*Symp.* 176c) because they have run out of stamina, and as such, they decide to drink moderately for the rest of the evening (*Symp.* 176c-e). Only Alcibiades and Socrates, with their mutual stamina, show any kind of contrast to the moderation of the others.

First, we are told by Plato that Alcibiades is already "very drunk" when he arrives at the Symposium (*Symp.* 212d) and Alcibiades refers to his own drunken condition many times (*Symp.* 212e, 214c, 215a). Shortly after arriving, Alcibiades disrupts the established protocol of moderation at the Symposium by quickly consuming more than four pints (over eight cups) of *undiluted* wine¹⁶ by himself¹⁷ (*Symp.* 213e-214a). Alcibiades has Socrates drink the same amount immediately afterwards, and Alcibiades refers to Socrates's inability to get drunk, I believe, to draw a comparison between Socrates and himself. After all this, Alcibiades not only has the wherewithal to give a eulogy to Socrates, but he is able to deliver what he claims he cannot deliver: "If I don't remember things in the right order, don't be surprised. It isn't easy for someone in my condition to list all the aspects of your peculiarity in a fluent and orderly sequence" (*Symp.* 215a). Not only is Alcibiades surprisingly coherent, he is also orderly and sequential. He gives a chronological account of his encounters with Socrates in his youth and describes the increasing intensity he felt towards Socrates. He speaks of the experiences he had with him first in the gymnasium, then in private, then during the battles they shared during times of war. When he finishes a point and moves on to the next, the transitions are smooth and chronological, not disorderly: "It was after these events had occurred that we served together" (*Symp.* 219e) and "So much for that incident; but 'what the stout-hearted man' did and endured next on campaign is well worth hearing" (*Symp.* 220c).

In fact, Alcibiades is of such sound mind that at the end of his speech he corrects himself for a point missed, saying, "There is something I forgot to say at the beginning" (*Symp.* 221d). At the beginning of his speech he makes a comparison between Socrates and statues of Silenus,¹⁸ stating that, like them, Socrates has a ridiculous outward appearance but a god-like centre. He is now careful to add that Socrates's "discussions are also very like those Sileni that you open up" (*Symp.* 221d). If his organisation and his ability to remember what he says and does not say at the beginning of

his speech is not enough to suggest that Alcibiades might be sober, Socrates claims at the end of Alcibiades's speech, "I think you are sober after all, Alcibiades. Otherwise you wouldn't have been able to conceal the motive of your entire speech by ingeniously disguising it in this way" (*Symp.* 221c).¹⁹

For those who are willing to admit the mutual stamina of Alcibiades and Socrates and yet still assume the "drunkenness" of the one and the "sobriety" of the other, I think it is significant that we hear from Eryximachus shortly after Alcibiades and Socrates's wine binge. Earlier in the dialogue, he is not concerned with any amount of liquor Socrates may or may not drink because he assumes Socrates's sober character will still stand (*Symp.* 176c). As for Alcibiades, the amount of liquor consumed is paramount; it becomes a question of character: "What sort of behaviour is this, Alcibiades?" he says (*Symp.* 214a). Alcibiades's disruption of Symposium protocol might not be about base consumption; it might be about "a divinely inspired release from normally accepted behaviour" (*Phaedrus* 265a).²⁰ Eryximachus needs to be reminded that "we must not let anyone disturb us or frighten us with the claim that you should prefer a friend who is in control of himself to one who is disturbed" (*Phaedrus* 245b).²¹ Eryximachus assumes Socrates's sobriety and Alcibiades's drunkenness, and Alcibiades is completely aware that his drunken appearance may distract his audience from the truth (*Symp.* 215d). He appeals to the audience's assumption that Socrates will remain sober no matter how much he drinks, therefore should anyone not believe Alcibiades's account of the truth, Socrates is always in a position to discern the truth. Alcibiades says to him: "If I say anything that is not right, Socrates, you must contradict me" (*Symp.* 217b). This deference is explicit in stating that the truth is still more important than appearances. We may assume that Socrates would contradict Alcibiades if he says anything falsely because Socrates says to Alcibiades, "I'll let you tell the truth; indeed I order you to" (*Symp.* 214e).²²

Alcibiades and Socrates: the Art of Speech-Making

Alcibiades has been faulted for making a rather passionate but poor attempt to relate the true nature of Socrates because of his use of images. I will argue that Alcibiades, like Socrates, gives careful consideration to his presentation of the truth.²³ He is only interested in telling the truth when he speaks about the particulars of Socrates's behaviour and how it relates to Silenus statues.

Socrates and Alcibiades's approach to speech-making begins with a fundamental claim that they are going to speak the truth. At the start of his speech in the *Apology*, Socrates states his intention to the jury: "from me you shall hear the whole truth.... Concentrate your attention upon this one question, whether my claims are fair or not. That is the first duty of the juryman, just as it is the pleader's duty to speak the truth" (*Apology* 17b-18a). Likewise, Alcibiades states his intention to speak the truth

at the Symposium: "Yes, I must tell the whole truth; so pay careful attention and, if I say anything that's not right, Socrates, you must contradict me" (*Symp.* 217b).²⁴ Both claim they will tell the "whole truth," both ask for the careful attention of the jurors,²⁵ and both invite interruption if their speeches appear to be "unfair" or "not right."

Both are also aware that the truth is often misunderstood and might appear comical. Defending himself against the charges of corrupting the youth of Athens, Socrates states in the *Apology*: "It is literally true, even if it sounds rather comical, that God has specially appointed me to this city" (*Apology* 30e). Alcibiades opens his eulogy to Socrates as follows: "The way I'll try to praise Socrates, gentlemen, is through images. Perhaps he'll think this is to make fun of him; but the image will be designed to bring out the truth, not to make fun" (*Symp.* 215a). Neither Socrates's claim that he is appointed to Athens by God, nor Alcibiades's recollection that he once saw divine statues in Socrates can be proven as fact to another person. However, both are well aware that the words and images used for describing such truths may be misunderstood.

Alcibiades's drunken appearance and Silenus-statue imagery at once reveal and conceal a truth, therefore they risk being misunderstood. The imagery of Silenus statues and Marsyas the flute-player is an attempt to reveal the truth, namely, that Alcibiades has caught sight of absolute Beauty through Socrates while existing in the temporal world. This significant truth is concealed with the audience's assumption that Alcibiades is concerned with the world of appearances when he speaks of the outward appearance and manner of Socrates (his drinking, sexual interest, and profession of ignorance). Even though he describes his experiences with Socrates in detail, he is *not* concerned with the world of appearances. He says he does not want to say anything that is false, but it is significant that he adds: "if I don't remember things in the right order, don't be surprised. It isn't easy for someone in my condition to list all the aspects of your peculiarity in a fluent and orderly sequence" (*Symp.* 214e-215a). In other words, the particulars, the events, and the order in which they occur do not matter as much as the truth.

Alcibiades has been faulted for using images to describe Socrates's nature and therefore, he has been accused of not speaking the truth, but of merely giving a particular account.²⁶ In Alcibiades's defense, he stands in good company. Both Socrates and Plato make use of images throughout various dialogues. Socrates uses the everyday imagery of "pack-asses, blacksmiths, shoemakers" (*Symp.* 221e) in his attempts to present the truth about abstract topics in his discourses. In the *Phaedrus*, Socrates says that describing the nature of the soul "would be a long tale to tell, and most assuredly a god alone could tell it; but what it resembles, that a man might tell in briefer compass" (*Phaedrus* 246a).²⁷ This tells us that making comparisons between concrete images and the abstract is essential to discourse, a complimentary but complicated relationship between the abstract, language, and sensual experience.

Alcibiades's speech also fulfills Socrates's standard of how imagery should be best

used to reveal the truth. In the *Phaedrus*, Socrates questions Phaedrus about Lysias's speech on 'love' (a disputed or abstract term) and asks whether the concept of 'love' is grounded in a concrete image at the beginning of the speech. If it is not, Socrates claims that the speech fails in its purpose because it does not establish a definition of the subject at the beginning: "Did Lysias at the beginning of his discourse on love compel us to conceive of it as a certain definite entity, with a meaning he had himself decided upon? And did he proceed to bring all his subsequent remarks, from first to last, into line with that meaning?" (*Phaedrus* 263d).²⁸

Alcibiades succeeds where a great orator like Lysias fails. He states at the beginning of his speech that he will praise Socrates through images. By comparing Socrates's nature with Silenus statues and Marsyas the flute-player, Alcibiades fulfills the requirement; a resemblance is established between an abstract term (the divinity of Socrates) and a "certain definite entity" in an effort to define and clarify the abstract point of the speech.

Alcibiades's speech "from first to last" maintains the imagery of Silenus statues and Marsyas the flute-player. The comment that he adds at the end of his speech, that Socrates's *discussions* appear ridiculous on the outside but are, in truth, divine once opened up (*Symp.* 222a), shows that he has brought "all subsequent remarks, from first to last, into line with that meaning." The body of his speech follows the original ridiculous-divine reference, showing continuity throughout. For example, Alcibiades's attempt to describe Socrates's virtue by describing his concrete behaviour in games of seduction and in war continues the point that Socrates appears to be one thing on the outside (ridiculous in his appearance and actions)²⁹ but divine on the inside. Alcibiades is also careful to set limits where comparisons end (again, good speech-making form). Take, for example, his clarification that Socrates is like Marsyas the flute-player, but that he differs from Marsyas since he does not bewitch people with music but with his words (*Symp.* 215c/d). This shows the same careful consideration that Socrates uses in his discourses and it shows that on the surface, both of their arguments appear to make absurd comparisons between the abstract and concrete to provide an account of truth.

It has also been argued that "the power to stun and possess is more important to [Alcibiades] than the ability to distinguish those who yearn for god and initiation."³⁰ The implication is that Alcibiades may have a keen interest in Socratic discourse so that he, in turn, may seduce others with his words. Although the power to stun and possess is significant to Alcibiades, he also knows his audience well, another requirement of good speech-making (*Phaedrus* 271c/d): They are those at the Symposium who "yearn for god and initiation" because they share with him and Socrates "the madness and Bacchic frenzy of philosophy" (*Symp.* 218b). He tells his companions that his experience with Socrates's discourses is analogous to being "bitten by a snake." He goes on to say that "someone who's had this experience is only prepared to say what it's like

to those who have been bitten themselves, because they're the only ones who'll understand and make allowances if the pain drives you to do and say shocking things" (*Symp.* 217e-218a). In other words, he does select his audience; his companions at the Symposium are most able to consider his meaning and understand his point, having had close contact with Socrates themselves. Furthermore, the power to "stun and possess" is significant; it is what an encounter with the divine is like according to the *Phaedrus*, according to Alcibiades's slavish behaviour, and according to mystical traditions. If the power to "stun and possess" is significant to Alcibiades, it might be because an encounter with the divine changed his perception of the world and himself, not necessarily because he values, above all else, its rhetorical effects on people.

Alcibiades is fully aware that he has two options available to him if he is to speak of something so abstract as the divinity he perceives in Socrates. He makes use of both options in an effort to cover his subject to the best of his ability, considering the nature of his experience, and in an effort to relate his experience to his audience. One is a concrete model relating to the world of sense imagery and one is a more abstract model relating to negative statements. He must speak in images or speak in the negative: either way he must speak if he is to tell the truth about his subject (Socrates), a truth that others may not be fully aware of. Therefore, the similarity that Alcibiades draws between Silenus statues, with their hidden god in the centre, and Socrates as having a god in the centre of his being is an attempt to relate an experience *beyond* the world of sense to the world of sense so that it may be better understood. The Silenus/Socrates comparison is not simply meant for persuasion; mere persuasion is not what Alcibiades's speech is all about. Alcibiades's claims to tell the truth are noteworthy since, according to Socrates, it is not a necessity of common orators: "There is, they maintain, absolutely no need for the budding orator to concern himself with the truth about what is just or good conduct, nor indeed about who are just and good men whether by nature or education...but only about what is plausible" (*Phaedrus* 272d). Alcibiades is not this kind of orator in the *Symposium*. The purpose of resorting to the Silenus-Marsyas imagery is to describe the truth of Socrates's being to those who have not had the depth of contact with Socrates that Alcibiades has had, but who nonetheless share the "Bacchic frenzy of philosophy."

For those who are reluctant to accept Alcibiades's account of Socrates's nature because it is related through images or comparisons, it is significant that Alcibiades attempts to forgo images by trying to describe Socrates in terms of what he is *not* like. For example, he says that what is "most amazing" about Socrates is that "he is like no other human being, either of the past or the present...this person is so peculiar, and so is the way he talks, that however hard you look, you'll never find anyone close to him either from the present or the past" (*Symp.* 221c/d). Speaking in negatives is the language often used by those who have perceived something beyond the world of sense. In mystical traditions, it is a common way to relate an experience with the

divine.³¹ However, just as speaking of the abstract through images has its own shortcomings, speaking in negatives does not appear to convey the truth very well either, since it is difficult to understand what something *is* by describing what it is *not*. The point is that Alcibiades's experience of perceiving the divine within Socrates is an experience well beyond the contradictions that exist in relation to images or words; and in this sense, speaking in negatives is appropriate considering the nature of his experience.³² Indeed, Alcibiades is aware of the shortcomings of language and images when he says of Socrates: "The best you can do is what I did, in fact, when I compared him, and his way of talking, not with human beings but with Sileni and satyrs" (*Symp.* 221d).³³

To conclude this section, Alcibiades's appearance and speech may appear to be comical and disorganised, but like Socrates, he intends to tell the truth about his subject and he is careful with this truth. He does not want his audience (or Socrates) to misunderstand his intention, and he is aware that the former, at least, probably will; he knows the limits and shortcomings of trying to place the abstract/divine world within the world of sense and speech. Like Socrates, he uses the world of sense and speech, in my opinion, to the best of his ability; staying consistent with comparisons, stating their limitations, and selecting his audience.

Alcibiades and the Ascent

For his not going beyond the love for a particular individual³⁴ towards a more universal conception of love, Alcibiades has been accused of being unable to make it past the first stage of the ascent.³⁵ Rosen, too, might disagree that Alcibiades is capable of making the ascent, since he concludes that Alcibiades's "fundamental obedience to the body makes it impossible" for him to comprehend the divine.³⁶ Another claim is that "the ultimate aim must be a non-physical relationship....One senses that this is a doctrine that Alcibiades can neither deny nor bring himself to accept; it is too alien to his character, to the formation of his personality," and from this the author concludes that "For Alcibiades, Socrates's rejection of his body could only mean, automatically, the rejection of his entire being."³⁷ However, Socrates's physical rejection does not mean that Alcibiades sees it as a rejection of his entire person, especially when we recall the "degree of understanding" that a young Alcibiades came to recognise and admire in Socrates. The dialogue *Alcibiades I*³⁸ allows more insight into Socrates's and Alcibiades's earlier relationship with one another. Here we find Socrates stating the rules on the "art of love":

One must first seek out the worthy object of love and discern whom one should love...then after deciding to love, not even to speak to him until the critical moment comes when he is ready to listen to

philosophical discourses; and then, when he is ready to listen, take him in hand and teach him the principles of love, and so generate in him a reciprocal love.³⁹

At some point after his initial humiliation, Alcibiades must have made the transition from being preoccupied with physical desires in his youth to becoming a "worthy object of love" and "ready to listen" to Socratic discourse. Otherwise, Alcibiades would not be able to make the claim that he knows Socrates and his discourses more than the others do, a claim Socrates does not dispute. Moreover, Socrates says to a young Alcibiades:

I hope to exert great influence over you by showing you that I'm worth the world to you and that nobody is capable of providing you with the influence you crave, neither your guardian nor your relatives, nor anybody else except me — with the god's help, of course. When you were younger, before you were full of such ambitions, I think the god didn't let me talk to you because the conversation would have been pointless. But now he's told me to, because now you will listen to me. (*Alcibiades* 105e-106a)⁴⁰

This reference to Alcibiades's readiness to listen to Socrates further supports the idea that a non-physical relationship with Socrates is not "too alien to [Alcibiades's] character" or to the "formation of his personality." This passage shows that he is at least capable of and ready to make the ascent.

There is further support in the *Symposium* that Alcibiades has gone beyond the love of a particular body. When he arrives at the Symposium, he garlands both Agathon and Socrates before his speech (*Symp.* 212e-213e). Initially, his intention is to garland only Agathon (*Symp.* 212e) to honour him for being the "wisest and most beautiful" man, but when he discovers Socrates's presence, he insists that some of the garland be shared with Socrates. Although it could be argued that the only reason Alcibiades garlands Socrates is because, as he says, Socrates will criticise him if he does not (*Symp.* 214e), it is nevertheless a symbolic act showing an obedience to the divine in Socrates and a recognition of the divine in both Agathon and Socrates. Despite his specific intention to garland Agathon, Alcibiades no longer references the particular (Agathon) when he shares the garland with both Agathon and Socrates. Secondly, he does not attempt to seduce Agathon or Socrates, despite his supposed drunkenness and Socrates's comment that Alcibiades is "jealous" and "violent" (*Symp.* 213c-d). Instead, the actual portrait we get of Alcibiades at the Symposium is indeed the "opposite" of what Socrates says (*Symp.* 214d); he does not do injury with violent outbursts or acts of jealousy (the behaviour of those with an attachment to a particular person) as

Socrates fears he will. And considering Alcibiades's description of Socrates's bravery (*Symp.* 221b), we should be suspicious of Socrates's fear and the plea that Agathon might need to protect Socrates should Alcibiades become jealous or violent towards him. Such behaviour does not manifest. Instead, Alcibiades quietly disappears from the Symposium amid the noise of an incoming party of people. Finally, Alcibiades's expression of Socrates in the negative, so as to distance him from human characteristics, supports the idea that Alcibiades is not slavishly attached to Socrates the man, but to Socrates the divine. In short, the act of garlanding both Agathon and Socrates, the absence of jealous or violent behaviour, and the description of Socrates in the negative suggests that Alcibiades has no *physical* attachment to a particular body and is able to recognise more than one body as beautiful.

As to the next level of the ascent, Alcibiades is aware that the mind and soul are indeed more beautiful than bodies. The Silenus-statue comparison, among other things, is an obvious reference that the physical beauty of Socrates pales in comparison to his inner beauty, and to Alcibiades, it is Socrates's inner beauty that makes Socrates stand apart from the rest of humanity. It is also what makes Alcibiades a slave to Socrates the divine. Moreover, Alcibiades's desire to become "as good a person as possible" and the "degree of understanding" he has in relation to Socrates's physical rejection of him shows that the cultivation of his soul is more important than the beauty of bodies.

The next level of the ascent is the recognition of beautiful laws and practices (*Symp.* 210c). Historically, Alcibiades is accused of profaning the Mysteries of Eleusis by satirising them in front of non-initiates. This event might be alluded to in the dialogue,⁴¹ suggesting that he has no recognition of beautiful laws and practices.⁴² However, his conviction that Socrates's discourses are "most divine" shows that Alcibiades is aware that there are "practices" for becoming "as good a person as possible." Moreover, Alcibiades's claim that he should do, must do, and has agreed to do what Socrates says shows a recognition of beautiful laws and practices.

Alcibiades's speech also refers to the peculiar customs of Socrates.⁴³ It has been said that Alcibiades refers to Socrates's "customary ways with great frequency"⁴⁴ and the interpretation is that Alcibiades's awareness of Socrates's peculiar behaviour means that he is more concerned with Socrates's outward appearance than the reasons behind such behaviour. But a major point of Alcibiades's speech, if you open it up, is that Socrates has these peculiar customs because he is devoted to, and associated with, the divine. In Waterfield's translation of the *Symposium*, this stage of the ascent is devoted to studying the particular actions of people:

The next stage is for him to value mental beauty so much more than physical beauty....And this in turn leaves him no choice but to look at what makes people's activities and institutions attractive and to

see that here, too, any form of beauty is much the same as any other, so that he comes to regard physical beauty as unimportant. (*Symp.* 210b/c)⁴⁵

This focus on people's activities as expressions of beauty might indicate that Alcibiades's study and description of Socrates's peculiar actions is not as insignificant as Rosen suggests; it is part of the ascent.

Turning to the final stage of the ascent, Diotima says:

Anyone who has been educated this far in the ways of love, viewing beautiful things in the right order and way, will now reach the goal of love's ways. He will suddenly catch sight of something amazingly beautiful in its nature; this, Socrates, is the ultimate objective of all the previous efforts. First this beauty always *is*, and doesn't come into being or cease....Second, it is not beautiful in one respect but ugly in another...nor beautiful here and ugly there because it is beautiful for some people but ugly for others. [Nor will beauty appear to him in the form of a face or hands or any part of the body; or as a specific account or piece of knowledge]....It will appear as in itself and by itself, always single in form; all other beautiful things share its character, but do so in such a way that, when other things come to be or cease, it is not increased or decreased in any way nor does it undergo any change. (*Symp.* 210e-211b)

Alcibiades's glimpse of the divine through Socrates not only relates to the soul's glimpse of the divine in the *Phaedrus*, it relates to this final stage of the ascent as well. We know that Alcibiades "caught sight of something amazingly beautiful in its nature" when he speaks of the statues he once saw in Socrates. The divinity that Alcibiades once saw in Socrates, I will argue, is the Form of absolute Beauty that Diotima refers to.

The Form of Beauty that Alcibiades perceives in Socrates is not related to Socrates's flesh nor to Alcibiades's physical sight; it can only be discerned by the soul. This is what Socrates basically taught Alcibiades — he helped him recollect the divinity of his true nature: "So, my dear Alcibiades, the soul also, if it is to know itself, must look to another soul if it is to know itself, and particularly to the place in it where the virtue of the soul, wisdom, is, and to anything else that is similar."⁴⁶ The soul is the faculty with which the Form of Beauty is known. As such, the soul of Alcibiades is able to discern absolute Beauty because its nature partakes in the divine and his soul recollects its divine heritage while in the temporal world. The nature of this encounter, soul to soul, means that absolute Beauty would not be tied to Socrates's body nor to Alcibiades's physical sight in any way, since these are physical things, whereas the soul

is not physical matter. Furthermore, this soul-to-soul encounter would explain why Alcibiades resorts to images, comparing Socrates to Silenus statues instead of comparing him to other human beings. Comparing Socrates with other humans leads Alcibiades to speak of him in terms of negatives. If Socrates cannot be compared to fellow humans, then the implication is that Socrates has eternal, not human, qualities despite his obvious human form. Since Alcibiades insists that Socrates is like no other human, then the beauty he sees in Socrates is not associated with the flesh and is not discerned by the senses, but by the soul. He makes a clear distinction that what he knows about Socrates is not based on physical qualities that "could be attributed to other people, too" (*Symp.* 221c). It is the soul's discernment that tells him that Socrates is not like anyone "past or present." This distinction implies that the beauty in Socrates is not subject to time, change, death, or even to Socrates's particular actions in the world. Socrates cannot be compared with other humans "past or present" because something in him stands apart, alone; an absolute Beauty "in and of itself." If Socrates stands apart from the rest of humanity, then the beauty (statues) that Alcibiades perceives in Socrates is the absolute Beauty at the final stage of the ascent — unchanging, immortal, not tied to the flesh or human conditions. That Alcibiades's slavish behaviour and speech appear so confusing to us is not because Alcibiades is drunk or physically preoccupied with Socrates or Agathon, but because he is acutely aware of the divine and his audience is not.

Is the Ascent Flawed?

Some interpreters find that Diotima's account of the ascent is too abstract because it does not account for personal Eros. In other words, it does not allow us to love the particular (another individual).⁴⁷ Nussbaum asserts that Diotima's ascent "starkly confronts us with a choice."⁴⁸ The interpretation is that there appears to be an incompatibility between personal Eros (physical attraction towards another) and divine Eros (the soul's attraction to the divine). Vlastos also concludes that a crucial step for the individual making the ascent is the individual's *decision* to move beyond the love of a particular body or not.⁴⁹ I will argue that it is not that Diotima/Plato does not allow for the love of a particular individual, just that the individual body must be seen (and thus loved) in a different way, as an expression of the divine and not as an object of the senses. I will also argue that this new way of seeing a body (or bodies) as an expression of the divine has little to do with an intellectual determination to see a body in this particular light.

Diotima claims that the individual, perceiving correctly with the help of a guide, "sees" that beauty is a quality that all bodies reflect. When beauty in general is discerned, the individual will "relax his intense passion for just one body" (*Symp.* 210b). This seeing and relaxing is a much more passive approach to the ascent than the idea

that an individual is faced with an intellectual choice between personal and divine Eros. Indeed, the language of the ascent suggests that the passivity of the individual is necessary to the ascent. For example, a guide is necessary (*Symp.* 210a-c). This requirement puts the ascending person in a passive role in relation to the guide, which is the nature of Alcibiades's relationship to Socrates. Another example of Alcibiades's passivity and the passivity of the ascent comes after the beauty of minds is recognised. At this stage, one is then "forced to observe the beauty in practices and laws and to see that every type of beauty is closely related to every other, so that he will regard beauty of body as something petty" (*Symp.* 210c).⁵⁰ In Waterfield's translation, the connotation of "force" is the same: the person ascending "has no choice" but to look at beautiful practices and thus come to regard the beauty of bodies as unimportant. Alcibiades feels forced to observe the beauty in practices and laws when he claims that he *must* do what Socrates says. Furthermore, Socrates's words have an effect on him not in an intellectual sense of understanding, but by way of inner understanding. He is overwhelmed; a wave of uncontrollable feeling comes over him when Socrates speaks (*Symp.* 215d).⁵¹ The language of Alcibiades and Diotima suggests that the ascent happens through revelation, not decision; therefore, there is no choice to be made between personal Eros and divine Eros. The passive observation and insight of the soul, not intellectual choice, determines whether one moves past the love of a particular body or not. The ascent appears to be a kind of evolution through perception, an indisputable realisation that one is *forced* to recognise and perceive; it is a disposition that inevitably occurs once the beauty of minds is perceived with the soul.

Diotima tells us that once the beauty of minds, and in turn, the beauty of laws and customs is perceived, non-attachment to particular persons or practices takes place: "Looking now at beauty in general and not just at individual instances, he will no longer be slavishly attached to the beauty of a boy, or of any particular person at all, or of a particular practice" (*Symp.* 210d). If non-attachment seems too strong an interpretation of one no longer being "slavishly" attached to the particular, the qualification "at all" implies a transcendence, a participation in — but a distance from — the world of particulars. This distance does not necessarily mean that we must decide to distance ourselves. It is not that we *should not* love a particular body if we are to make the ascent. It might mean that the love of a particular body *will not* be one's foremost preoccupation and, if so, the ascent may be played out. If the beauty of a particular body is one's foremost preoccupation, possibly even if it is a preoccupation concerned with the decision that one should abandon particular attachments to make the ascent, then one might not be passive enough to let the ascent happen. What is incompatible with the ascent is the "slavish attachment" to the physical beauty of a particular person; such attachment means that the beauty of all bodies will not be recognised. However, once the beauty of all bodies is recognised, then the beauty of the particular body is allowed since it now is understood within this context.

If Diotima's ascent seems too abstract, the belief that one is able to move beyond sensual attractions toward a more universal conception of love is a common one. The idea can be found in Christian, Buddhist, and Hindu teachings.⁵² It may seem too abstract because many of us are mostly unaware of, or question, the divine nature within ourselves and others since we often identify with our senses and sense experiences to the point that we are unable to perceive things from the perspective of our soul.⁵³ The teachings might lead us to believe that sensual attractions have to be abandoned to make the ascent, but perhaps the teachings simply show us that it is possible for the particular to be viewed in a different way, from a different point of view, which, in turn, reshapes how we perceive things from a sensual point of view. The flaw of the ascent might be our perception of it as an intellectual choice or dilemma we are faced with, coloured by the idea that somehow we lose something if we transcend our passion for a particular body. Instead, the ascent shows us that we are able to have a different awareness of the particular, one that arises through the soul rather than the senses. Alcibiades has this different awareness of the particular.

In line with Diotima's meaning, Alcibiades also has a certain distance from or non-attachment to the world of particulars in the *Symposium*. Alcibiades is distanced from his audience when claims to know Socrates and his discourses more than anyone else. He is distanced from his audience when he is unable to adequately express his encounter with the divine. He is distanced from particular practices when he does not follow Symposium protocol (drinking undiluted wine, not passing the jug around). Alcibiades no longer perceives Socrates (and possibly Agathon and the others) with his physical sight but with his soul. As such, he cannot truly distance himself from Socrates (the divine).

The Trouble with the Truth

Socrates accuses Alcibiades of concealing the true motive of his speech, which he says is to make trouble between Socrates and Agathon (*Symp.* 222c/d). Alcibiades says, "I'm warning you, Agathon, not to be deceived by him, but to learn from what we've suffered and be cautious, and don't, as the proverb puts it, be the fool who only learns by his own suffering" (*Symp.* 222b). It has been argued that Alcibiades warns Agathon about Socrates because his motive is to condemn Socrates for rejecting him.⁵⁴ I do not agree that hurt or confused feelings are the motive for Alcibiades's warning to Agathon. Instead, Alcibiades's warning to Agathon is to convey the truth.

Alcibiades's speech and warning to Agathon serves the purpose of telling his audience that the full divinity of Socrates can only be discerned by looking beyond appearances. Otherwise, you face the outer casing, the feigned erotic attraction, the image of Socrates only.⁵⁵ This applies to Alcibiades and his speech as well. If Agathon or the others get caught up in the appearance of either the imagery of his speech or the image

of his person, they might miss the significance of his speech. Is not Alcibiades's caution to Agathon from one who knows to one who does not know?⁵⁶ Is it not the same caution Socrates makes to the young Alcibiades when he tells him to be sure that he is not exchanging "gold for bronze?"⁵⁷

Socrates accuses Alcibiades of trying to cause trouble between Socrates and Agathon, but the latter tells the truth when he says that Socrates deceives other men (including Alcibiades) into "thinking he is their lover and then turns out to be the loved one instead of the lover" (*Symp.* 222b). In *Alcibiades I*, the relationship between Alcibiades and Socrates is explicit in showing that Socrates once played the role of the lover in relation to Alcibiades: "I'm sure you are wondering what I could possibly have in mind — why I don't give up on you? The others [suitors] have all been sent packing, so what do I hope to achieve by persisting?" (*Alcibiades* 104c/d). But in the *Symposium*, we hear that Alcibiades is rejected by Socrates "the lover," and as such, Alcibiades, by pursuing him, becomes the lover to Socrates "the beloved" (*Symp.* 217c). Alcibiades's warning that Socrates might be deceiving Agathon into believing that Socrates is the lover (when the reverse is the case) appears to cause trouble on the surface, but it is a warning that tells the truth: that it would be a mistake to take Socrates or his words at face value. It is a fair warning to Agathon, according to Alcibiades's experience with Socrates.

If we still believe that Alcibiades is just trying to cause trouble, it should be noted that Socrates is often seen in Platonic dialogues as one who causes trouble by pointing out people's ignorance. As uncomfortable as that might be for those at the receiving end of Socratic discourses, it is done with the intention of revealing a truth: that people are often more ignorant than they know or believe themselves to be. Like Socrates, Alcibiades tries to show people that they might not be as perceptive as they think they are with regards to Socrates or himself.

But then what are we to make of Socrates's claim that Alcibiades is causing trouble because he thinks Socrates should love only him and that he (Alcibiades) should love only Agathon (*Symp.* 222d)? Is Socrates saying that Alcibiades is a slave to human attachments? Not necessarily. Earlier in the dialogue, when Socrates claims to be "terrified" of Alcibiades because he has a "mad attachment" to his lovers (*Symp.* 213c/d), Alcibiades says to his audience, "Don't you realize that the truth is quite the opposite of what he said? If I praise anyone else, whether god or human, while he's around, it's he who'll beat me up" (*Symp.* 214d). The claim that Alcibiades has a "mad attachment" to his lovers is mentioned again. This time, it's implied by Socrates's comment that Alcibiades thinks Socrates should love only him and that he (Alcibiades) should love only Agathon (*Symp.* 222d). If the truth is the opposite of what Socrates says with regards to Alcibiades appearing to be madly attached to his lovers, then Alcibiades does *not* think that Socrates should love only him and that he (Alcibiades) should love only Agathon. In other words, he does not have a mad, human attachment to his lovers,

although it might appear so, just as it appears that Socrates is erotically attracted to good-looking boys. His motive is disguised by Socrates's accusation that Alcibiades has a mad attachment to Agathon and Socrates. But if Alcibiades's attachment is a recognition of the divine, then this is what Socrates really means when he talks about Alcibiades's mad attachment. If the truth is the opposite of what Socrates says, then Alcibiades suffers from divine, not human, madness. Alcibiades's motive is to show his knowledge of the divine. What is "ingeniously disguised" is that all that Alcibiades has said about Socrates is true about himself as well. This is summed up nicely by Socrates in the *Phaedrus*:

A human being must understand speech in terms of general forms, proceeding to bring many perceptions together into a reasoned unity. That process is the recollection of the things our soul saw when it was traveling with god, when it discarded the things we now call real and lifted up its head to what is truly real instead. (*Phaedrus*, 249b/c)⁵⁸

Conclusion

It has been said that Alcibiades's "ability to resist Socrates's enchantments is clear evidence that he lacks a genuine comprehension" of the divine.⁵⁹ But Alcibiades is *not* able to resist Socrates's enchantments. Therefore, he does have genuine comprehension of the divine. He says, "If it weren't for the fact that you'd think I was completely drunk...I'd take an oath on the truth of what I'm saying about the effect his words have had on me — an effect they still have now....So I force myself to block my ears and go away, like someone escaping from the Sirens" (*Symp.* 215d-216a). Alcibiades is *still* deeply affected by Socrates's words. Moreover, in the *Apology*, Socrates claims it is his god-given duty to constantly reproach men for their attachments, especially if they profess the desire to rise above attachments:

Are you not ashamed that you give your attention to acquiring as much money as possible, and similarly with reputation and honor, and give no attention or thought to truth and understanding and the perfection of your soul? And if any of you disputes this and professes to care about these things, I shall not at once let him go or leave him. No, I shall question him and examine him and test him; and if it appears that in spite of his profession he has made no real progress toward goodness, I shall reprove him for neglecting what is of supreme importance, and giving his attention to trivialities....This, I do assure you, is what my God

commands....For I spend all my time going about trying to persuade you, young and old, to make your first and chief concern not for your bodies nor for your possessions, but for the highest welfare of your souls. (*Apology* 29e-30b)

According to this passage, Alcibiades, as one who cares about the "perfection of his soul," cannot resist nor escape Socrates. But if this is so, what are we to make of Alcibiades's admission that he gets "carried away by the adoration of the masses"? (*Symp.* 216b) Does this not show that he is pulled towards the human, not the divine, world? Not necessarily. Alcibiades gets carried away by the masses, but to be carried away by the masses is to be a slave to Socrates (the divine). Alcibiades claims to be "runaway slave" to a master who turns up "suddenly wherever I least expect [him]" (*Symp.* 213b). This state of affairs is strongly supported in the *Apology* passage above. If Alcibiades is drawn toward the physical world, Socrates is always there to draw him back toward the ascent. The *Apology* passage also supports Alcibiades's claim that he is telling the truth when he says that Socrates will beat him up if he praises anyone other than Socrates.⁶⁰ This is not someone who has turned his back on Socrates, but someone unable to turn his back on Socrates and his duty to perfect his soul.

Alcibiades's speech has been described as a human account of truth. "It is in fact an adventure story of the lively vicissitudes of virtue under stress."⁶¹ I agree that Alcibiades appears to be an example of "virtue under stress" in contrast to Socrates's effortless virtue in action. However, this does not mean that Alcibiades's speech is "just a love story."⁶² It does not mean that Alcibiades is unfit to make the ascent, lacking "genuine comprehension" or unable to discern the divine. On the contrary, Alcibiades possesses genuine comprehension because he has experienced the divine and he knows it.⁶³

One problem for Alcibiades (and for anyone who aspires to become a better person) is the difficulty of making (and maintaining) the ascent, of cultivating one's divine nature, while existing in the world of senses. The struggle involves trying to discern when appearances reflect the divine and when they do not. This is especially difficult when the divine and the particular intersect through language and images. Another problem is the difficulty of relating this experience to others. The experience is often incomprehensible, and this reflects badly on Alcibiades, although it should not, because the nature of the experience defies expression. To assume that Alcibiades is an example of someone who has not made the ascent, to assume that he has turned his back on Socrates in favour of the world of senses, is a hasty judgement that misses the point. The judgement should not be that either Alcibiades is part of the sensual world or he is part of the spiritual world; the two worlds are not mutually exclusive. Alcibiades, as an example of what it means to make the ascent, illustrates the ongoing struggle between distractions in the sensual world and their impact on the soul's evo-

lution. Alcibiades is involved in that struggle, in that evolution. He has not and cannot turn his back on it. Therefore, I do not doubt that Alcibiades has made the ascent and has caught sight of absolute Beauty while incarnate. Alcibiades has been too intimate with the divine and all other things pale in comparison. This is genuine comprehension. Although it is open to question whether he maintains the ascent, maybe this is not possible or even necessary while one exists in the world of changing appearances. The struggle might not be an indication of weakness of character, but a fact of life in the sense that "The path up and down are one and the same."⁶⁴ In other words, the sensual world is where the ascent is played out — it is the soul's environment for evolution. Therefore, the world of appearances and the divine are not mutually exclusive, but mutually dependent on one another to give an account of the truth.

Furthermore, the character of Alcibiades "ingeniously" disguises yet illustrates the fundamental Platonic doctrine that we should not be unduly fooled or distracted by appearances in our search for truth. Plato makes this point in the *Symposium* by ingeniously distracting us with what he essentially tells us not to be distracted by: the appearance of Alcibiades. The caution that we should not be misled by appearances might mean that, despite appearances, Alcibiades is quite in contact with the truth of his subject, although the audience might not realise this. If the audience concludes that Alcibiades's appearance and presentation of the truth show that he is removed from the truth, such an assumption, in turn, further removes the audience, not Alcibiades, from the truth.

In the end, what Alcibiades offers his audience is proof of the experience with the divine, with absolute Beauty, and with divine madness. Alcibiades, with his professions of confusion and drunkenness and with his imagery of Silenus statues and Marsyas the flute-player, produces a ridiculous portrayal of himself and Socrates. But what is concealed is a divine, sobering message to Agathon (and to us?) not to be fooled by appearances. Likewise, it is a message about what it means to make the ascent and the struggle involved for someone living in the world of appearances. These cautions risk being ignored because we, like those at the *Symposium*, like those at Socrates's trial, sit in judgement, distracted by the ridiculous appearance of Alcibiades — drunk, spirited, beautiful. We fail to notice the true nature of his speech: what it means to know the divine.

Notes

¹ Martha Nussbaum, 'The Speech of Alcibiades: a reading of the *Symposium*,' in *Fragility of Goodness* (Cambridge, 1986), 166.

² All references to the *Symposium* are taken from the following translation unless otherwise noted: Plato, *The Symposium*, trans. Christopher Gill (Penguin, 1999), 58.

³ Daniel E. Anderson, *The Masks of Dionysus: A Commentary on Plato's Symposium* (State University of New York Press, 1993), 116. Anderson says of Alcibiades, "Instead of seeing Socrates's behaviour as a compliment, he sees it as an insult to his youthful charms."

⁴ As when Diotima reproaches him, in his youth, for finding more value in these things than in the cultivation of his soul (*Symp.* 211d).

⁵ Plato, *Socrates's Defense (Apology)*, trans. Hugh Tredennick, in Edith Hamilton and Huntington Cairns (eds.), *Collected Dialogues* (Princeton University Press, 1996), 16. Socrates says: "It is literally true even if it sounds rather comical...God has specially appointed me to this city" (30e). Subsequent in-text citations are to this edition.

⁶ Rosen, *Plato's Symposium* (Yale, 1968), 282. Rosen notes that Socrates "is prevented by his daimonion from abandoning Alcibiades precisely because of the young boy's hybris or pleonexia." Rosen goes on to say that "Extreme hybris, or the desire to be a god, is not simply a vice, but a quality deserving study, and in a sense, of cultivation." Rosen's reference is from *Alcibiades I* 103a5, 104e6 ff.

⁷ See Gill's introduction to his translation of *The Symposium*, *op. cit.*, xv.

⁸ Anderson, *The Masks of Dionysus*, 116.

⁹ Stanley Rosen, *Plato's Symposium*, 290.

¹⁰ Italics are mine. Plato, *Phaedrus*, trans. R. Hackforth (Cambridge University Press, 1952), 92.

¹¹ William James's description of the mystical state supports this point that an encounter with the divine leaves one in a state of "slavish" passivity: "the mystic feels as if his own will were in abeyance, and indeed sometimes as if he were grasped and held by a superior power." From the essay 'The Mystic's Way,' in *Philosophic Problems: An Introductory Book of Readings* (New York: MacMillan, 1957), 564. Originally, James's essay appears as a chapter in his book *The Varieties of Religious Experience*.

¹² Plato, *Phaedrus*, trans. Alexander Nehemas and Paul Woodruff, in John M. Cooper (ed.), *Plato: The Complete Works* (Cambridge: Hackett, 1997), 527.

¹³ Anderson, *The Masks of Dionysus*, 116. Anderson believes that Alcibiades's slavish manner towards Socrates indicates that Alcibiades has "a misunderstanding of Socrates's character and the dialectic."

The fact that Alcibiades cannot adequately describe what he saw will be discussed in the section on speech-making.

¹⁴ Plato, *Phaedrus*, trans. Nehemas and Woodruff, 527.

¹⁵ *Ibid.*, 524.

¹⁶ The wine is supposed to be diluted with "five parts water to two parts wine" according to Gill's note in his introduction, xii.

¹⁷ Also against Symposium protocol; the wine jug is to be passed around.

¹⁸ Alcibiades compares Socrates to the satyrs Silenus and Marsyas. In appearance, satyrs are rumoured to be half man and half beast. They were devoted to drunkenness, debauchery, sexuality, music, and the god of wine, Dionysus. The statues of Silenus Alcibiades refers to have this outward appearance. The dolls contain a god hidden in the centre.

¹⁹ I will discuss what Socrates says about Alcibiades's motive in the section 'The Trouble with the Truth.'

²⁰ Plato, *Phaedrus*, trans. Nehemas and Woodruff, 542.

²¹ *Ibid.*, 523.

²² Incidentally, Socrates and Alcibiades use imperatives with one another: Socrates "orders" Alcibiades to tell the truth and Alcibiades tells Socrates that he "must" contradict him should he speaks falsely.

²³ Nussbaum would probably disagree with this point. She describes Alcibiades's speech as "disorganised and tumultuous," moving "from imaging, describing...back again many times over." *Fragility*, 187.

²⁴ This intention to speak the truth or stand corrected is mentioned earlier at 214e: "If I say anything that isn't true, interrupt, if you like, and point out that what I am saying is false. I don't want to say anything that is false." Note that this claim is slightly weaker than the claim to tell the "whole truth" at 217b. Also, Socrates's choice to interrupt or not at 214e becomes a command to interrupt at 217b.

²⁵ Like Socrates in the *Apology*, Alcibiades addresses those at the Symposium as jurors.

²⁶ Nussbaum, *Fragility*, 185-7.

²⁷ Plato, *Phaedrus*, trans. Nehemas and Woodruff.

²⁸ *Ibid.*

²⁹ The ridiculous outward appearance of Socrates that Alcibiades refers to is Socrates the erotic lover, his stance in combat, his all-day meditation, his going around without many clothes on in winter, his ability to drink without getting drunk, his ability to cheerfully forgo food, and his "comical" speeches which refer to pack-asses and blacksmiths (219e-221b).

³⁰ Rosen, *Plato's Symposium*, 299.

³¹ F.C. Happold, *Mysticism: A Study and an Anthology* (Penguin, 1963 [reprint 1990]), 86. Happold says of divine visions that "Dominating the vision of Transcendence is the recipient's feeling of his own littleness, his complete unworthiness....He is filled with awe and self-abasement....How can It be known? To name It is presumption. It can only be spoken in negatives." This quotation also reinforces the slavish nature of the recipient of a divine vision.

³² William James says: "The handiest of the marks by which I classify a state of mind as mystical is negative. The subject of it immediately says that it defies expression, that no adequate report of its contents can be given in words. It follows from this that its quality must be directly experienced; it cannot be imparted or transferred to others. In this peculiarity mystical states are more like states of *feeling* than like states of intellect." See 'The Mystic's Way,' 563-4. Italics are mine. Note that James's reference to a state of feeling describes the effect that Socrates's words have on Alcibiades: his "tears flood," his "heart pounds" (215e). Although Socrates has this effect on others, according to Alcibiades, I argue that he still stands alone in his recognition that through Socrates he encountered the divine.

³³ Happold notes: "Though the negation may be truer than the affirmation, it is possible and legitimate to make affirmations about the nature of Divine Being. One must, however, be conscious of their limitations and realise that they can never be more than symbols, images, approximations, always inadequate, always in a sense, untrue. When we speak of God as Goodness, Beauty, Truth, or Love, we are not guilty of 'laughable presumption,' provided we realise that what we are actually doing is translating into inadequate word-concepts something which springs from inner experience." *Mysticism*, 64.

³⁴ David M. Johnson, 'God as the True Self: Plato's *Alcibiades I*,' in *Ancient Philosophy*, 19:1, 1999, 2. Johnson is in no way alone in holding this opinion of Alcibiades.

³⁵ Gill comments on page xxxvi of his introduction that "The key contribution" of Alcibiades's speech at the Symposium is the "oblique light that it throws" on Diotima's account of the ascent. He concludes that "Alcibiades would not have understood the mysteries even if he had heard them, since his own conception of love is so different."

³⁶ Rosen, *Plato's Symposium*, 299.

³⁷ Peter H. Von Blanckenhagen, 'Stage and Actors in Plato's *Symposium*,' in *Greek, Roman, and Byzantine Studies*, 33, 1992, 66-7.

³⁸ There is a debate as to whether the early *Alcibiades* dialogues are Plato's or not. For the sake of this paper, I will treat them as Platonic dialogues.

³⁹ This excerpt from *Alcibiades I* is taken from John Dillon's 'A Platonist *Ars Amatoria*,' in *The Classical Quarterly*, 44:2, 1994, 388.

⁴⁰ Plato, *Alcibiades*, trans. D.S. Hutchinson, in Cooper (ed.), *Plato: The Complete Works*, op. cit., 560. Unless otherwise noted, all quotations taken from *Alcibiades* are from this edition.

⁴¹ For example, when he orders uninitiates to put "big doors" on their ears during the latter part of his speech (*Symp.* 218b). Whether this is an allusion to the event is uncertain. Gill notes that the *Symposium* took place immediately before the accusation that Alcibiades profaned the mysteries (see 81, note 137), but it is difficult to tell whether Plato is alluding to this specific event or not.

⁴² Walter M. Ellis, *Alcibiades* (London: Routledge, 1989), 58-61. Ellis also points out

that whether Alcibiades did in fact profane the Mysteries is very much open to interpretation, based as it is on "circumstantial evidence." But in the end he concludes that Alcibiades did, as a "joke in poor taste," profane the Mysteries. As for the later charge of Alcibiades mutilating the herms statues, Ellis concludes that it is "highly doubtful" that he had anything to do with it.

⁴³ See endnote 23 for the list of Socrates's strange customs.

⁴⁴ Rosen's quotation is taken from John E. Ziolkowski's 'The Bow and the Lyre,' in *The Classical Journal*, 95:1, 1999, 19-35. See footnote in Ziolkowski's essay on page 30.

⁴⁵ Plato, *Symposium*, trans. Robin Waterfield (New York: Oxford University Press, 1994), 53-54.

⁴⁶ David M. Johnson, 'God as the True Self: Plato's *Alcibiades I*,' 10. Passage originally taken from *Alcibiades I* 133b7-10.

⁴⁷ Nussbaum, *Fragility*, notes on page 166-7 that this is Vlastos's view. Rosen holds the same view but for different reasons. See Rosen's *Plato's Symposium*, 279.

⁴⁸ Nussbaum, *Fragility*, 198.

⁴⁹ I have taken Vlastos's point from Nussbaum's *Fragility*, 166-7.

⁵⁰ Italics are mine. Plato, *Symposium*, trans. Waterfield, 53-54.

⁵¹ Alcibiades has been faulted for being carried away by his feelings in relation to Socrates. But as I have argued, passivity — not choice — characterises the ascent. See also endnote 26 of this essay for William James's claim that a mystical experience is characterised by feeling and not intellect.

⁵² For an in-depth exploration of the relationship between Eastern and Western philosophy and mysticism, see W.T. Stace's book *Mysticism and Philosophy* (New York: St. Martin's, 1960). Also see F.C. Happold, *Mysticism: A Study and Anthology*, which is quoted in this essay.

⁵³ In Buddhism, individuals have an essential "buddha-nature" beneath the sensual self; in Hinduism, there is a Universal Self behind the sensual self. In Christianity, St. Paul makes the same point: "For you are the temple of the living God" (Corinthians 2:6:16). Diotima and Socrates ask that we recollect our divine heritage as well.

⁵⁴ Rosen, *Plato's Symposium*, 320.

⁵⁵ Ziolkowski suggests that we should visualise "Alcibiades's Silenic simile for Socrates...as a Marsyan/Dionysian figure on the outside, revealing an Apollonian *daimon* inside." ("The Bow and the Lyre," 33). I see Alcibiades with this duality as well.

⁵⁶ Ziolkowski notes that Socrates was once an "inept" pupil of Diotima's: "The pupil eventually becomes an expert who can guide others and help them when they are bitten by philosophy" (34). My argument is that this is the role of Alcibiades in relation to Agathon when he warns Agathon about Socrates.

⁵⁷ Alcibiades's recounting of Socrates's caution not to exchange "gold for bronze" is the only direct quotation of Socrates's words in Alcibiades's speech; otherwise, his speech

is mostly descriptive. See 218d-219a.

⁵⁸ Plato, *Phaedrus*, trans. Nehemas and Woodruff, 527.

⁵⁹ Rosen, *Plato's Symposium*, 300.

⁶⁰ Incidentally, if Alcibiades is a runaway slave, then we might be even more inclined to believe Alcibiades when he says that the truth is the opposite of what Socrates says. Socrates says, "I haven't been able to look at or talk to a single attractive man without his getting so jealous and resentful that he goes crazy and shouts at me and almost beats me up" (213d). This does not sound like the actions of a runaway slave.

⁶¹ Lowry Nelson Jr., 'Alcibiades's Intrusion in Plato's *Symposium*,' in *The Sewanee Review*, 94, 1986, 203.

⁶² Nussbaum, *Fragility*, 187.

⁶³ Rosen, *Plato's Symposium*, 300.

⁶⁴ Heraclitus, Fragment 60, trans. Diels-Kranz, taken from Reginald Allen, *Greek Philosophy: From Thales to Aristotle* (New York: Free Press, 1966), 41.

Sources

Allen, Reginald E. *Greek Philosophy: From Thales to Aristotle*. New York: Free Press, 1966.

Anderson, Daniel E. *The Masks of Dionysus: A Commentary on Plato's Symposium*. New York: State University of New York Press, 1993.

Dillon, John. "A Platonist Ars Amatoria." In *Classical Quarterly*, 44:2, 1994, 387-92.

Ellis, Walter E. *Alcibiades*. London: Routledge, 1989.

Happold, F.C. *Mysticism: A Study and an Anthology*. London: Penguin, 1963 (reprint 1990).

James, William. "The Mystic's Way." *Philosophic Problems: An Introductory Book of Readings*. Ed. Maurice Mandelbaum, Francis W. Gramlich, and Alan Ross Anderson. New York: MacMillan, 1957, 563-576.

Johnson, David M. "God As the True Self: Plato's *Alcibiades I*." *Ancient Philosophy*, 19:1, 1999, 1-19.

Nelson, Lowry Jr. "Alcibiades's Intrusion in Plato's *Symposium*." *The Sewanee Review*, 94, 1986, 197-204.

Nussbaum, Martha. *The Fragility of Goodness*. Cambridge, 1986.

Plato. *Alcibiades*. In *Plato: The Complete Works*. Ed. John M. Cooper. Trans. D.S. Hutchinson. Cambridge: Hackett, 1997.

— *Phaedrus*. Trans. (with introduction and commentary) R. Hackforth. Cambridge: Cambridge University Press, 1952 (Reprint 1997).

— *Phaedrus*. Trans. alexander Nehemas and Paul Woodruff. *Plato: The Complete*

Works. Ed. John M. Cooper. Cambridge: Hackett, 1997.

— *Socrates's Defense (Apology)*. Trans. Hugh Tredennick. *Collected Works*. Ed. Edith Hamilton and Huntington Cairns. Princeton University Press, 1961 (Reprint 1996).

— *Symposium*. Trans. (with introduction and notes) Christopher Gill. London: Penguin, 1999.

— *Symposium*. Trans. Robin Waterfield. New York: Oxford University Press, 1994.

Rosen, Stanley. *Plato's Symposium*. Yale, 1968.

Von Blanckenhagen, Peter. "Stage and Actors in Plato's *Symposium*." *Greek, Roman, and Byzantine Studies*, 33, 1992, 51-67.

Ziolkowski, John E. "The Bow and the Lyre." *The Classical Journal*, 95:1, 1999, 19-35.

McDowell and Kim on Functionalism

Michael Brent, University of Toronto

In his most recent book, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*, Jaegwon Kim offers a complex argument in favour of a functional model for the reduction of the mental.¹ He claims that the traditional model of reduction is problematic, and that we ought instead to be satisfied with local reductions of functionally characterised mental properties. In an important paper, 'Functionalism and Anomalous Monism,' John McDowell argues that, given a certain robust understanding of the rational person, a functionalist characterisation of mentality is out of the question.² In what follows, I will first piece together Kim's argument and explain its main parts, while reserving critical commentary for later sections. I will then explain McDowell's argument against the functionalist characterisation of mentality and show how it applies to Kim's model of reduction.

1. Kim's Functional Model of Reduction

Kim is perhaps most famous for his work concerning the notion of *supervenience*. That notion can be understood as follows:

Mental properties supervene on physical properties in that necessarily any two things in the actual world that are indiscernible in all physical properties are indiscernible in mental properties.³

In the past, Kim has argued that the supervenience relation between mental and physical properties provides a positive account of the way in which minds are related to bodies.⁴ However, he no longer claims that the supervenience relation itself provides a positive account as such. He now argues that supervenience merely characterises the relation between mental properties and physical properties. It is a type of relation that is implied by a variety of mutually exclusive theories concerning the mind-body prob-

lem, and it does not provide a positive answer to the metaphysical difficulty of explaining *how* our mentality is related to our physical nature. Kim maintains that the notion of "physical realisationism" explains exactly how mental properties supervene upon physical properties. Physical realisationism is the claim that whenever a mental property is realised or instantiated in some system, it is done so physically. This claim provides the foundation for the functional model of reduction. We can identify three main parts to the model: (a) a particular conception of physicalism; (b) a particular conception of mental properties and intentional content; (c) the functional model itself. Let us examine each in turn.

1.2 Physicalism: According to Kim, physicalism is the common and pervasive multi-layered ontological model of the world that stems from the natural sciences. In that model, the bottom level⁵ is thought to consist of the basic bits of matter out of which all material things are composed. As we ascend upwards through the levels, we encounter atoms, molecules, cells, living creatures, planets, etc., and the various properties that characterise each type of entity. This mereological ordering relation generates a micro-macro hierarchy of ontological levels: entities and properties at any given level can (in principle) be exhaustively decomposed into entities and properties that belong at any other lower level.

Kim urges us to recognise an important distinction between the micro-macro hierarchy of ontological *levels*, and the relationship between *first-order* and *second-order* properties. He claims that the micro-macro hierarchy of ontological levels must be kept conceptually distinct from the ordering relation among properties. In particular, he argues that the realisation relation does not track the micro-macro relation. The reason is that "*both second-order properties and their first-order realisers are properties of the same entities and systems....[F]or something to have a second-order property is for it to have one or another of its realisers, that is, a first-order property satisfying the specification that defines the second-order property.*"⁶ Consider the following example.⁷ The macro-level property of "being a water molecule" is the micro-entity-based property of having two hydrogen atoms and one oxygen atom in a certain bonding relationship. Although the properties and relations of its micro-entity-based parts fully constitute this macro-level property, it is nevertheless a property had only by the macro-level entity. In this case, then, the property of "being a water molecule" is a molecular property that is not had by any of its individual constituent atoms.⁸

1.3 Mental Properties and Intentional Content: Kim's conception of a mental property is the functionalist conception, whereby a mental property is specified by its extrinsic role⁹ as a causal intermediary between sensory inputs and behavioural outputs, and not by the constituents of the system in which it is realised. To use a common example, for an organism to be in pain, says Kim, "is for it to be in some inter-

nal state that is typically caused by tissue damage and that typically causes groans, wincing, and other characteristic pain behaviour.¹⁰ The mental property of "being in pain" is thus a second-order functional property defined in terms of the causal relations amongst the first-order properties in which it is realised: system x has the property of "being in pain" if and only if x has some first-order property P such that when x is exposed to tissue damage P causes it to exhibit behaviour that is characteristic of a system that is experiencing pain. Now, the second-order functional property of "being in pain" can be realised across multiple systems, so long as the system has some first-order property P that causes it to exhibit characteristic pain behaviour when exposed to tissue damage. According to Kim, such functional properties are thus abstract in nature and we need not conceive of them as properties in and of themselves. Rather, such properties are merely defined over the local system in which they're realised and do not necessarily refer to an actual property that can be realised across multiple systems.¹¹

Kim's functionalist conception of mental properties also involves a certain understanding of intentional content. Intentional content is content that is *about* or *directed at* some object or event. For example, one might have a desire to eat a piece of chocolate cake. The intentional content of that desire is about or directed at the chocolate cake that one wishes to eat. Kim says that intentional content is evident in such propositional attitudes, and that

[if] we are asked to design a structure that can perceive, process information received through perception, store it, make use of it to make inferences and guide action, and so on, we don't necessarily have to make a replica of a system that is known to do these things....Arguably we have already done so, in certain *robots* and *computer-driven devices*.¹²

Since we know that robots and computer-driven devices are not conscious, we can see that Kim has divorced intentional content from conscious experience.¹³

1.4 Functional Reduction: Kim argues that this conception of mental properties allows for a functional reduction of the mental. The notion of functional reduction can be understood as follows: a particular second-order functional mental property M is first specified by its extrinsic role as a causal intermediary between a sensory input and a behavioural output.¹⁴ Since the causal powers of the mental property *are* the causal powers of its physical realiser,¹⁵ we can identify that mental property with its particular physical realiser on that occasion P_1 , in virtue of the identical causal roles that they play. The mental property M can thus be understood as a second-order functional property concept that characterises a set of first-order properties that can be instantiated in the system.

Consider functional reduction in the context of the following mental property R : my perceptual belief *that it is raining*. According to Kim, we can specify that mental property by its extrinsic role as a causal intermediary between a sensory input and a behavioural output. Next, we can identify R with its particular physical realiser in the following way: system x has R if and only if x has some first-order property P such that when x has a certain sensory input, P causes x to exhibit a certain behavioural output characteristic of R . In the current case, we can understand "system x " as a particular person: me.¹⁶ The first-order property P must, therefore, be some complex neurobiological property involving my central nervous system. The causal powers of my mental property R are thus nothing but the causal powers of my complex neurobiological property P . As a result, R is a second-order functional property concept that characterises the causal relations amongst my neurobiological properties.

We can conclude this section with a brief summary. Kim's argument begins with the claim that whenever a mental property is realised it is done so physically, thereby accounting for the supervenience relation between mental and physical properties. His functional model itself involves three main parts. First, the model occurs within the physicalist ontological framework stemming from the natural sciences. Second, it utilises a particular conception of mental properties as causal intermediaries between "sensory inputs" and "behavioural outputs," and an understanding of intentional content as divorced from conscious experience. Finally, the notion of functional reduction involves the claim that a particular instance of a mental property can be identified with its physical realiser on that occasion in virtue of the identical causal roles that they play.

2. Mental Properties Are Properties of Persons

For a moment, let us grant Kim that a particular realisation of a mental property is a second-order functional property identical with and thus functionally reducible to a particular realisation of a complex neurobiological property (which is itself a property of a particular person). How exactly are we to understand this identity claim? Recall the mental property R : my perceptual belief *that it is raining*. Suppose that my belief is a second-order functional property identical with and functionally reducible to some complex neurobiological property. If my belief is indeed reducible in this manner, then it appears that it is also identical with and similarly reducible to some complex property of my central nervous system.¹⁷ And if this, too, is the case, then it appears that my belief is further identical with and functionally reducible to some complex property spread out over my body. Finally, if my belief is some complex property of my body, then my belief is ultimately a property of a particular person.

Now, we can anticipate a possible objection to this thought.¹⁸ That objection might go as follows: it seems odd to say that a mental property is functionally reducible to some complex property of a particular person, because we can imagine that certain

parts of the person might be removed while the mental property remains the same. As a result, the objection continues, we must find some essential part of the particular person that is necessary for the maintenance of the mental property. Since that part is commonly thought to be the brain, it can't really be the case that a mental property is functionally reducible to some complex property of a particular person. Rather, a mental property must be some complex property that ultimately resides in the brain.

However, in claiming that a mental property is a property of a particular person, we need not deny that the brain is essential to the production and maintenance of mental properties. It is perfectly clear that if one were to cut off my head and throw it in the garbage, I would no longer believe anything at all. But brains do not function in isolation from their bodily and external environments. A bad science-fiction story might include an experiment in which a human embryo is successfully raised in total isolation from the world, with all bodily movement restricted during development. It seems plausible to think that the embryo could not develop into a child that would have the ability to form and express beliefs with intentional linguistic behaviour. If it did somehow reach that stage of development (this is a *bad* science-fiction story, after all), we can assume that the child will be severely developmentally retarded.

This provides us with an appreciation of why Kim points out that mental properties are complex properties had by human persons.¹⁹ In order for a person to believe something about the external environment, she must have a properly functioning body that includes the relevant perceptual apparatus, a central nervous system, and a brain. Now, according to McDowell, once the notion of a person enters into our discussion, we can include a robust conception of rationality that both governs our understanding of the intentional actions and utterances that persons engage in and constitutes the inner workings of thought in general. If we hold on to that conception of rationality, says McDowell, then a functional reduction of the mental is out of the question. Let us examine his argument.

3. McDowell on Functionalism

As a preface to McDowell's argument, consider the following quotation from Donald Davidson:

We explain a man's free actions, for example, by appeal to his desires, habits, knowledge, and perceptions. Such accounts of intentional behaviour operate in a conceptual framework removed from the direct reach of physical law by describing both cause and effect, reason and action, as aspects of a portrait of a human agent.²⁰

McDowell claims that a functional account of the propositional attitudes is excluded

by "the constitutive ideal of rationality."²¹ The constitutive ideal of rationality is a conceptual framework that guides our explanation and understanding of the intentional actions and utterances of other persons.²² It allows us to claim that persons are essentially rational agents — that is, they have the ability to form beliefs on the basis of rational reasons.²³

We can understand the strength of McDowell's claim as follows: Consider the notion of "deductive rationality." Deductive rationality is a capacity that is instantiated to a certain degree of perfection within different rational individuals. In particular, it is a capacity to hold beliefs both when and because they follow deductively from other beliefs that one holds. McDowell claims that this capacity should be understood as an active exercise that consists in an approximate understanding of a normative structure. The understanding that one has of that structure enables one to determine what follows from what, and, hence, what ought to be believed on the basis of other beliefs, for deductively connected reasons.²⁴

Consider the following example. Suppose a person comes to believe *that the ground is likely dry enough for a picnic* on the basis of her perceptual belief *that it is sunny outside*.²⁵ With the constitutive ideal of rationality firmly in place, we can say that her understanding of the normative concept of, say, probabilification explains why she holds the latter belief on the basis of the former. It is important to emphasise that her *understanding* of that normative concept is an active capacity that constitutes the inner workings of her rational thought in general.

This presents a problem for Kim. Recall that he conceives of functionalisable mental properties as causal intermediaries between "sensory inputs" and "behavioural outputs." In the context of the above example, Kim could only claim that that person came to believe *that the ground is likely dry enough for a picnic* on the basis of the causal efficacy of her perceptual belief *that it is sunny outside*. This causal relationship between one belief and another loses sight of the fact that a person's understanding of a normative concept is an active rational ability. Instead, it conceives of a person's understanding as passive, in the sense that the relationship between beliefs is the result of causal interaction among those beliefs and not a positive effort on the part of the person.

Now, suppose that a person comes to believe *that the ground is likely soaking wet* on the basis of her perceptual belief *that it is sunny outside*. If we maintain our grip on the constitutive ideal of rationality, then we can say that her *misunderstanding* of the normative concept of probabilification explains why it is irrational for her to hold the latter belief on the basis of the former. According to McDowell, the kind of explanation and understanding of the intentional actions and utterances of other persons that we can attain with the constitutive ideal of rationality is *sui generis*. It is unique because it allows us to recognise that reasons for action, belief, and so on can be causes for action, belief, and so on, but in a distinctive manner, insofar as a belief is caused by a

person's *understanding* of a normative concept (as it was applied to a particular perceptual belief). This kind of causal relation is unique because it involves a person's understanding of normative concepts and rational relations. Unfortunately, Kim could not claim that the latter belief was irrational on the basis of the former, for the concept of irrationality does not operate amid merely causal relations between beliefs. A causal relation as such does not concern normative relations; it does not concern what one ought to do, believe, and so on.

There is a second part to McDowell's claim. He argues that, in addition to a person's understanding of the normative structure, the capacity of one belief to stand in a rational relationship to another depends necessarily upon the intentional content of those beliefs. As a result, the normative structure that one has grasped cannot be abstracted away from the relations that exist between the intentional content of the beliefs that one holds. We cannot, in other words, capture that normative structure within the relations of a system described in non-intentional terminology.²⁶ In the first example above, the intentional content of the perceptual belief was about the outside environment; that is, the relevant condition of that environment played a necessary role in that person's understanding of the relationship between her two beliefs. As a result, that content cannot be eliminated from our explanation of the rationality or irrationality of the relation between her two beliefs.

This second part of McDowell's claim has devastating consequences for Kim's functional conception of mentality. Recall that Kim conceives of intentional content as divorced from conscious experience. He claims that robots and computer-driven devices have the ability to utilise intentional content to process information and make use of it for the guidance of action. Such abilities, however, should not be confused with the rational capacities of a conscious person. The kind of perception that conscious persons enjoy is very unlike the loose kind of "perception" that we can attribute to non-conscious robots. Our perceptual experience presents us with an environment that is constituted by particular objects and events. It presupposes an understanding of the distinction between subject and object (even if that understanding is not explicit during experience), it is permeated by our linguistic and conceptual resources, and it includes an understanding of normative concepts and rational relations. As such, perceptual experience involves a very strong kind of intentional content. Following McDowell, with the constitutive ideal of rationality firmly in place, we can recognise that the conscious experience that persons enjoy involves this strong kind of intentional content. Unfortunately, Kim's understanding of intentional content does not capture the kind of intentional content that is distinctive of our unique conscious experience of the world.

This leads to further difficulties for Kim's functional conception of mentality. Kim argues that phenomenal qualities of conscious experience are most problematic for the functional model of reduction.²⁷ This argument, however, is cast under the assump-

tion that intentional content can be attributed to non-conscious systems and hence divorced from conscious experience. McDowell claims that one cannot separate intentional content from conscious experience.²⁸ To see if that is indeed the case, we must consider whether our conscious experience involving intentional content ever *lacks* phenomenal qualities. Consider my perceptual belief *that the sun is setting*. The intentional content of this particular belief is phenomenally loaded. I can perceive *that the sun is a particular shade of orange*, or *that the temperature is dropping*, or *that the crickets are chirping*, and so on. Such phenomenal qualities are an essential aspect of the intentional content of my perceptual belief, and any explanation of that perceptual belief must therefore account for their presence. In the case of the conscious experience of a rational person, it seems difficult to claim that phenomenal qualities are not an essential part of the perceptual belief that one has about that very experience. Perhaps the perceptual experience that persons have is necessarily phenomenally loaded? If this is true, we can deny that intentional content can be separated from conscious experience, and we can claim that phenomenal qualities are, thus, an essential aspect of the intentional content of perceptual beliefs. In so doing, we can recognise that both the intentional content and phenomenal qualities of the perceptual beliefs of conscious experience will fall through Kim's reductionist net. And since many of the intentional states that persons have are related to their conscious experience of the world, Kim's position runs into serious trouble.

Let's conclude this section with a brief summary. Kim's functional conception of mentality is problematic for a number of reasons. First, it is a causal account of mentality. As such, it cannot make sense of the thought that normative concepts function within a framework whose occupants stand in rational relation to one another. It is only from within the ambit of such a normative framework that one belief can be understood as rationally justified in the light of another. If we want to insist that rationality is indeed an essential hallmark of mentality, then the functionalist conception becomes problematic because it loses sight of rationality. Second, Kim's conception is a passive account of mentality. As such, it cannot make sense of the thought that a person's rational or irrational understanding of the normative framework is an active capacity that can cause action, belief, and so on in a unique manner. Third, the normative framework itself that one has grasped cannot be abstracted away from the relations that exist between the intentional content one enjoys. In other words, we cannot capture that normative framework within the relations of a system whose occupants are described in non-intentional terminology. As a result, Kim's understanding of intentional content as divorced from conscious experience ignores the fact that intentional content itself is inextricably tied to a person's phenomenal experience of the world about her. Given that Kim's functional conception of mentality runs into difficulty when attempting to account for phenomenal qualities, it will also encounter problems when explaining intentional content that is tied to perceptual experience. It

appears, then, that Kim's functional model of reduction fails as an account of the relationship between mental states and the person enjoying them.

4. McDowell on Physicalism

Donald Davidson's notion of the constitutive ideal of rationality plays an essential role in McDowell's argument. The constitutive ideal of rationality is a conceptual framework that guides our explanation and understanding of the intentional actions and utterances of other persons. It allows us to claim that persons are essentially rational agents. Kim, on the other hand, does not entertain the possibility of such a robust conception of persons as essentially rational agents. What might lead Kim to ignore the constitutive ideal of rationality and our nature as essentially rational persons?

According to McDowell, a particular aspect of physicalism is at the root of the problem. Recall that Kim understands physicalism as the common and pervasive multi-layered ontological model of the world that stems from the natural sciences. From within this metaphysical framework, Kim claims that entities and properties at any given layer can (in principle) be exhaustively decomposed into entities and properties that belong at any other level. McDowell argues that the fundamental issue within physicalism is the notion that all of reality is objective, such that it can be described from no particular point of view.²⁹ He claims that this notion stands in tension with our natural intuition that mentality is both real and characterised by an essential subjectivity. The reality and essential subjectivity of the mental stems from the fact that it is persons (i.e. human subjects) who enjoy mental states. Understood in this manner, we can say that mental states are always *someone's* mental states, and are thus partially constituted by that person's unique perspective or point of view of the world about them. If we maintain that mentality is essentially subjective, the idea that reality is describable from no point of view is very problematic.³⁰

McDowell argues that the notion that reality is objective places pressure on us to try and force mentality into an objective framework. That pressure, McDowell says, is one that we do not necessarily have to succumb to. McDowell claims that it is mere prejudice to think such a model exhausts our ontological options. We can be comfortable with the thought that mentality is both real and characterised by an essential subjectivity if we let go of the presumption that the understanding of the world ascertained via the natural sciences is the only legitimate account of what exists. In other words, we need not think that the ontological model of the world proffered by the natural sciences is *the exhaustive* account of reality.

5. Conclusion

We have seen that Kim's functional account of mental properties as extrinsic causal

intermediaries between sensory inputs and behavioural outputs is problematic for a number of reasons. It is a causal and, hence, passive conception of a person's capacity for understanding, and it loses sight of the fact that persons are essentially rational agents. It engenders a conception of intentional content as divorced from a person's conscious experience of the external world, and it fails to recognise the strong kind of intentional content that persons enjoy. According to McDowell, Kim's commitment to physicalism might hinder an appreciation of Davidson's notion of the constitutive ideal of rationality. With that ideal firmly in place, we can recognise that persons are essentially rational agents and maintain that mentality is both real and characterised by an essential subjectivity. And by letting go of the thought that the ontological model of the world proffered by physicalism exhaustively accounts for all of reality, we can be comfortable with the recognition found through the utilisation of that Davidsonian ideal.³¹

Notes

¹ Jaegwon Kim, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation* (Cambridge: MIT Press, 2000).

² John McDowell, 'Functionalism and Anomalous Monism,' reprinted in his *Mind, Value, and Reality* (Cambridge: Harvard University Press, 1998), 325-340.

³ Note that the concept of *indiscernibility* plays a crucial role in this definition of supervenience. As Kim recognises, we can also define it with respect to the idea of a *difference*, as in, "no mental difference without a physical difference" (see Kim, *Mind in a Physical World*, 10).

⁴ *Ibid.*, 11: "It has sometimes been thought — as I myself did at one point — that the answer is yes, that what might be called supervenience physicalism is a possible position to take on the mind-body problem."

⁵ As Kim points out, this model assumes that there exists a bottom level. See *ibid.*, 15 ff. 23.

⁶ *Ibid.*, 82 (italics in the original).

⁷ *Ibid.*, 84.

⁸ This will be related to mental properties below (§1.3).

⁹ It should be emphasised here that Kim is not very clear on what exactly is meant when we say that a property is "extrinsic." He claims that properties have extrinsic roles insofar as they are *relational* properties of those individuals that have them (see page 21, especially ff. 35). The contrast here is with non-relational or intrinsic properties, presumably those properties that do not bear a relation to *anything* external to the individuals that have them.

¹⁰ *Ibid.*, 19.

¹¹ This relates to Kim's distinction between *levels* and *orders* in the following way: the second-order functional property of "being in pain" and the first-order property *P* are properties of the same system, and both are realised at the same ontological level. Hence the second-order functional property of "being in pain" can be understood instead as a *concept* that characterises a set of first-order properties that can be instantiated in some local system.

¹² *Ibid.*, 103 (italics added).

¹³ This is an important point and will be expanded on below (§3).

¹⁴ According to Kim, this is to "functionalise" the mental property. He claims those properties that resist functionalisation will also resist capture within his particular version of reduction. I hope to show that many mental properties resist functionalisation and, hence, reduction. See below (§3).

¹⁵ The identification of the causal powers of *M* and *P_j* is referred to as "the causal inheritance principle." See pages 54-5 for a discussion of the principle. Unfortunately, an examination of it is beyond the scope of this paper.

¹⁶ At one point, in reference to supervenient psychological properties, Kim says, "In the latter case the physical base properties, presumably certain neurobiological properties, are at the same level as the psychological properties: they are both had by human beings and other sentient creatures" (*ibid.*, 117). In our above example, then, the "system *x*" is a human being and the first-order property in which *R* is realised is presumably some neurobiological property *P*.

¹⁷ Given that this is a belief about the external environment, its formation requires that the subject possess the relevant and properly functioning sensory organs, as well as the physiological infrastructure that serves to connect those organs with the brain. Since there is nothing inherent in the concept of a mental property that forces us to restrict the spatial location of some such realisation *to the brain*, we can claim that it is a property of the person and doing so follows naturally from the bodily involvement.

¹⁸ There is a second possible objection. One might argue that the extra-neuronal bodily involvement in a perceptual belief shows only that (a) perceptual beliefs have complex *causal* conditions, (b) that more peripheral parts of the body and central nervous system are but part of the cause of any given perceptual belief, and (c) that such beliefs are thus ultimately a property of the brain. While it is true that such beliefs have complex causal conditions, one need not view the relationship between, say, a particular belief about the external environment, the person's bodily conditions, and the environment itself as constituted by a progression of causes. Rather, the relationship that a person has with her environment can be thought of as *rational* (or *irrational*). This is a point that McDowell raises and one that will be expanded on below (§3).

¹⁹ It is important to point out that Kim himself does not present this argument, but he does clearly say that psychological properties are properties of human beings (see ff. 16 above).

²⁰ Donald Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), 225.

²¹ McDowell, 'Functionalism,' 325-40. For the notion of "the constitutive ideal of rationality," see Davidson, *op. cit.*, 223.

²² This is similar in vein to Daniel Dennett's point that propositional attitudes come into view only from within the context of the intentional stance, which enables us to capture the patterns that are distinctive of our unique way of living (it is "similar" because Davidson is a realist about the propositional attitudes, whereas Dennett might be an instrumentalist). In other words, the concepts of *believing that X*, *desiring that Y*, and so forth, are intelligible only given a postulate of rationality. See Daniel Dennett, *The Intentional Stance* (Cambridge: MIT Press, 1987).

²³ This does not mean persons are solely rational agents. It leaves room for, say, emotions.

²⁴ McDowell likens the active abilities of the understanding to the Kantian notion of "spontaneity." See John McDowell, *Mind and World* (Cambridge: Harvard University Press, 1998). At one point (page 182), he characterises Kantian spontaneity as "the freedom that consists in potentially reflective responsiveness to putative norms of reason."

²⁵ As it stands, this is not a deductive argument. To render it a deductive argument, we would have to include a third belief about the relationship between it being sunny outside, on the one hand, and the conditions of the ground, on the other. For example, we might say she believes *that when it is sunny outside, the probability that the ground is dry increases* (or something to that effect).

²⁶ This claim, McDowell says, is consistent with the fact that there is no mechanical test for logical validity in general (see 'Functionalism,' 327). The validity problem is the question of whether a proof of the conclusion from the given premises exists. The precise statement of the validity problem is: "There is no algorithm (i.e. a computer program) that, given an arbitrary argument in first-order predicate logic, will tell you whether it is valid or not." The validity problem is distinguished from the problem of proof-checking, for one can program a computer to verify that a putative proof is in fact a proof. I am indebted to Alasdair Urquhart for clarification on this matter.

²⁷ *Ibid.*, 102.

²⁸ McDowell phrases this point as follows: "The substance, and not just the form, of a propositional attitude's representational bearing on the world contributes intelligibly to the attitude's explanatory powers" (*ibid.*, 334). In other words, the explanatory capacity of a propositional attitude is not independent of its bearing on the world, where "bearing on the world" is related to the empirical "substance" that one perceives in one's conscious experience of the world.

²⁹ *Ibid.*, 336-9.

³⁰ Indeed, it seems difficult to imagine how a "point of view" can be taken seriously if it is *not* from someone's perspective.

³¹ I would like to thank Prof. William Seager and an editor from this journal for some very helpful comments made to an earlier version of this paper.

Sources

Davidson, D. *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

Dennett, D. *The Intentional Stance*. Cambridge: MIT Press, 1987.

Kim, J. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge: MIT Press, 2000.

McDowell, J. "Functionalism and Anomalous Monism." reprinted in *Mind, Value, and Reality*. Cambridge: Harvard University Press, 1998.

McDowell, J. *Mind and World*. Cambridge: Harvard University Press, 1998.

The Value of 'Authentic': A Defence of Authenticity in Music

Carl Joakim Gagnon, McGill University

In the following, we shall be concerned with finding a viable definition of "authentic"¹ performance and with seeing whether authenticity, so defined, is a valuable artistic goal. With a view to focus, I will ignore further, non-aesthetic benefits of authentic performance. I will also be deliberately vague about the exact relation between the benefits of authentic performance and other aesthetic benefits, because my focus is on finding a definition and defending its value. Moreover, if this value is supposed to be substantial, I will need to show that other, especially aesthetic, values do not usually override it. Finally, I am primarily concerned with discussing authentic performance as a living alternative to "conventional" performance. So even if it makes aesthetic sense to keep traditions of authentic performance alive in order to maintain a "benchmark" against which to measure the qualitative progress of our performances, my ultimate defence of authenticity will need to go beyond this. Such justifications provide no reason for which authenticity belongs outside the museum; I want to argue that it does.

The latter sentence touches on why I think musicians and people interested in music should care about this apparently rarefied topic. Educated musicians aware of historical scholarship and playing original instruments have produced some of the finest, freshest, and most exciting classical recordings of the last decade.² However, most people have scoffed at the idea that such music is in some sense "authentic." There is an obviously naive usage of loaded terms like "authentic": my gut-stringed violin is not some fabulous keycard to Bach's spirit. But I think there is also an interesting and non-naive understanding of authenticity that can help explain the artistic success of so-called authentic performance. This may have implications for other forms of music, too. Since Kathleen Brennan co-authored the melody of the song *Anywhere I Lay My Head* with Tom Waits's guttural voice in mind, there might be a sense in which this song will sound best when it is played by Waits and his band. But popular music has a weaker tradition than European classical music in performing structurally complex and through-composed works, where exact and authentic reproduction of a

composer's conception can conceivably be an overriding artistic goal. In improvised music, the work is often not even detachable from the performance of it, at least not in any interesting way. Who can see John Coltrane studying Richard Rodgers's handwritten score of *My Favourite Things* in order to enrich his performance of the piece? It should be kept in mind, then, that this paper concerns classical music, and that its conclusions cannot be thought to apply indiscriminately to all other kinds of music. This is a topic for future research.

Following James O. Young in his article 'The Concept of Authentic Performance,'³ we might say that there are at least two basic questions to be asked with reference to authentic music. First, is authentic performance an attainable artistic ideal? And second, if it is, is this ideal worth attaining? Young answers these questions with a resounding no. I will argue later that it is possible to give a "strong" response to Young's challenge, by showing a clear sense in which a realisable performance can be called *completely* authentic, yet remain an eminently valuable artistic goal. Aside from this, however, there is a "weaker" response. Whether complete authenticity is attainable or not, a performance can surely be *more* or *less* authentic, in relation to whatever ideal standard our definition of authenticity provides (we shall explore such standards shortly). In this case, the first question collapses into a version of the second: should we value *more* authenticity in performances?

This collapse will take place as long as we have some possibility of approaching the ideally authentic (set of) performance(s), as determined by our "standard of authenticity." Two considerations are worth noting. First, we leave the possibility open that the ideally authentic realisation of whatever we take to be realised (e.g. the score, an earlier performance, or the composer's intentions) might be a *set* of performances, rather than a single performance. Second, a certain minimum of authenticity must be realisable if other aesthetic concerns are not always to override the concern for authenticity. If we manage to find some definition of authentic performance that satisfies these conditions, we will have effected the "collapse" of question one into question two. If this understanding of "authentic" still seems to involve a valid artistic ideal, we will have met Young's challenge and successfully defended the value of authentic performance.

Both Young and his opponent, Stephen Davies,⁴ discuss various possible definitions of authentic performance. One such — call it (1a) — states that an ideally authentic performance reproduces music as it was heard at the time of its composition. Suggestion 1a is easily refuted by considering unperformed works, since it seems as if, on this definition, unperformed works cannot be performed authentically at all, and there is no obvious reason why unperformed works should have a different status in this respect than performed works. Perhaps we can revise (1a) to (1b): "Ideally authentic performances reproduce music as it was or would have been heard at the time of its composition." Proposition (1b) will certainly effect the "collapse": *something* can be

said about how most compositions would have been performed at the time of their composition. But it fails in describing an attractive artistic goal. The quality of the instruments and musicians was, for much of Western musical history, often very low by modern standards. The same applies to other important aspects of performance. The time and care that went into the production of a piece like Mozart's *Don Giovanni* were minuscule, for instance.^v Why would we want to relive *that*?

Perhaps (1b) fails because it does not make any reference to the intentions of the composer. Mozart must have *wanted* better singers and fewer mistakes. Proposal (2) therefore runs: "An ideally authentic performance is one that sounds as the composer intended it." Clearly, if we take *all* the composer's intentions to be equally relevant, we end up in immense epistemological *and* practical difficulties. How are we really to know these intentions? Can a few remarks by Beethoven, if they seem to evince a change of heart, be enough to radically alter the score of one of his symphonies? To avoid such difficulties, we might limit the relevant intentions to those publicly expressed in a score. In the above example, Mozart might have wanted better singers in order to make fewer mistakes in *the performance of the score*. As Davies points out, this might still be too inclusive, since not all information about the composer's intentions given in the score is usually determinative for an authentic performance. Instead, what types of information should be taken as determinative depends on the musical tradition the composer is part of. Fairly exact tempo indications might, for example, be taken as determinative in much 20th-century music but only regarded as a recommendation in early-19th-century music. Our proposal (3), then, runs: "An ideally authentic performance is one that faithfully executes the determinative intentions publicly expressed in a score by a composer."⁶

Proposal (3) works better than proposal (2) because it differentiates more clearly between the composer's intention generally and those that are relevant to the work; not just any change in the composer's intentions will translate into a change in the work. Furthermore, proposal (3) recognises that as a musical work becomes the possession of the public, it ceases to be that of the composer. This is a gradual and usually slow process, but after a work becomes sufficiently a part of the public domain, a composer's intentions no longer affect the work — an attempted revision then creates a separate work. If Steve Reich were to demand today that his magisterial "Drumming" (1971) be performed on bagpipes, his demand would quite simply not be accepted, although a separate work might be created.

Despite the public nature of some such works, I think that it is still reasonable to regard the composer as having some sort of "privileged access" to her own works, sometimes even after they have been published. We accept certain of J.S. Bach's revisions of the *Well-Tempered Clavier*, even when scribbled into student copies and not the autograph score itself.⁷ Even more tellingly, Milan Kundera can quite justly praise conductor Charles Mackerras for playing many of Czech composer Leos Janacek's operas in

ways that seriously diverge with the score accepted by Janacek himself, since Janacek was bullied into revisions for commercial reasons.⁸ We may accept Mackerras's version as more authentic because it complies more closely with the determinative intentions of the composer as set out not only in the score, but also in letters, diaries, etc. We must, of course, expect very unambiguous evidence from the composer in order for other statements of intention to override those found in the score. The score, after all, is usually our best guide to the composer's intentions. This method is not as simple as just accepting determinative intentions as they are expressed in a score, but probably closer to the way we usually think about the "standards of authenticity" against which a performance's faithfulness is to be measured.⁹

If I am right about this, then all the determinative intentions need not be submitted "through" the score, although they need, of course, be clearly and publicly expressed. We should also note that "publicly" here need not mean that the intentions were actually expressed publicly, but only that they could be publicly accessed. A composer might have worked on a score privately and then shelved it. It is with all these reservations that we must reread proposal (3): "An ideally authentic performance is one that faithfully executes the determinative intentions publicly (in principle) expressed in (usually) a score by a composer."

Proposal (3) seems to give a plausible account of one "standard of authenticity." It does not, as noted above, define authentic performance *per se*, but only an ideally authentic set of performances. As such, it certainly represents an attainable artistic ideal whereby performers can reasonably aspire to a certain minimum of authenticity; on first glance, there seems nothing unreasonable in aspiring to be *completely* faithful to the determinative intentions publicly expressed in a score, since these intentions under-define the performance considerably. In this case, our question one will collapse into question two: should *more* authenticity in a performance be valued? What *is* the value of "authentic"?

Stephen Davies gives one plausible answer. According to Davies, "a performance is better for a higher degree of authenticity (other things being equal) *whatever* the merits of the composition."¹⁰ Yet, the value of authentic performance lies in the actual and accurate *transmission* of the work. Therefore, the reason authenticity as such is seen as valuable is that musical works tend to have features *worth* transmitting.

We could add to Davies' account that, although we take it to be true that an authentic performance will, other things being equal, always be better than an inauthentic one, this does not, interestingly, have to mean that the value of an authentic performance is completely independent of the value of the work performed. Davies claims that "in any particular instance, authenticity in performance is valued independently and irrespective of the aesthetic value of the work itself."¹¹ There is a good argument to be made, however, that it is aesthetically more important to have an accurate performance of one of J.S. Bach's masterpieces than, for example, one of J.K.F.

Fischer's works, or one of Bach's lesser works.¹² In his *Violin Concerto in A Minor*, Bach exhibits something like full control over musical forms, styles, and effects. Such a musically accomplished work is one for which it is more important to transmit accurately, for reasons analogous to why an extremely carefully worded poem is even less likely to be improved by a misprint than a casually written one.¹³ Less musical "content" and less achieved musical form means, bluntly, that there is less to be lost. The reason it does not immediately occur to us that Bach's works are actually more in need of authentic performances than Fischer's might be that the strength of Bach's works are often such that, even in their altered form, many of their qualities "shine through" — especially formal or "architectonic" ones. But this does not mean that the margin to Fischer's works would not be even greater when played authentically — really, it implies the opposite.

All this does nothing to weaken the central thrust of Davies's argument: for a given work, an authentic performance as described above will always be better than an inauthentic performance, other things being equal. It does mean, however, that we should be especially concerned to see the works we care most about performed authentically. If so, we have shown that on our definition (3) of authenticity, first, question one can be collapsed into a version of question two, since a performance can aspire to at least *some* authenticity; and second, that question two ("should we value more authenticity in performances?") can be answered in the affirmative. Indeed, if all this is so, our loudest yes to question two should be heard in precisely those instances where our most cherished works are concerned.

If all this is so — for it is now time to bring a crucial assumption out in the open. What underlies and guarantees the value of authenticity is the fact that it communicates a work to a listener. But as we have defined it, it might seem as if all authenticity proper can do is to faithfully execute the determinative intentions of a composer by vibrating air in a certain way. It says nothing about whether the same sounds (i.e. the same vibrations in air) are *heard* in the same way by listeners; that is, it does not consider whether the musical communication between performer and audience breaks down. This communication is crucial, recall, because it "underlies and guarantees the value of authenticity" as we defined it: an authentic performance is valuable because it communicates properties of a work to a listener.

James O. Young raises an objection somewhat like this in the article mentioned above.¹⁴ Young argues that *just vibrating air* is not an attractive artistic ideal, because he cannot see what artistic goals it would attain. In one sense, the objection is, I think, good: if authenticity, on our reading, is to be a valuable pursuit, this certainly seems to require actual *communication*. Such considerations seem to be crucially overlooked in Paul Thom's reply to Young, published two years after Young's article.¹⁵ Thom wants to take "the *work* rather than any kind of performance as the reference point against which authenticity is to be judged."¹⁶ For him, authenticity involves complying with

the set of specifications that make up the work. Simply carrying out the directives of the scores might be an attainable ideal, but it is certainly not clear that this would be a valuable artistic goal if the music did not "get through" to listeners more or less as the composer intended. Conductor Roger Norrington's comment, which Thom cites in his defence, seems only to reinforce Young's point: "We find that the more accurately we reflect such gradations, the better the music sounds."¹⁷ If Thom cannot give a convincing account of the value of this "sound," then we must grant Young's point: if the goal of authentic performance is to faithfully communicate what the composer has "to say," we cannot only worry about following directions. We must also worry about how our performances *sound* to a contemporary audience.

It is worth noting that our definition (3) might have the conceptual resources to deal with such a demand, while Thom's might not. For we can plausibly claim that a composer surely intends for her work to be heard "properly," that is, so as to understand the musical "content" of the work. So players must aim at actually communicating the composer's conception and not simply at producing a set of acoustic events. Whatever the argument, Stephen Davies seems to agree. Davies explicitly agrees to the following characterisation of his view offered by Young: "An authentic performance makes a piece sound as it would have sounded at the period of its composition, had conditions been ideal,"¹⁸ where "sounds" clearly means not just vibrations in air, but what one subjectively *hears*. Can we plausibly hold that definition (3) is attainable if we accept this new requirement for communication?

The crucial question is whether Young's objections to Davies's proposal stand. Young tries to show that "older" (the boundaries are left undefined) musical compositions cannot sound to us as they did to their contemporaries. Young claims, plausibly, that beliefs and attitudes shape the way we hear. This allows him to introduce the distinction between "hearing" and "hearing-as," which forcefully makes the point that not every vibration of air is heard in the same *way* by two listeners, especially if there are large differences in "belief-systems." We might have *knowledge* of a wide variety of 18th-century beliefs that presumably shaped their musical appreciation, but how can we have the 18th-century experience without actually *holding* those beliefs?

This will not fully refute Davies's account. Even perfect contemporaries have different beliefs. Yet we still allow that contemporaries from different "historical" (e.g. different generations), geographical, or social backgrounds can come to form similar judgements about performances and, through them, about works. These judgements are not taken to be simple expressions of personal, idiosyncratic experience.¹⁹ This is presumably because, first, there are beliefs that we do not think significantly alter how we hear music or understand our aesthetic experience of it; and second, even where there is an "aesthetically relevant" difference in beliefs, different beliefs can be sympathetically appreciated. Nothing mysterious is meant by "sympathetic appreciation" here. It simply refers to our ability to appreciate differences in context when passing

aesthetic judgements, as when we understand and appreciate the piety of a 15th-century work without ourselves believing in the God it glorifies.

On Young's account, however, passing the same general set of aesthetic judgements on performances actually seems to require having the same beliefs (I speak of "aesthetic judgement" and not "experience," unlike Young, to emphasise that we are only talking of that part of experience that is relevant to aesthetic judgement):

By listening to a great deal of their music we share some of their experiences and come to have some of their beliefs — but not all of them....Our inability to share their beliefs makes it impossible to hear music of earlier periods as it was heard.²⁰

Perhaps it is correct that just knowing (in the sense of "being aware of the fact") that an 18th-century listener heard this sound as such-and-such will not make me have the same type of aesthetic experience as her.²¹ From this it does not follow, however, that I actually need to *adopt* her beliefs from the same aesthetic judgements — if this were true, any person's aesthetic judgements would have no relevance outside his unique and personal "belief-box." So Young's argument is about *degrees*. It must claim that, after a certain point, beliefs and attitudes become so different that it is impossible for us to sympathetically entertain them, and that this is the case in the appreciation of "early" music.

As such, the argument is both unconvincing and impossible to wholly refute, but the following considerations bear on it: (1) In order to refute it, we need not claim that we have the same experience as earlier listeners, since only beliefs that regard aesthetic judgement or experience are relevant. (2) In order to refute it, we need not even claim that we have the same *aesthetic* experience as earlier listeners, as we do not have the same aesthetic experience as our contemporaries. (3) For a Western listener, earlier parts of the Western musical tradition's approach to harmony, pitch, rhythm, and melody are far closer to "home" than those of, for example, different parts of 20th-century non-Western music; cannot this broadly shared "attitude-base," plus historical understanding, produce reasonably similar aesthetic judgements? (4) The broad musical palette of the modern listener indicates that it is fully possible to entertain "contrary" aesthetic beliefs about, for example, harmony, as applied to different *kinds* of music: a dissonance that would have sounded jarring in Mozart, and absurd in Palestrina, sounds pleasant in the music of Thelonius Monk. (5) Finally, we only need to reject enough of Young's account that authentic performance can *usually* also communicate the "substance" of the work better to the listener, and that this is *usually* a very important aesthetic goal.²²

If we accept this, then we can clearly reach at least a very substantial degree of authenticity using Davies's definition above, so this definition surely gets something

right. But aside from the fact that we might never experience a *completely* authentic performance, so construed, there are plenty of problems. The most prominent is the following: Davies's above account emphasised that the standard of authenticity was not simply air vibrations, but the way these were heard by contemporaries. But which contemporaries are we talking about? The general educated public? As Young points out, "much great music was poorly regarded by early listeners."²³ Moreover, composers might not like the way their performances were heard and played in their time.²⁴ Perhaps authenticity should be defined relative to the way the composer would have heard the piece? Then authenticity becomes almost impossible to measure, unless we speak of determinative, publicly "available" intentions — but then we give up talk of *sound* altogether. More generally, even if these intentions could be known, to whose modern "hearing-as" should we compare it? Could a performance be authentic for one person, but inauthentic for another, because the latter did not hear it as the relevant portion of the composition's contemporaries did?

These considerations force us to conclude that we cannot define authenticity in terms of how a performance sounds *to* some audience. But Young's arguments, above, showed us that it is central to the value of "authenticity" that a performance does sound as it might have done to the composer — that it communicates, and doesn't simply vibrate air. We said above (following Davies) that the reason authenticity as such is seen as valuable is that musical works tend to have features worth transmitting. In response to the problems raised above, we can now add that the fact that authentic performances allow for the possibility of transmitting aesthetic features of the work also makes authentic performance as a whole a valuable endeavour. In order to realise this potential, however, the audience has work to do. As we have seen, actual communication relies heavily on sympathetic appreciation, which does not come for free: whatever else it demands, it certainly requires patience, knowledge of, and, of course, sympathy for the beliefs and attitudes of a different time.

This also offers us some attractive conclusions, for it means that pulling out the harpsichord is not necessarily the best way to play Bach for all audiences. Authentic performance does, however, communicate maximally to that part of the general audience who are better trained in "Baroque listening." Moreover, since "how it sounds" guarantees the value of the project, but does not enter into the actual definition of authentic performances, this also means that the fact that a performance does not at all sound to an "untrained" audience like it sounded to a contemporary of the composition does not make the performance less authentic.

Therefore, we can stick with our account of authentic performance as faithfully adhering to the determinative, publicly expressed intentions of a composer. Before concluding, however, we should consider whether this actually justifies the peculiar choice of instruments and interpretations of the authentic-music movement. After all, perhaps many musicians performing on modern instruments and in modern ways can

say that they are "just reading the score," especially if the score does not specify instrumentation.

If there were such a thing as reading the score *simpliciter*, however, it would certainly not always be identical to reading modern scores. Insofar as the sound a musician produces today is determined by the score, this involves relying on a set of interpretative conventions that differ from various earlier periods in more or less arbitrary ways. Long-lived conventions still in use today include the rule that a notated accidental (flat or sharp) applies to all subsequent tones of the same pitch within a measure, but do not include little *sforzandi* at the beginning of musical phrases or certain rules for embellishment. Yet, what the composer intended to be read off the score has not changed. "Just reading the score," therefore, means reading the score as it would have been read under the interpretative conventions of the time.²⁵

This conclusion implies that authentic instruments are required, since Bach and his audience had no notion of modern ones. Moreover, since the composer's determinative intentions are what matter, and since the composer could not intend to use instruments she did not know of, authenticity requires using only those kinds of instruments the composer was acquainted with, even within her own epoch. The limits must be loosely drawn, however. Although the composer had never heard a particular make of flutes, one could still use these flutes for performance, as long as they remain the same *kind* of flute. Contrariwise, a work composed in the early 1830s by a composer who had not yet heard the new Boehm flutes (whose sound was significantly different from earlier flutes) might rather — other things being equal — be played more authentically on an older flute.

Authentic-performance methods (including both instruments and certain ways of manipulating them, e.g. hand-crossing) are not, then, to be adopted because of the way they sound but because the score "said so," either explicitly or implicitly. (The advantages of not adopting performance methods simply because of the sound they produce are laid out by Jerrold Levinson.)²⁶ However, if the instrumentation of a work can sometimes be implicit, so surely can the kind of clothes performers wore, the price of admission, etc. Hence, "because the score said so" is not quite an adequate description of our standard. Rather, we refer back to the score — and, to fully understand the implicit conditions of the score, the era — for guidance in those matters that are generally thought to be of aesthetic importance. This guideline does not raise any questions about the relative merit of authentic and modern performance methods and authentic- and modern-sounding performances. It is simply a matter of settling on what part of the score's explicit and implicit instructions are generally agreed to have real aesthetic significance, which is a point we have no reason to suppose will divide "authenticists" and "non-authenticists." Both presumably agree that a certain piece played on harpsichord, for example, sounds different in an aesthetically relevant way from the piece played on piano, even if they disagree as to whether this difference is

for the better or worse.

Intuitions about what is of aesthetic importance can change, of course. Although the exact shape, size, and colour of the performance location does not usually matter much to most moderns for musico-aesthetic reasons (except acoustics; and it might matter for comfort, or for other "pretty-surroundings"-type aesthetic reasons), we might imagine an alternative world where this is taken to be crucial. Performance methods might therefore be adopted, even if they make no difference to the sound of the performance, when we have only intuitive aesthetic reasons to do so (as Levinson's argument suggests). Comfortably, however, nothing crucial hinges on this: if our aesthetic intuitions were to change — or if we simply "misread" them — our notion of what goes into authenticity might change, too.

Our proposal (3) needs to be slightly revised to take account of how many aesthetically important features of a score require using the same standards of interpretation to *read* them as the composer needed to *write* them. Embellishments — not written out only because the composer expected them to be played out naturally — are only one example. These standards are usually those of the composer's time, and are adopted only implicitly, since she usually knows that any deviation from these standards will have to be explicitly addressed in order to be understood. Our proposal (3') runs: "An ideally authentic performance is one that faithfully executes the determinative intentions publicly expressed in a score by a composer, or implied by the standards of musico-aesthetic interpretation implicitly adopted by the composer."

This seems to answer our questions one and two — *can* and *should* performance be authentic? A performance certainly can be authentic if authenticity is regarded as a property that comes in degrees. On our definition, it also seems as if authenticity can be substantially and perhaps completely attained, as definition (3') significantly underdetermines the sound of the performance. Other possible conditions, such as performance methods, can more easily be satisfied. The possibility of these conditions also shows why authenticity should not be defined in terms of air-vibrations, while a host of other problems arising from "hearing-as" show why authenticity should not be defined in terms of sound. However, we have seen that the possibility that educated, sympathetic listeners would appreciate something of the same aesthetic "content" as the contemporaries of its composer through the same performance events is crucial for the value of authentic performance. By incorporating this into the "value-theory" underlying our definition of authenticity, instead of into the notion of authenticity itself, we avoided some of the problems Davies seemed to face. This does not give us a complete account of the values of performance *per se*, or how these interrelate. But it should do something to show the substantial value of *authentic* performance in communicating a work. And it would be strange, after all, if the aesthetic content as shaped by a Bach or a Mozart were to be regularly improved upon by the whims and fancies of modern conductors and performers — even when judged by modern ears.

Notes

¹ For convenience, I will use the expression without quotation marks from now on, although I recognise that "early-music performers now almost always enclose it in quotation marks to indicate that they know better than to use the term naively" (B.D. Sherman, heading 'Authenticity in Music,' in *Encyclopaedia of Aesthetics*, 1998). As mentioned in the succeeding paragraph, I will, however, argue that a non-naive interpretation of authenticity is possible.

² Examples come by the dozen, but British violinist Andrew Manze and his partnerships with groups like the Academy of Ancient Music and Romanesca have been among the most widely recognised. The recognition and awards stretch from those internal to the classical community to Grammy nominations and Amazon.com's Classical Artist of the Year designation.

³ J.O. Young, 'The Concept of Authentic Performance,' in A. Neill and A. Ridley (eds.), *Arguing about Art* (New York: McGraw-Hill, 1995), 74-84.

⁴ Stephen Davies, 'Authenticity in Musical Performance,' in Neill and Ridley (eds.), *Arguing about Art*, 62-73.

⁵ See Davies, 'Authenticity,' 65, and Young, 'Concept,' 76-77.

⁶ The phrase, substantially, is found on page 393 of J. Levinson, 'Authentic Performance and Performance Means,' in *Music, Art and Metaphysics* (Ithaca: Cornell University Press, 1996), 393-408.

⁷ C. Wolf, *Johann Sebastian Bach - The Learned Musician* (New York: Norton, 2000), 382.

⁸ Milan Kundera, *Les Testaments Trahis* (Paris: Gallimard, 1993), 226-27, 296-298.

⁹ Franz Kafka represents an interesting problem in this regard, however. His last "revisions" to a large part of his work, including his three great novels, was a demand for destruction. Are we right to say here, with Max Brod, that the earlier, undestroyed "version" of the work is preferable, and that even clearly specified instructions for "revisions" are to be disregarded? Kundera would say no (cf. Kundera, 329; although Kundera admits that he too might not have "trouvé la force d'obéir sans restriction au 'testament' de Kafka," 330). To some extent, this hinges on the limits we choose to draw in distinguishing between private and public possession of artworks.

¹⁰ Davies, 'Authenticity,' 72.

¹¹ *Ibid.*, 71-72. The rider "in any particular instance" is meant to show that the value of authenticity in music *as such* is valued because works tend to be.

¹² This is, of course, intended as an argument about the relative value of *accuracy* in the two performances — my argument is not that a performance of "good" Bach is aesthetically more valuable than a performance of "bad" Bach.

¹³ This is a slight emendation to an example given by Davies on page 376 in *S*.

Davies, 'Authenticity in Performance — A Reply to James O. Young,' in *British Journal of Aesthetics*, 28, 1988, 373-376.

¹⁴ See Young, 'Concept,' 82-83. Young also wants to claim that defining authentic performance in terms of vibrations of air does not even amount to an *attainable*, far less artistically attractive, goal. His arguments for the attainability of authenticity so defined seem tenuous. For our present purposes, it is suffice to say that defining authenticity by degrees, as we have, avoids any such problems.

¹⁵ Paul Thom, 'Young's Critique of Authenticity in Musical Performance,' in *British Journal of Aesthetics*, 30, 1990, 273-276.

¹⁶ Thom, 'Young,' 273. Author's emphasis.

¹⁷ *Ibid.*, 274-75.

¹⁸ Young, 'Concept,' 79. Cf. Davies, 'Reply to Young,' 373: "James O. Young criticises a view which he correctly attributes to me...according to which an authentic musical performance 'makes a piece sound as it would have sounded at the period of its composition, had conditions been ideal.'"

¹⁹ Davies makes this point in 'Reply to Young,' 374.

²⁰ Young, 'Concept,' 80-1.

²¹ We should also note that having the same *type* of aesthetic experience need not necessarily mean under-appreciating a composer who was misunderstood by her contemporaries. The composer, *ex hypothesi*, must have had the same type of aesthetic experience as her contemporaries (since Young's argument rests on contemporaries, of the same general sort of background, having a certain general type of aesthetic experience), but certainly need not have disapproved of her own works.

²² Perhaps it does not strike us as immediately obvious that authenticity is simply playing "as written." I discuss and try to justify this notion below.

²³ Young, 'Concept,' 81.

²⁴ Peter Kivy makes a related point in 'On The Concept of The "Historically Authentic" Performance,' *Monist*, 1988, 279-290.

²⁵ At least usually. There might be special instances where the composer disagreed with the performance practices of his times or where she was not aware of them in detail. The ultimate goal is, of course, to read the score as the composer intended it to be read. In general, however, if the composer wished that the performer would deviate from common practices, she would presumably have specified this clearly in the score.

²⁶ Levinson, 'Authentic Performance and Performance Means.'

Sources

- Davies, S. "Authenticity in Musical Performance." *Arguing about Art*. Eds. A. Neill and A. Ridley. New York: McGraw-Hill, 1995.
- "Authenticity in Performance — A Reply to James O. Young." *British Journal of Aesthetics*, 28, 1988, 373-376.
- Levinson, J. "Authentic Performance and Performance Means." *Music, Art & Metaphysics*. Ithaca: Cornell University Press, 1996.
- Kivy, P. "On the Concept of the 'Historically Authentic' Performance." *Monist*, 1988, 279-290.
- Kundera, M. *Les Testaments Trahis*. Paris: Gallimard, 1993.
- Sherman, B.D. "Authenticity in Music." *Encyclopaedia of Aesthetics*, 1998.
- Thom, P. "Young's Critique of Authenticity in Musical Performance." *British Journal of Aesthetics*, 30, 1990, 273-276.
- Young, J.O. "The Concept of Authentic Performance." *Arguing about Art*. Eds. A. Neill and A. Ridley. New York: McGraw-Hill, 1995.
- Wolf, C. *Johann Sebastian Bach - The Learned Musician*. New York: Norton, 2000.

Amor Fati: Consciousness in Nietzsche's Gay Science

Mark Sentesy, Carleton University

In *The Gay Science*, Friedrich Nietzsche tries to sketch a way of living in the face of nihilism.¹ By outlining the conditions of his society's nihilism, he mounts a radical attack on epistemology and metaphysics, but also prepares a solution. He argues that there is no "reality" in the traditional sense: there are no "things" (110). For us, reality is made up of appearances, of surfaces that we constitute. Language and communication, on the one hand, and consciousness, on the other, are key tools in Nietzsche's critique of epistemology and of the solution that he offers. The insight that relates these tools to one another is that consciousness is not individual, but developed under the pressure of the need to communicate. This will serve as a lens through which we will try to glimpse a cluster of ideas at or near the centre of Nietzsche's thought. Nietzsche's critique is so radical that it challenges the possibility of both knowledge and consciousness. On the other hand, Nietzsche wants to refine a higher form of consciousness: *amor fati* ("love of fate"), a "yes" to reality as such. I argue that this problem can be resolved by distinguishing two kinds of consciousness. Nietzsche says that consciousness penetrates and merges with its object, but this cannot yield self-consciousness, and so does not make consciousness possible. A prior shape of the mind, however, allows us to grasp what is impenetrable and unique about the object, enabling us to remain outside the object. I argue that this is the meaning of Nietzsche's metaphor of distance, in which his project of incorporating consciousness into the human creature, of "life as a means to knowledge,"² intertwines with the idea of writing one's life.

To prepare the ground for Nietzsche's critique of consciousness, and to show to what extent it undermines the idea of truth, we must outline his idea of communication. We will discuss the relationship between names and essences, the genealogical critique of truth claims, and the dynamics of conversation.

The way that Nietzsche uses language in general shows us where the tip of his criticism lies. As he says in Section 381, his style is abrupt and sudden; he swoops down

on an insight and is back in the air clutching the rabbit in his claws. If we were to sit and brood on an insight, wonders Nietzsche, what would happen? We would find it familiar and tolerable; we would grow sleepy and become forgetful, even comfortable. Nietzsche's non-systematic writing is a call to awake! For him, what are the dangers of sleep, then? Why avoid nearness, familiarity, forgetfulness, and comfort?

Naming

Nietzsche goes to considerable effort to show the extent and character of our relationship with "reality." He shows traditional concepts of reality to be untenable. Writing to the Realists, who think that they perceive reality, he says that everyone is like "an artist in love. And what is 'reality' to an artist in love?" (57). An artist, that is, like Romeo, who finds the stars, pilgrims, gods, and angels in Juliet's figure and grace.³ The world is of radical value to him. By expressing it, he enchants us by filling it with treasures. To be sure, they are metaphors, but they are tangible: he paints realities. We think the sound of "G-O-D" has the weight of reality itself, or more. Juliet is but a particular bundle of flesh and skin, but is worth *dying* for. Nietzsche means that if we consider the human world from a non-human perspective,⁴ all our thoughts seem to be arbitrary metaphors. Why not use the word "canoe" for both canoe and paddle? Why divide a place into "lake" and "shore"?

Nietzsche shows that how things are named constitutes reality.⁵ In Section 58, he describes how, once assigned, the name, reputation, appearance, and value of, say, a tree, grows into the tree and, for us, become its very body. Once we forget the origin of these aspects of the tree's appearance, they become its very essence. We see in the tree only the significance that our ancestors assigned to it. Nietzsche's point is twofold: (1) what we experience is a world of meaning, and (2) that the meaning in this world is human. Since we constitute "things," all essences are like clay in our hands. Seeing how the meaning of our tree originated in our ancestors does not extract the meaning they put into it: we still see the tree in the same way, but now our eyes are sceptical. Scepticism never abolished any meaning; it only made it look strange and problematic. The inherited, meaningful world remains before us like a veil, but a dark one. By changing the veils over a thing, in the long run we change the "thing itself"; the appearance of the tree *to us* — its name, behaviour, and value — is the tree itself. Naming is a term that we can use to describe giving a veil to something, specifying a value, appearance, and behaviour — in short, giving a thing its essence. But one must see it, or see in it a new thing, before giving it a new name. Once named, the new thing becomes visible to others (261). Nietzsche's position, then, is not that there are no things (he is not a nihilist in this sense), but that what things really are, independent of humanity, is irrelevant to us and, moreover, is impossible for us to know. That is his scepticism. But if human nature is immutable, it could be that naming has a natural

or proper order in relation to humanity. Natural convention could reintroduce a kind of grounded truth and morality.

However, Nietzsche says that people's characters are utterly plastic as well: "Here a large mass of second nature has been added; there a piece of original nature has been removed" (290). We can add and remove pieces of ourselves, conceal what is ugly or make it sublime, and generally use it wherever it would be suitable. Even if our drives are consistent, we can shape how they operate and how they appear. For example, in science, the will to truth is turned against itself and negated by destroying the belief in Truth (God) (344). The same applies to character, for the trait we call recklessness could be moulded into religious fervour. A sickly constitution could be turned into an ascetic virtue or used as a tool of scepticism. —

This analysis of human nature considerably deepens Nietzsche's criticism of names. Not only would the names we give to things be arbitrary if we could consider the things themselves, but the names are arbitrary among people, since "human nature" is plastic. It follows that we can find no fixed or natural conventions for naming and, hence, for knowledge. Rather than nature determining how we name things, "it is enough to create new names and estimations and probabilities in order to create in the long run new 'things'" (58). This observation, according to Nietzsche, is substantiated by how we experience the world.

In Section 166, Nietzsche says that we create our own world so totally that, when a man experiences anything, he responds to and remembers *only* what he constituted. Furthermore, in Section 127 he says that even what a man experiences as pleasure and displeasure depends on his mind's (mostly unconscious) interpretation of it. A man can remove some, but not all, of his presuppositions from the reality he experiences, but he "cannot take anything out of things that did not belong to [him] before; what belongs to others remains behind" (242). Other things constitute one's world, besides oneself: other people, history, drives, or the fact of experience. But history and drives, as we will see, are inseparable from the "self," so we could interpret them as something that belongs to oneself — as long as one takes possession of it. How to take possession of oneself is a central and very personal issue for Nietzsche. But first we must see how he understands the self.

History

Nietzsche might philosophise with a hammer, but he swings it with style and technique. *A Genealogy of Morals* is the classic example of Nietzsche's technique of historical deconstruction, but he performs the same tricks in *The Gay Science*.⁶ His basic technique is to portray a particular mindset or doctrine — in short, a particular will to power — by a history of its origin. This history contradicts, shames, and undermines the dignity of the will to power as it is in the present.

In his attack on the common definition of knowledge, he locates its value and its character in a detestable, irrational origin. Nietzsche says that what both common people and philosophers "want when they say 'knowledge' [is for] something strange...to be reduced to something *familiar*" (355). The motive underlying their need for knowledge is fear of what is foreign. They believe that everything predictable, everything that submits to the rule of the herd, is safe. To be known is to be predictable. But what we know is only what the herd believes to be useful (354). In the same parable, he criticises this kind of knowing, saying, "What is familiar is what we are used to; and what we are used to is most difficult to 'know' — that is, to see as a problem; that is, to see as strange, as distant, as 'outside us'" (355).

It is as though Nietzsche asks every speaker, "What makes you say that? What *deceives* you into believing that it is *necessary*? Why do you need it to be true?" Nietzsche disassembles all truth claims by unveiling their history, or, more specifically, by locating their origin in an irrational drive. What we call truth and knowledge are defined by our drives, which, although irrational and usually unconscious, use reason as a tool to accomplish their ends.

In Section 34, Nietzsche says that the view and drives of every great person relativises the whole of history. How are we to understand this? He means that each individual⁷ who sets eyes upon history, or becomes part of it, recasts and rediscovers it.⁸ This is the same process as renaming a tree: if the great person recasts history, it is recast in the mould of his interpretation of reality.

Nietzsche uses his hammer both to knock down and to build up. All criticism is already laden with value and, in a sense, already offers the solution. *The Gay Science* is no exception. Nietzsche attacks certain buildings, as it were, burns some, and topples others. The destruction of a city in a war leaves rubble, but we build over the ruins and use their materials to build with, in the same way that we transform human nature. Nietzsche uses his criticism to transform the human creature. For his positive doctrine to come into view, however, we must show how Nietzsche conceives of distance.

Conversation

Based on his critical techniques, we can characterise Nietzsche's critique of epistemology as a critique of communication. He reinterprets conversation by including its ambiguities, omissions, and mysteries, its lies and truths and their motives, its deliberate and accidental deafness, blindness, and forgetfulness, and its habits.⁹ "One does not only wish to be understood when one writes; one wishes just as surely *not* to be understood," for example to be understood only by one's disciples (381). By writing for a particular audience, writers "at the same time keep away, create a distance, forbid 'entrance,' understanding" (381). They use their style to create distance, and distance yields power. When seen at a distance, a thing appears whole (78), serene, and quiet

(60), and this is a source of its power. Yet distance is suspicious, for it smoothes imperfections and dims noise. If you meet with a politician or a friend and he is distant, it is because he is *hiding* something from you. People use morality and lofty ideals and other ways of distancing to hide their boredom, sickness, tameness, and ugliness (352, 359, and 377). But the same applies to people who are too friendly: what are they hiding? Having reasons for one's actions ready on the lips betrays that there are secret reasons lying underneath (194). Noise prevents reflection (280); hastiness robs it of its dignity (6) and makes a person's character unconvincing (230).

We can read Plato quite effectively by looking for what Socrates does not explain, or carefully leaves out; for what he chooses to refute, or does not refute. Every word, every phrase at once reveals something and conceals something else. We choose certain words and intonations over others because we want to draw attention to certain things and ward it off of others. Total honesty is impossible; lies are as common as birds. In Section 190, Nietzsche says that even to praise is to say, "I am your peer," though praise seems innocent. For example, gifts are wrapped in many layers of meaning. Do you, the giver, place the receiver under obligation? Do you imply that you are richer, more generous, or more creative? What if the gift is inappropriate? Too appropriate? Do you give it under obligation? Or in order to ward off contempt?

Consciousness and Scepticism

These criticisms come together in Nietzsche's writing on consciousness. Here we will see Nietzsche making some of the same intellectual moves as we have mentioned so far. This time he uses them to breach the very stronghold of modernity — the subject — from within. Since Descartes, the subject, and, more precisely, consciousness and its self-reflexivity, has not only been the ground of understanding, but that of certainty as well. Nietzsche attacks the modern prejudice that, since consciousness is fixed and universal, if one sees a complete thing in his mind's eye, he has certain knowledge of it, or, as with Kant, of its relation to consciousness. This prejudice arose because the subject was isolated from the world and its perception of objects was immediate, whole, clear, and pure. Nietzsche says:

The great certainty of the natural sciences in comparison with psychology and the critique of the elements of consciousness — one might almost say, with the *unnatural* sciences — is due precisely to the fact that they choose for their object what is *strange*, while it is almost contradictory and absurd to even *try* to choose for an object what is not-strange. (355)

The *cogito*, the thinking subject, was divided into two moments: first, the perception

of the object before the mind's eye, and second, the operation of the mind reflecting on the object. Descartes was sceptical of the first moment. For him, the objects before the mind were suspicious, but the operations of the mind were transparent. They belonged to me. I could choose not to believe in anything that appeared before my mind's eye, but I could not doubt that it was my mind that chose.¹⁰

Nietzsche's scepticism, however, undermines both moments of thought. We have already seen how we cannot find objects of knowledge that are independent of the subject. Nevertheless, to distinguish an object, it must be separate from the subject. In other words, I might be able to identify my enemy well, but only if I can also identify my friends. If I can identify no friends, I suffer from a sceptical nihilism: *suspicion*. That is, if the operations of my mind are less mine than they are social activity, or if they are not in my control, are unconscious, or are deceptive, "I" cannot think, and there is *no cogito*. Who is the "I" that thinks? If the knower cannot be isolated and is infected with drives, it seems there can be no knowledge. But as we shall see, Nietzsche can give up Descartes's isolated "I" and still have knowledge, on the condition that we can see mental content and its inter-related signs as strange and distant.

Consciousness, History, and Instinct

Knowledge does not seem to be necessary for life, however. Nietzsche suggests in Section 354 that we could live quite well if nothing ever rose into consciousness — if there were no consciousness. He makes a stronger point in Section 11, calling consciousness weak, undisciplined, and quick to make false judgements. It is dangerous to humanity, so our instincts give it little leeway. They tyrannise it and make it extremely valuable to us so that its delicate parts are treated gingerly and do not cause catastrophe under the intoxication of a false feeling of strength. Our instincts carefully channel its growth until it becomes as confident and strong as instinct itself. The evolution of consciousness aims at "*incorporating* knowledge and making it instinctive" (11).¹¹

What does this mean? In Section 110, Nietzsche mentions that "knowledge became a piece of life itself, and hence a continually growing power." Knowledge will be incorporated into life, into the human creature. The process of coming to know should be as spontaneous and free as a lion's conquest of its prey, with the same clarity and innocence and lack of malice. Just as, according to its taste, in its hunger a lion identifies food by its skin, movement, and smell, so too, according to its taste, in its desire consciousness comes to know its objects by their appearances, movement, and smells. The metaphor Nietzsche uses for knowledge is not one of digging, uncovering, or ascending, but of instinct, of body. Instinct, he implies, is learned behaviour. All behaviour presupposes a valuation of the world — in other words, a particular taste.

It follows that, for Nietzsche, instinct has a history. "The whole of human and

animal past continues in me to invent, love, hate, and to infer," he says in Section 54. Scholars seek a kind of knowledge that is determined by their heritage, which has become in them an unconscious drive and has "entered their blood" (348). The past lives and acts in the present. Therefore it is impossible to isolate the subject, for it no longer has possession of itself, but is controlled by, and unconscious of, its past. Of course, according to Nietzsche, the past is shadowy. "All of us harbour concealed gardens and plantings," (9) which grow plants and bugs and fungi whether we know it or not. Even a sunny past is difficult to see; an action, especially a great one, such as the murder of God, could take centuries before it has an effect, and especially before it is perceived (125). Our desires and parts of our own minds hide from us and grow unruly and strong in secret, moving the floor of consciousness. Thus, according to Nietzsche, the objects of consciousness are dishonest and its operations are playthings of the drives of the organism.

Consciousness is Social

In Section 354, Nietzsche outlines how the mind's operations open it and its content to deconstruction. Here he says that the whole development of consciousness occurred under the pressure of the need to communicate, though it did not necessarily originate in the need for communication.

What did the creature need to share? It needed to share its actions, thoughts, feelings, movements, and needs in order to protect itself and remedy its distress. Consciousness is a kind of surface or waiting room where these things are prepared, or "known," so that they can be communicated clearly. It is only here at the surface that thoughts can take the form of words. The vast majority of thoughts are inchoate and unconscious. Consciousness itself, then, "is really only a net of communication between human beings" (354).

Nietzsche dislikes the idea that one can become conscious of something as-yet unconscious *by expressing it*:

They write and write their insufferably
sagacious larifari,

As if to *primum scribere*,

Deinde philosophari. (First write, afterward philosophise) (Poem 34)

To communicate something does not mean you understand it or even know you are communicating. Expression can generate content and can also generate consciousness of that content.¹² Nietzsche says that "Man, like every living being, thinks continually without knowing it; the thinking that rises to *consciousness* is only the smallest part of all this" (354).¹³ But I argue that one can be so unaware of what one says or does as

to be effectively unconscious of it. Nietzsche's position, that one needs consciousness to communicate *clearly*, is still strong. But for Nietzsche this shading off of consciousness towards unconsciousness qualifies all knowledge. Knowledge also shades off into ignorance.

Consciousness is social — how are we to understand this? In Section 354, Nietzsche goes on to say that everything in consciousness can only be general, average, shallow — a reduction of everything to the superficial, an over-simplification. It cannot be conscious of anything singular or unique, even when it is conscious of itself. Thus, as he says in Section 354, everything in consciousness is already mediated; that is, related to other objects, words, or thoughts that are relativised, categorised, deflated, and corrupted, so that they can be communicated to some other person. It is impossible to hold something as unique in the mind, for as it becomes thought, it is entangled in a web of concepts and communicable signs. Signs are not restricted to words, but may include gestures, gifts, and art. This aspect of animal consciousness¹⁴ makes a surface-and-sign world out of the whole of conscious thought (354). I argue that each conscious thought admits of a critique of the whole milieu of communication in which it is shared and the history of that milieu. In other words, saying that authors "wish to be understood when [they] write...[and they] wish...just as surely *not* to be understood" (381) is a critique of all thought. Our conscious mind hides its content from us, for its content is already related to signs. All the falsification involved in communication occurs as the thought rises into consciousness. Our thoughts are revealed and concealed according to our taste and are ready to be communicated. For before saying anything, in becoming conscious we have already entered the milieu of conversation and begun to communicate.

Singularity

At this point, we must ask whether the milieu of communication can appear whole, or if it is boundless. This is the condition of possibility of self-reflection, and therefore of knowledge. Another question we must ask is, can it be unique? Nietzsche says yes when he says that we must order ourselves to correspond with a single taste, to a single style (290). This does not require total self-control, since it also requires that we be able to lose ourselves (305) to the drives that create our taste. Nietzsche's hope is that we become supremely aware, supremely alive. He likens himself to waves that move with overflowing vitality, with *Übermut*, or over-mood:

It seems to me that most people simply do not believe in elevated moods....But to be a human being with one elevated feeling — to be a single great mood incarnate — has hitherto been a mere dream....Nevertheless history might one day give birth to such peo-

ple...[for whom this] might perhaps be the usual state...a perpetual movement between high and low, the feeling of high and low. (288)

Shortly after saying this, using the word "gaiety" of *la gaya scienza* (the gay science) to describe the enjoyment of the discipline, Nietzsche says one must order one's character according to a single taste (290). This does not require tyrannical self-control, since it also requires that we be able to lose ourselves (305) to the drives that create our taste.

If the process of coming to knowledge is to be incorporated in the human creature, like a lion coming to satisfy its hunger, then, like hunger, consciousness must be able to be singular. Nietzsche yearns

to experience the history of humanity as a whole as *his own history*...if one could finally contain all this in one soul and crowd it into a single feeling — this would surely have to result in a happiness that humanity has not known so far: the happiness of a god....This godlike feeling would then be called — humaneness. (337)

But as we have seen, his critique of knowledge as communication seems to imply that uniqueness and consciousness are antithetical. To save the gay science, we can distinguish two kinds of knowledge: (1) knowledge developed for communication and familiarity, unable to know anything unique, as we discussed above, and (2) a kind of knowing that is singular and incommunicable and has a very different relationship with memory. As we have seen, Nietzsche observed that consciousness developed under the pressure of the need to communicate, but did not originate in that need (354). Nor can thought be reproduced entirely in words, which suggests that something about thought is incommunicable (244).

But Nietzsche's critique of consciousness left us with the problem of how we can be conscious at all. How can we relate parts of our experience to one another, or to signs, if we are lost in a tumult of drives and thoughts? We must be able to separate the knower from his object, to take distance from it, in order to know it. If we cannot, our situation is somewhat like a swimmer caught in a rip tide — without being able to see land, we can neither take our bearings nor detect that we are being swept out to sea. To see something from a distance is usually to see it beside something else, allowing us to discern our relation to it and to compare the two things with one another. Seeing one thing beside another already relates them and turns them both into signs. Something like this process must occur in order for a thing to enter consciousness: a pre-conscious reflexivity on all our mind's content. If the origin of consciousness corresponds to a kind of knowledge that is distinct from the first, what is it? In Section 335, Nietzsche says that actions are unknowable, that is, *impenetrable*. It follows that

the need to communicate makes us *penetrate* the object as we come to know it. But there is a kind of consciousness that allows us to remain outside the object and swoop down upon it, as Nietzsche does. It seems to follow that actions are impenetrable because they are unique — they can neither be known nor communicated. But we have an awareness of actions, though they stand outside the web of communicable thoughts. This is crucial if a person is to love his whole fate.¹⁵

But Nietzsche's critique of consciousness left us with the problem of how we can be conscious at all. How can we relate parts of our experience to one another, or to signs, if we are lost in a tumult of drives and thoughts? We must be able to separate the knower from his object, to take distance from it, in order to know it. If we cannot, our situation is somewhat like a swimmer caught in a rip tide — without being able to see land, we can neither take our bearings, nor detect that we are being swept out into the sea. To see something from a distance is usually to see it beside something else, allowing us to discern our relation to it, and to compare them with one another. Seeing one thing beside another already relates them and turns them both into signs. Something like this process must occur in order for a thing to enter consciousness: a pre-conscious reflexivity on our mind's content. As Nietzsche says, in becoming conscious of things, we already relate them to one another. This also means that we constitute them pre-consciously.

This pre-conscious moment could be the spring of the hidden "Yes" that Nietzsche describes at the end of Section 377, saying, "The hidden Yes in you is stronger than all Nos and Maybes that afflict you and your age like a disease"; the Yes is a faith. But if so, the Yes he seeks is also conscious, or overflows into consciousness. He seems to suggest this in his description of those humans who will prepare "the age that will carry heroism into the search for knowledge and that will *wage wars* for the sake of ideas and their consequences" (283). There he describes attributes of the courageous and moderate preparatory "human beings who know how to be silent, lonely, resolute, and content and constant in invisible activities,...distinguished...by cheerfulness, patience, unpretentiousness, and contempt for all great vanities" (283). Their discipline is conscious, but a consciousness aware of a deep and sustaining affirmation of life, the hidden "Yes."

In Book 4, which was originally the last book of *The Gay Science*, the second last aphorism is titled "The Greatest Weight." This is where Nietzsche first introduces the idea of eternal recurrence:

What, if some day or night a demon were to steal after you into your loneliest loneliness and say to you: "This life as you now live it and have lived it, you will have to live once more and innumerable times more; and there will be nothing new in it..." Would you not throw yourself down and gnash your teeth and curse the

demon who spoke thus? Or have you once experienced a tremendous moment when you could have answered him: "You are a god and never have I heard anything more divine."...The question in each and every thing, "Do you desire this once more and innumerable times more?" would lie upon your actions as the greatest weight. (341)

When you feel most isolated, most singular, the demon confronts you with the impenetrable fact of eternal recurrence, making you both extremely self-conscious and conscious of the demand of each and every thing. I argue that in the tremendous, overflowing moment at which you could adore the fact of eternal recurrence, you consciously take the hidden "Yes" upon yourself.

But this primordial constitution, this hidden "Yes," while a prerequisite to knowledge, is not knowledge itself, and nor is it conscious. How can this unconscious "Yes" rise into consciousness? In what follows, we shall try to show how one becomes conscious and how this "Yes" can become the conscious *amor fati* (love of fate). The problem of knowledge implies the need for clear reflection — clear consciousness — and in *The Gay Science*, distance and self-reflection are associated with *writing one's life*, that is, with artistry and confession. The life of knowledge is for Nietzsche the life of aesthetic, reflective distance.

Aesthetic Distance, and Writing One's Life

The question of whether for Nietzsche knowledge is possible or not comes to this: does he believe that we can distance ourselves from things? The problem is that, "What is familiar is what we are used to; and what we are used to is most difficult to 'know' — that is, to see as a problem; that is, to see as strange, as distant, as 'outside us'" (355).

On his terms, however, the problem of knowledge is subsumed in the problem of life. An important metaphor in Nietzsche is language as life. In Poem 52, titled "Writing with One's Feet," he says

Not with my hand alone I write:
My foot wants to participate.
Firm and free and bold, my feet
Run across the field — and sheet.
(Poem 52)

In Section 299, he says "we want to be the poets of our life," and in Section 301 he echoes this, saying that we ourselves are the poets who continually create our lives. In Section 376, a man, who was the author of his life, looks back over it with an

autumnal satisfaction.

That Nietzsche *wants* distance to be possible is plain in Poem 25:

...who *I* am, I cannot tell:
My eye is much too close to me,
I am not what I saw or see...
If only I could sometimes sit
Farther away...
(Poem 25)

This desire is also plain in Section 107, at the end of Book 1, when he says "at times we need a rest from ourselves by looking upon, by looking *down* upon, ourselves and, from an artistic distance, laughing *over* ourselves or weeping *over* ourselves." But how do we understand this artistic distance? We find his answer in Section 93, appropriately, in a conversation. This should remind us of his writing on consciousness. He gives the players no names, instead calling them A and B, which heightens the weirdness of the dialogue. It is likely that it represents an inner struggle. In response to B's questions, A says to B that he writes to get rid of his thoughts and he knows no other way of doing so. Though it is a pressing and embarrassing need, he says, he has no explanation *why*, nor does he know *if* he wants to, but he *must* get rid of his thoughts. The section suddenly ends in rupture, with B (not A) saying, "Enough! Enough!" The tension and abruptness of the section signals that Nietzsche was aware of the importance of this murderous, self-reflective distance. But — to get *rid* of his thoughts! That implies that writing puts a distance, indeed, an unbridgeable gap between the writer and the thoughts expressed in his writing.¹⁶ By writing, making art, or naming the world, we can take an aesthetic distance from ourselves and see ourselves on the surface of our artwork. Nietzsche says that artists spend their time devising ways of

moving away from things until there is a good deal that one no longer sees and there is much that our eye has to add if we are still to see them at all...or giving them a surface and skin that is not fully transparent — all this we should learn from artists. (299)

This applies to poets as well. Nietzsche notes that a poet has "*vis contemplativa* and the ability to look back upon his work" (301), just as an artist does. But it is not only one's thoughts that grow distant when one expresses them; one loses oneself as well. "As a human being 'communicates himself' he gets rid of himself, and when one 'has confessed' one forgets" (351). I argue that if one communicates with something physically outside, the relations of guilt and conscience in a confession, for example, are nonetheless internal and reflect upon the whole man: a guilty man cannot escape him-

self, for he cannot escape his own memory. But, with the help of a confessor, he takes an internal distance from himself and his error. In other words, the confessor mediates his error and helps him understand it as though from outside. He gives the guilty man perspective on both his error and his guilt, distance from which he can reflect upon them and upon himself, rather than being lost in them. The distance Nietzsche describes is opened by external expression, but the result is self-reflection.

However, just as poets (and playwrights) are liars (80, 91, and 299), so is self-reflection, for it conceals much, and the observer either fills in or ignores the gaps. But self-reflection presents a man with a surface, a skin, a unity. In the motion of self-reflection, a mind moves away from itself. From there, it sees the content of consciousness as a unity and at the same time falsifies it. Nietzsche holds that we can take this aesthetic, reflective distance through expression. So reflection means taking distance from, which makes a thing foreign and strange, but also makes it possible to see a kind of unity in it. Pre-conscious thoughts are already mediated and formed by our drives — they are our drives. A person looking at a tree from a distance would not see branches and leaves everywhere, as he would were he nearby, but he would see a self-enclosed, identifiable shape that is related to other shapes. The tree itself would appear self-enclosed because of the man's perspective. In this way, his perspective constitutes the unity of the tree. Nevertheless, it would appear unified, and so could be knowledge. So, too, if we see ourselves as though from a distance, then our drives, like branches, would appear as a surface, a unity, and they would become thoughts. If a man looks at himself from a distance, he could see his own relation with things.

Consciousness

Seeing or knowing a thing requires distance, and the distance makes it appear somehow unified. But distance implies something hidden, a lie. Just as we can only see surfaces, we can only know surfaces. Complete understanding of a thing is impossible, for we cannot see every surface because a surface is a nothing, an edge by which a thing appears to what is outside it. We take distance from our consciousness, thereby falsifying it, but in the same instant see it as though a whole and know its content. Since for Nietzsche the milieu of consciousness itself is like a surface on which things can be understood, then if our drives change our tastes, it is simply the surface which alters, and we see different things. That is, if our drives change, how and what we know changes with them — but not *that* we know.

In Section 367, Nietzsche is careful to distinguish two ways a person can look at a work of art. One is to look at it and forget the world; the other is to be aware of oneself as a witness of it. The witness is, of course, better, for he is aware that he is looking at a work of art, standing just here, at a particular hour, in a certain mood. He is not lost in it and is probably aware that it has an origin (if not what the origin was),

and therefore it seems strange and arbitrary, like its origin. Similarly, in self-reflection, a man should not simply gaze at himself but be aware that he is watching himself. Knowledge requires something more than an outside, distant view of an object. Otherwise, simply watching a person or animal would provide knowledge. For knowledge, you must be aware of what is hidden as well. How can a man be aware of what is concealed by distance, if he is at a distance?

To do this, I argue, I need to be both far from something *and* in something. The first makes it appear whole and relates it to other things, and the second shows its complexity and disunity. How is it possible to have both at once? I must be able to remember the complexity of something while looking at it from afar. This improves the detail of my thoughts. But, as Nietzsche says, I must also be aware of myself watching. While I reflect on my life, as on an object, if I remember that "this is myself" who I observe, then suddenly the strangeness of Nietzsche's second kind of knowledge comes upon me. Therefore, memory also makes me aware that from a distance, I see a surface. Relating-consciousness *penetrated* that surface and, dissolving it in relations, constituted it and became it, whereas singular consciousness did or could not penetrate the surface and remained above it. The desire for familiarity that drives the former is the urge to appropriate the object to oneself, while the desire for distance, for singularity or uniqueness, that drives the latter is the desire to appropriate the whole of oneself to the world in one's very singularity.

Memory is the substance of this mental activity, underlying it and giving it content. Memory is the space that opens between the self that sees and the self that is seen, or between the subject and the object. It is significant that memory loss seems to lead to loss of distance from the world as well, as, for example, in Alzheimer's disease or amnesia. Though memory itself does not make a thing strange, it is the condition of possibility of knowledge. Consciousness is the interplay of memory and self-distancing, driven by the desire for familiarity and the desire for singularity.¹⁷ Insofar as knowledge must be a relation of a subject with an object, only the desire for uniqueness can give rise to it.

In Nietzsche we can distinguish two levels of consciousness: one pre-conscious and one conscious. The latter is the realm of knowledge. It seems that someone could have a thought and yet not know or be conscious of it — so the thought would not be knowledge, *per se* — until he reflects on it. Before being fully conscious of his thoughts, he can articulate them. He becomes conscious of them afterward by seeing them as strange.¹⁸ He can mediate his thoughts to himself through the expression of them, or he can mediate them internally. In this mediated self-reflection, they can appear strange, problematic, and distant, "outside." As we saw, this is the condition of possibility of consciousness for Nietzsche (355), and is, as it were, the opposite of the relating process by which I become conscious of a thing. I do not penetrate, but rather *unravel* a thing's relations and see it as unique and thus strange. The desire for famil-

arity is the desire to dissolve or forget the object of knowledge. But the desire for distance is a constituent of the desire for knowledge as such. This distance gives us the power to name it.

But Nietzsche wants this process to occur without external mediation. In the dialogue between A and B, it is plain that A writes to communicate thoughts that he has already had (93). When the demon comes and announces the doctrine of eternal recurrence, you cannot live and *then* answer (341). Your "Yes" to life must be given *while you still live*. Nietzsche wants communication to be chosen by a self-conscious, self-critical being. Nietzsche's ideal is like the subjective thinker in Kierkegaard, who is aware of the object of thought, but also of his relation to the object. In Nietzsche, the greatest men are aware of their distance from — and so their freedom from — the entanglements of the world. Only by self-distancing, which makes self-reflection possible, can they see themselves beside and distinct from things in the world. Seeing both self and object as singular and strange is the condition of knowledge, and if there were no primordial affirmation of singularity, one could fall into familiarity, into nihilism. For Nietzsche, it is a way of looking at something and at the same time being aware of looking at it, being distant from it. The man is conscious of himself, but as an other, as an *enemy*.¹⁹ In Section 11, Nietzsche announced the new horizon of knowledge: its *incorporation* into life. But what must become instinctive, strong, natural, and complete is *consciousness*, for "consciousness is the last and latest development of the organic and hence also what is most unfinished and unstrong" (11). Consciousness must develop slowly in order to be strong. I argue that the higher man is he who is aware of his pre-conscious act of constituting the world. When the mind becomes truly, wholly *conscious*, then that pre-conscious act, that Yes, becomes the conscious Yes of *amor fati* (love of fate), the total acceptance of life in the face of the demon (276 and 341).

Conclusion

"What could I say about any essence, except to name the attributes of its appearance?" Nietzsche asks in Section 54. It is important that this epistemology of surfaces does not presuppose any access to what might lie underneath. If there were a "thing in itself," for Nietzsche it is completely out of our reach, for the idea of an "underneath" cannot correspond to a reality independent of the human. Even if you were to open up someone's body, you could never find an "underneath," but only more surfaces — for example, the surface of the heart, under the surface of which is simply a different view of the heart.²⁰ Nor is it necessary that our trust in these surfaces, as it were, collapse under the weight of suspicion and make "knowing" impossible and reality absurd. As we said in the section on naming, these surfaces *cannot* collapse, for we cannot but live in a world of meaning and interpretation (374, 58). We cannot experience the world outside of interpretation. Whether all life interprets or it's possible to not interpret, we

cannot know (374).²¹

I tried to show how Nietzsche associates knowledge with naming and aesthetic distance. These issues join in the problem of consciousness, which Nietzsche understands by its trajectory, by the supremely conscious humans to come. He destroys the traditional conception of reality in three blows: first, by showing that our experience of reality is of the meaning we put into it, its surfaces, and its signs; second, that as in a conversation, whenever something is revealed, named, or seen, something else is simultaneously concealed; finally, by drawing attention to the fact that we are saturated with drives, history, and the mediations of consciousness. For both the objects of our thoughts and how we think about these objects are determined by these distortions. Nietzsche does not distinguish knowing, feeling, and willing from one another. They are bound up in the idea of Section 354 — that consciousness itself is social. Here we find that all knowledge admits of a critique of the milieu of communication in which it arose and will be expressed.

I tried to argue that in Nietzsche's epistemology, knowledge is possible through distance and reflection, and that this is consciousness proper. But self-reflection is concerned with what is impenetrable, and it escapes expression, for the consciousness that grew to communicate dissolves its object and ends knowledge by making it familiar. Expression must be of something whose uniqueness is seen but not penetrated. I tried to argue that for Nietzsche the reason for this is that consciousness must grow strong in order to have full possession of itself. Until that moment, one experiences only what is already familiar: oneself. Only then can the mind face itself as whole, unified, and singular. Only then — not from nothing, nor in a wasteland of meaning, but with the tremendous affirmation in flood from one's origin — only *then* can one truly answer "Yes" to the demon and proclaim one's *amor fati* (276).

Notes

¹ The nihilism Nietzsche identified is that the will to truth, e.g. modern science, has destroyed its metaphysical ground — the belief in truth (§344) — in its pursuit of certain, objective knowledge. However, to be of any value, the will to truth requires the belief in truth, or it is nihilism. In rejecting God, science thereby rejects the belief in truth, destroying its own principle. See §125, 343, 347. All subsequent in-text references are to sections of Friedrich Nietzsche, *The Gay Science*, trans. Walter Kaufmann (New York: Random House, 1974).

² *Ibid.*, §324.

³ William Shakespeare, *Romeo and Juliet* (Don Mills: Academic Press, 1964), I.v.93.

⁴ Nietzsche's aphorism in §224 of *The Gay Science* — "animals as critics" — does exactly this, but in a slightly different context. He explicitly examines knowledge from the perspective of animals in his essay 'On Truth and Lies in a Nonmoral Sense,' in Daniel Breazeale (ed. and trans.), *Philosophy and Truth: Selections from Nietzsche's Notebooks of the Early 1870s* (New Jersey: Humanities Press, 1979), 79.

⁵ In a general way, on the importance of names and taste, Nietzsche resembles Confucius, who differs mainly in his emphasis on the rightness of each. See A.C. Graham, *Disputers of the Tao: Philosophical Argument in Ancient China* (Chicago: Open Court, 1989), 23-29.

⁶ Nietzsche, *The Gay Science*, §135 presents some of the ideas, and uses the same techniques, that are in his *A Genealogy of Morals* and in Part 5 of *Beyond Good and Evil*, both of which can be found in *Basic Writings of Nietzsche*, Walter Kaufmann (trans.) (New York: Modern Library, 1992). Cf. also *The Gay Science*, §149.

⁷ To Nietzsche, those who are not great are not individuals but herd animals without their own point of view.

⁸ Nietzsche did not use the words "reinvented" or "re-created," possibly because there is a world given to the individual that he does not create himself, though he constitutes everything he receives from it.

⁹ This technique is an attack on the philosophy of G.W.F. Hegel, who, it seems to me, argues that consciousness is communication and, furthermore, that on the level of Spirit, all is, as it were, understood.

¹⁰ René Descartes, *Meditations on First Philosophy*, in *Meditations and Other Metaphysical Writings*, trans. Desmond Clarke (London: Penguin, 1998), 23-4. The operations of the mind are doubting, understanding, affirming, denying, willing, not willing, imagining, and sensing (26).

¹¹ In §110 of *The Gay Science*, Nietzsche poses the question, "To what extent can truth endure incorporation? That is the question; that is the experiment." The character of this experiment is beyond the scope of this essay, but it is the attempt to reconcile the impulse for truth, once it becomes life-preserving, and error, which is

already life-preserving.

¹² This is central to the philosophy of Hegel, for whom what is expressed is oneself.

¹³ See also *Gay Science*, §333.

¹⁴ Whether animal consciousness is the same as human consciousness is beyond the scope of this paper. I suspect that memory in animal consciousness differs from its human counterpart.

¹⁵ Hannah Arendt was a student of Nietzsche in this respect. Her insights into surfaces and uniqueness are profound and remarkably similar to Nietzsche's.

¹⁶ See also Friedrich Nietzsche, *Beyond Good and Evil*, in *Basic Writings of Nietzsche*, trans. Walter Kaufmann (New York: Modern Library, 1992). In Section 296, the last section of the book, Nietzsche mourns the thoughts he has just set into words.

¹⁷ Plato's *Meno*, particularly the theory of recollection, is relevant here. Plato, *Meno*, trans. G.M.A. Grube, in John M. Cooper (ed.), *Plato: Complete Works* (Cambridge: Hackett, 1997). It would be interesting to examine Nietzsche's writing on memory. It would also be interesting to study forgetfulness in the existentialist and absurdist playwrights. For example, loss of memory is central to Samuel Beckett's *Waiting for Godot*. It is because of their loss of memory that his characters can only accept a truth in pure faith and carry no consistent perspective that could allow them to name or set their own meaning — they have lost *themselves*, for memory is the substance of consciousness. Samuel Beckett, *Waiting for Godot* (New York: Grove, 1954). The role of forgetting in self-reflection is beyond the scope of this paper.

¹⁸ A certain kind of forgetting is crucial when one looks at oneself, if it is to be a self-reflective glance. What if consciousness was a function of memory?

¹⁹ This echoes Hegel's phrase, "Destiny is consciousness of oneself, but as an enemy," which I found quoted in Jean Hyppolite, *Introduction to Hegel's Philosophy of History*, trans. Bond Harris and Jacqueline Bouchard Spurlock (Gainesville: University Press of Florida, 1996). Hyppolite did not give a reference.

²⁰ See Nietzsche, *The Gay Science*, §256 and §59 for complementary statements about skin and surface.

²¹ If some existentialists hold that reality is nonsensical, or life meaningless, Nietzsche points out that they could never experience them that way. Perhaps they suffered *fear* of an absurd reality.

Sources

Beckett, Samuel. *Waiting for Godot*. New York: Grove, 1954.

Descartes, René. *Meditations on First Philosophy*. In *Meditations and Other Metaphysical Writings*. Trans. Desmond Clarke. London: Penguin Books, 1998.

Graham, A.C. *Disputers of the Tao: Philosophical Argument in Ancient China*. Chicago: Open Court, 1989.

Hyppolite, Jean. *Introduction to Hegel's Philosophy of History*. Trans. Bond Harris and Jacqueline Bouchard Spurlock. Gainesville: University Press of Florida, 1996.

Nietzsche, Friedrich. *A Genealogy of Morals*. In *Basic Writings of Nietzsche*. Ed. and trans. Walter Kaufmann. New York: The Modern Library, 1992.

— *Beyond Good and Evil*. In *Basic Writings of Nietzsche*. Ed. and trans. Walter Kaufmann. New York: The Modern Library, 1992.

— *On Truth and Lies in a Nonmoral Sense*. In *Philosophy and Truth: Selections from Nietzsche's Notebooks of the Early 1870s*. Ed. and trans. Daniel Breazeale. New Jersey: Humanities Press, 1979.

— *The Gay Science*. Trans. Walter Kaufmann. New York: Random House, 1974.

Plato. *Meno*. Trans. G.M.A. Grube. In *Plato: Complete Works*. Ed. John M. Cooper. Cambridge: Hackett, 1997.

Shakespeare, William. *Romeo and Juliet*. Don Mills: Academic Press, 1964.

Decision-theoretic Consequentialism and the Nearest and Dearest Objection: A Critique

Jared Will, McGill University

In a paper titled 'Decision-theoretic Consequentialism and the Nearest and Dearest Objection,' Frank Jackson tries to counter the objection that consequentialist ethics is dehumanising. He does so by embedding concepts from Decision Theory, a type of Rational Choice Theory, into the consequentialist theory of right action. In order to evaluate Jackson's theory, I will present his argument in the same manner that he does, first describing the theory and arguing that it is not *prima facie* implausible, then showing how it meets the so called Nearest and Dearest Objection — the objection that Consequentialism is dehumanising. I will then argue that Jackson not only inherits the flaws inherent in Decision Theory, but that even if we grant the validity of Decision Theory, his theory in no way refutes the Nearest and Dearest Objection. That is, not only is his theory unsound in virtue of being a version of Decision Theory, it is also untenable as an ethical theory.

Consequentialism, broadly construed, is the moral theory that evaluates right action in terms of the consequences of an action. Thus, an act is morally right if and only if, of all available actions, it brings about the greatest amount of good in the world. An action that meets these conditions is said to *maximise utility*. Further, the results of an action must be evaluated impartially; that is, the value of the outcome cannot in any way be dependent on *who* benefits from the act. As a result, not only are the particular people affected by an action irrelevant in considering its utility, but the status of the particular agent is also irrelevant. It does not matter who performs the act, so long as it is the act that maximises utility, wherein utility is determined impartially. The determination of the utility of a consequence, which is impartial by definition, is referred to as the consequentialist value function.

The Nearest and Dearest Objection is one of the most well-known objections to Consequentialism. It can be summarised as follows: because Consequentialism requires that we assign *impartial* values to the outcomes of actions, we are not permit-

ted to pay any *special* attention to those people whom we care about — our friends and family, our nearest and dearest. If we place more importance on their well-being than on that of others, we would violate the impartiality condition of Consequentialism. Moreover, because caring for and paying special attention to our nearest and dearest to a large extent constitutes what makes human life valuable and worth living, Consequentialism dehumanises and devalues life. Because these types of action necessarily constitute *the good life*, Consequentialism deprives us of the good life because it does not permit us to act in this manner. Thus, Consequentialism seems to be in conflict with what makes life worth living.

Jackson's project is to argue that, properly construed, Consequentialism is not guilty of this charge. He begins by reiterating the point that the consequentialist tells us to do whatever maximises impartial value in the world. He wants to investigate how the values that we assign to the outcomes of our actions affect the determination of what ought to be done, while maintaining the goal of maximising utility.¹ It is into this interaction between assigned values and determinations of right action that Jackson believes Decision Theory can be embedded.

Jackson begins his analysis by introducing a hypothetical situation that he thinks will demonstrate the intuition behind his theory. A patient comes to a doctor with an irritating, but non-lethal, skin infection. The doctor has three drugs available with which to treat this condition. Based on her knowledge of medicine and the condition of the patient, she is able to make the following assessment: drug A is very likely to relieve the condition without any side effects, but will not cure it. Either drug B or drug C will completely cure the condition, while the other will kill the patient, but the doctor has no way of determining which drug will have which effect. Jackson proposes that the first thing the doctor should do is rank the value of the possible outcomes of her action. She does so in the obvious way, highest to lowest: complete cure, partial cure, and death. Now, the problem becomes how one should move from this assessment of value to a decision about what to do. It is here that Jackson proposes that the ethicist should take a lesson from the decision theorist.

Subjective probability is the probability that a particular result will occur, *as calculated from an agent's limited epistemic stance*. According to Jackson, when making moral decisions an agent should assess the values of every possible outcome, multiply the value of each impartially evaluated outcome² by the subjective probability of each result and sum them for each possible action. That is, add up the results, as specified, for every possible outcome for each possible action, and choose the action that receives the highest value. In the case of the example provided:

$$\begin{aligned} & \text{Pr}(\text{partial cure}/\text{drug A}) \times V(\text{partial cure}) + \text{Pr}(\text{no change}/\text{drug A}) \times V(\text{no change}) \\ & \text{Pr}(\text{cure}/\text{drug B}) \times V(\text{cure}) + \text{Pr}(\text{death}/\text{drug B}) \times V(\text{death}) \\ & \text{Pr}(\text{cure}/\text{drug C}) \times V(\text{cure}) + \text{Pr}(\text{death}/\text{drug C}) \times V(\text{death}) \end{aligned}$$

The agent, on Jackson's view, should choose whichever of the above formulas yields the highest value, thereby recovering what the agent ought to do from a ranking of the value of the possible outcomes *and* the agent's subjective probability function. Thus Jackson formalises his decision procedure with the following rule of action: an agent ought always to maximise: $\sum_i Pr(O_i/A_j) \times V(O_i)$, where Pr is the subjective probability function, V is the consequentialist value function, O_i are the possible outcomes of each action, and A_j are the actions available to the agent. The result, which we knew long before this formula was even conceived, is that the doctor should choose to prescribe drug A. Thus, the basic idea is that the extremely low value ascribed to death, together with a 50 per cent chance of it occurring when either drug B or C is taken, sufficiently lowers the values of the latter two formulas such that the first, prescribing drug A, receives the highest value.³

From this, we see that Jackson believes that Consequentialism should command us to maximise *expected* utility. For him, the notion of expected utility captures the link between the consequentialist value function (utility) and what an agent ought to do (expected). (If the reader is wondering how these functions and values are defined, we are in full unison. Jackson leaves them undefined, which is a crucial problem for his theory. More on this below; my current objective is simply to explicate Jackson's theory as he presents it.)

Having presented the skeleton of his theory, Jackson spends a considerable portion of his paper fleshing it out and arguing for its plausibility. He stipulates that both the subjective probability and the value function are to be understood in terms of what the agent *in fact* believes at the time of the action, rather than what the agent *ought* to know and believe at that time. His reason for doing so is clear. Jackson is trying to put forward a theory of right action that an agent can follow when making moral decisions.⁴ (This is a reflection of Jackson's meta-ethical view that any sound ethical theory needs to be practicable.)

Jackson here argues that even though his theory is based on what an agent in fact believes, he can still account for culpable ignorance — the idea that there are cases in which an agent can be morally blamed for not being aware of certain facts when he or she acts. Jackson argues that postponing action and obtaining more information is almost always an available option, one that has its own expected utility.⁵ Thus, if abstaining from action pending further information is the action that has the highest expected utility and one acts regardless of this fact, one can be said to be culpably ignorant in Jackson's view.

Jackson raises and dismisses another view regarding the nature of an agent's beliefs at the time of action. This view is that of the typical staunch consequentialist, wherein an agent's beliefs are irrelevant and all that matters is that he or she performed the action that in fact had the best consequences.⁶ In this view, the explicit command is to maximise utility, as opposed to Jackson's proposal of maximising *expected* utility.

Jackson raises two serious problems with the staunch consequentialist position. The first is that such an account gives us the intuitively wrong answer in cases such as his drug example, wherein prescribing drug A will clearly *not* have the best possible results (because it will not cause a complete cure), though we would be appalled if the doctor had prescribed drug B or drug C. The second problem with an account that makes an agent's beliefs irrelevant is that, for Jackson, an ethical theory, a theory about right action, must include prescriptions for ethical *decisions* with respect to actions. The fact that a particular action does in the end have the best consequences is not very helpful unless the agent can *know* that that action will have the best consequences.⁷ So, again, because Jackson believes that an ethical theory has to be practicable, an agent's beliefs are most relevant. He gives us another example here, which he takes to further support his claim that what is important for a proper view of Consequentialism is the *expected* utility of the consequences of one's actions.

Again, we have a doctor and a patient who complains of a minor and non-lethal skin irritation. This time, the doctor only has two drugs available, X and Y. She knows that if she prescribes X, there is a 90 per cent chance of a cure but a 10 per cent chance of death, while if she prescribes Y there is only a 50 per cent chance of a cure but no possible side effects. Jackson takes it as obvious that she should choose Y, even though it is in fact far *less* likely to benefit the patient and to produce the best possible results. He believes that by implementing his notion of subjective probability along with the standard consequentialist value function, thus making an agent's beliefs relevant, he has made sense of the fact that we very often pursue the course of action that we know is in fact less likely to accrue benefit. That is, that we do not always *maximise* utility; we tend to "play it safe" more often than not.⁸ On his view, this behavior, which is intuitively laudable, is now actually prescribed by his theory.⁹ Therefore, according to Jackson, Consequentialism should not be construed such that it demands that people should always aim at the *best* consequences. Most of the time the most prudent and, on his view, moral action is the one that will definitely *not* have the best consequences.

Jackson points out that Decision-theoretic Consequentialism is not committed to any view about the requisite mental processes that occur prior to action. He believes that sometimes the best actions have to be spontaneous.¹⁰ In such cases, one would still have to have consequentialist motivations, those of promoting the greater good, without consciously reasoning out what the best action will be. For example, if I dash into a busy street to rescue a toddler, I was motivated by my desire to save the child, though I did not have time to consider what I should or should not do in that split-second. A possible objection could be raised here, namely, how could a subjective probability function factor into giving a spontaneous action a higher value than any other? The agent in such a case would not have any beliefs about the probability of the results of any action in that situation. And Decision-theoretic Consequentialism requires that an agent maximise *expected* utility. Jackson responds, somewhat laconically, that sponta-

neous action is not action without belief, but action without a conscious review of one's beliefs.¹¹

Further, Jackson points out that Decision-theoretic Consequentialism is not committed to the view that maximising expected utility is the right *motive* for action. He argues that an agent has the right motive whenever his or her beliefs and desires conform to the consequentialist value function, that is, whenever he or she desires to promote an impartially defined value, believes that action X is the best way to go about it, and therefore performs X. But conforming to the consequentialist value function does not entail conferring any extra value on the maximisation of expected utility. The fact that an action, when run through Jackson's formula, comes out on top does not give us an independent reason to desire it. We should desire to perform that action simply because it is the best way for us to bring the most good into the world under those circumstances. We should be motivated by the *high* utility that we can bring about, not the fact that this or that action *maximises* utility.¹²

It may seem that Jackson is picking nits, but his distinction becomes relevant in light of an objection that Michael Stocker raised against all traditional modern ethical theories. Stocker argued that all modern ethical theories (specifically Consequentialism and deontology) require that we are motivated by the prescriptions of the theory itself, and that if our actual empirical goal motivates us, we have in fact deviated from the theory. He calls this problem "moral schizophrenia."¹³ Jackson is arguing that his theory makes no such demand because it allows people to be motivated by the good results that they can bring about. This issue is being discussed here only because it has been widely influential in modern ethical debate, though it has little bearing on what I have to say about Jackson's theory and the independent implications of it.

Jackson raises and dismisses one more possible objection before he explicitly links his account to the Nearest and Dearest Objection. It could be objected that because the subjective probability function will vary over time, as the agent's epistemological status varies, we will end up with an equally varied sequence of 'oughts,' of what the right action would be. This objection presupposes that 'oughts' should be fixed. Jackson responds, quite rightly I believe, that the only 'ought' that matters is the one at the time of the action, and therefore all of the others become superfluous hypotheticals. This follows simply from Jackson's demand that a moral theory be normative — that is, be practicable — meaning that all that is relevant is what an agent should or should not do at the time of action.¹⁴

Finally, we get to the main point of Jackson's project. After presenting his theory and defending it from what he takes to be less serious objections, he seeks to show how it takes the wind out of the sails of the Nearest and Dearest Objection, which he takes to be a potentially devastating objection to Consequentialism. Once again, however, instead of putting his point directly, he presents it through the use of two examples. Unfortunately, I think it is necessary to work through his examples in order to under-

stand his point and to get a grip on what seems to have motivated his whole project.

Jackson's first example in this context is yet another involving a doctor's dilemma. This time, we have three patients (A, B, and C), but only one drug and only enough of it to treat one patient. Again, after reviewing everything she knows, the doctor comes up with the following scheme: Person A will accrue considerable benefit from the drug, while between B and C, one will be killed and the other will be cured, though she has no way of knowing which patient will receive which effect.¹⁵ Again, the obvious answer is that she should give the drug to Patient A. Why? Because that is the action with the highest *expected* moral value, even though she knows that there is another course of action that would yield better results; she simply does not know *which* course of action would do so, and the risks of picking the wrong action are far too high to even be considered. The important thing to notice is that she maintains an impartial assignment of value to the well-being of each of her patients. Her preference function does not assign a higher value to A's well-being; rather, it is simply the case that there is a greater subjective probability that she will be able to benefit A, in this case due to her epistemic limitations with respect to B and C.¹⁶

So, Jackson thinks that the relevant question to ask is whether the special attention we pay to a particular group (our nearest and dearest) can be justified by appealing to the fact that we have a much higher subjective probability of helping them because of our special epistemological status with regard to them. That is, can Consequentialism justify this kind of behaviour while maintaining that the values we place on people must be impartially determined? And, of course, he shows why he thinks the answer to the above question is yes by way of another example.

Jackson entreats us to imagine that we are a police commissioner assigned the task of devising a plan to control the crowd at a soccer match. We have the choice between two plans, sector and scatter. The rationale for the scatter plan is as follows: each person in the crowd is of equal value, so it would be immoral to require an agent to pay special attention to a small group and thereby neglect the rest of the crowd. Instead, each agent should wander the crowd and do good wherever he or she can.¹⁷ This is the typical consequentialist rationale pushed to its logical conclusions in a practical situation. The other option is the sector plan, in which each member of the police force is assigned to a special section of the crowd, making that section his or her personal responsibility. Jackson argues that in our daily lives we follow the sector plan, paying special attention to our friends and family, though we might feel justified in neglecting them if we find that we are in a position to make a big difference elsewhere.¹⁸

The point of this example is that it enables Jackson to argue that we will have a probabilistic justification for paying special attention to our nearest and dearest whenever a sector plan gives us the most likely way of benefiting the world and the people in it. He offers a list of five conditions under which such a plan, as opposed to a scatter plan, would be most effective: whenever (i) getting to know the particular individ-

ual that we wish to benefit is important for producing good results, (ii) achieving good results requires coordinating a series of actions in order to take advantage of the results of positive and negative feedback loops, (iii) achieving good results requires mutual trust and respect, (iv) there is a high chance of one agent nullifying the efforts of another, and (v) there is a natural way to assign agents to sectors that reflect their natural inclinations and enthusiasms.¹⁹

Jackson takes it as obvious that the sector plan is the most efficient in the circumstances of our day-to-day interaction with the people in the world (i-iv), and that we are naturally inclined to try to benefit our nearest and dearest (v). Thus, Jackson thinks that the special attention that we pay to our nearest and dearest, which we take to be an integral part of a life worth living, can be explained without having to attribute a biased value to the people to whom we pay special attention. Instead, this attention can be seen as a feature of our greater subjective probability of benefiting those people — as a feature of our epistemological relationship with the world. In other words, a considerable degree of focus on family and friends is consistent with living moral lives in consequentialist terms (that is, attempting to maximise expected utility).²⁰

Jackson raises and attempts to refute three objections that could be raised against his thesis that the special attention that we pay to our nearest and dearest can be justified probabilistically. The first, which he attributes to Bernard Williams, is that Consequentialism entails a notion of negative responsibility that makes *who* does an action somewhat irrelevant.²¹ That is, because Consequentialism commands us to maximise utility, it could often be the case that it commands us to prevent someone from acting immorally. In such a situation, if we were in a position to prevent it, but abstained, we would be on equally blameworthy grounds as the person committing the act. So in traditional consequentialist terms, the direct causal link between agent and action is not a necessary factor in assigning praise and blame. This presents a potential problem for Jackson, because if *who* does an action is irrelevant, then subjective probability cannot be a factor in determining right action. Jackson's response is that Williams fails to distinguish between value and *expected* value. *Who* performs an action has no bearing on the value of the outcome, but it certainly might have a large influence on the *expected* value of the outcome, which is a feature of the agent's perspective. For Jackson, *who* performs an act does matter because Consequentialism must be an ethical *decision* theory; that is, it must be a theory that is practicable from within an agent's subjective perspective.²²

The second objection comes to light in a familiar example in the current literature on Consequentialism, an example first presented by Peter Railton. It is an objection to Consequentialism broadly construed and one to which Jackson thinks his version is immune. In this example, a man has to decide between spending \$500 to visit his wife, who is feeling depressed, and donating that money to a charity that will dig

a well in a drought-stricken Third World village. Traditional Consequentialism would require this man to donate the money because it is the action that will provide the greatest number of people with the greatest benefit. This is intended as an objection because we are supposed to be appalled by the fact that an ethical theory requires this man to neglect his miserable wife.²³ But Jackson argues that it is not such a simple decision. He believes we must take into account how likely it is that this man can benefit his wife versus how likely it is that with an extra \$500 the charity group will be able to do more good than this man can do for his wife.²⁴ The decision involves more than the utility of the intended outcome. It crucially involves the agent's epistemic relationship to his wife as compared to the efforts of charity groups in far-off lands.

The third objection seems to me to be the most pointed. How can Jackson's Decision-theoretic Consequentialism justify the fact that our nearest and dearest is the group that it happens to be? That is, even if we grant that our epistemic relation to that group justifies the special attention we pay them, is there any justification for our choice of this specific group as the one that we will benefit? If no such justification can be found, then Jackson's theory will have the rug pulled from beneath its feet.

Jackson responds by way of an appeal to the "empirical facts of human psychology."²⁵ He first considers the possibility that choosing to benefit those we love, though it might not be the best choice in consequentialist terms, springs from the right type of character, namely, a character that generally makes one seek to produce positive outcomes. Jackson sees that this answer involves ducking the question to a fairly obvious degree, because our 'character,' however it gets defined, is not particularly relevant in consequentialist considerations. However, he attempts to make it relevant in consequentialist terms in the following manner: He argues that it is often the case that our actions will only have good results if they are followed up in the appropriate manner. Thus, in deciding which action I should choose now, if I know that in the future I am more likely to follow up on action X than on action Y, I should probably choose to perform X. Further, this decision involves making a judgment of one's character. Moreover, we generally know that we do better when it comes to these sorts of issues of character when we are engaged with projects involving those we love and respect. Therefore, because our nearest and dearest naturally comprises those we love and respect, we have a justification for our nearest and dearest being the group that it is, namely, because it *is* in fact the group that we are naturally most likely to benefit.²⁶

Jackson's position can be summarised as follows: Consequentialism assigns values in an impartial, agent-independent manner. However, the lives we consider worth living give a central place to a small group of people. In other words, people have an agent-relative outlook, one that *prima facie* clashes with the consequentialist value function. But Jackson believes that Consequentialism can explain this agent-relativity in terms of the role that subjective probability plays in the recovery of what one ought to do once the agent-independent values have been assigned. Thus, we are command-

ed to maximise *expected* utility, and because we often have a greater likelihood of benefiting a particular small group to whom we feel devoted, the special attention that we pay that group can be justified in consequentialist terms. As such, Consequentialism is not dehumanising, which is the charge of the Nearest and Dearest Objection.

My objections to Jackson's argument will fall into two categories. I will first consider the problems that accrue to his account by virtue of the fact that it is a version of Decision Theory, a branch of so-called Rational Choice Theory. I will then argue that even if we grant the decision-theoretic aspects of his theory as valid, his theory fails to refute the Nearest and Dearest Objection. Once his thesis is refuted on both of these levels, it will be clear that his argument is not only invalid, but also irrelevant. Not only does the insertion of Decision Theory into Consequentialism prove to be invalid under scrutiny, but the success or failure of that reconstruction of Consequentialism is irrelevant because even if such an account were valid, the new version of Consequentialism is by no measure immune to the Nearest and Dearest Objection.

There are three separate, though related, problematic issues with Jackson's notion of subjective probability. These are problems that Jackson inherits from Decision Theory. The first is that probability is a term that refers to the *objective* chances that an event will occur. The probability of an event is completely independent of both a subjective perspective and the probability of other events. Furthermore, probability is a feature of random results, such as the result of flipping a coin or spinning a roulette wheel, wherein there are no factors that determine which result will occur. It is strictly a matter of chance. It is probabilistic. Thus, the notion of subjective probability is self-contradictory. As soon as perspective becomes relevant, one can no longer be dealing with a probabilistic event, which must always be objectively quantifiable.

However, to be charitable, we can grant that Jackson is using the term 'subjective probability' metaphorically, so that when he says 'subjective probability' he really means 'perceived likelihood.' We could allow that he is referring to what the agent *believes* to be the likelihood of a certain action having a certain outcome (presumably calculated with reference to past frequencies and other beliefs about the nature of his or her circumstances). Not only is this the way he describes what he means by subjective probability, but it is something that we all do everyday when we are making decisions with any degree of reflection. It seems that fairness would require us to grant Jackson this charitable interpretation of his notion of subjective probability. But now we arrive at the second problem with that concept.

Though it makes sense to interpret subjective probability as perceived likelihood, the latter concept could never be quantified with the kind of accuracy that Jackson's normative formula requires. In order to implement his formula, $S_i Pr(O_i/A_i) \times V(O_i)$, when deciding which action is morally right, one would have to assign a precise value to this likelihood. However, we could never make such a quantitative assignment in any meaningful way. For example, what would it mean to say that action X has a 90

per cent likelihood of producing result Y? It could only have meaning in randomised circumstances where this sort of figure can be calculated; for example, if one drops a ball into a spinning roulette wheel that has 10 slots, nine of which are red, we can say that the ball has a 90 per cent chance of landing in red. But in this case we are dealing with a legitimately probabilistic event and not a mere perceived likelihood. The simple fact of the matter is that life is not random, and because the results of goal-directed action cannot be probabilistic, they cannot be assigned a percentage chance of occurring in any meaningful way. Thus, if we assign a 'probability' to the results of each of our actions, we must be doing so in a completely arbitrary manner. A formula such as Jackson's cannot produce anything but arbitrary results, and it is obvious that arbitrary results cannot constitute the normative principle of any ethical theory that purports to reflect the values that we hold dearly.

The third problem with Jackson's notion of subjective probability is that it is dependent on the overly simplistic assumption that an agent can know and consider every action available to him or her at the time of an action. Jackson's theory, in telling us that we should consider each action available to us, sum the expected utilities of each action's possible results, and then perform the action that maximises expected utility, implicitly assumes that we can be aware of which actions are available to us. Jackson would of course respond that the only relevant factor is which actions an agent *believes* to be available to her at the time of the action. Granted, but it is naive to assume that an agent will only perceive a relatively small number of available actions in a given circumstance. (This problem does not crop up in Jackson's examples, and this is due to a related feature that makes them fail to represent most of our moral decisions, which will be considered below.)

If an agent were to narrow down the list of available actions to those that he or she thought to be most likely to produce beneficial results and test those actions with Jackson's formula, he or she would not have any guarantee of maximising expected utility. Every action that the agent perceived to be available would have to be considered in order to guarantee maximising expected utility, and this list of available actions would usually be so long that one would spend a lifetime doing calculations before any decision could be reached. (Consider the prime minister of Canada deciding on the most ethical way to stimulate economic growth in sub-Saharan Africa.) Thus, Jackson's theory fails to meet one of his requirements for any ethical theory, namely, that it must be practicable, and, therefore, Jackson's theory cannot achieve its goal. The theory does not provide an agent with a method for deciding which action, of *all* those available, maximises expected utility.

An evaluation of Jackson's concept of a consequentialist value function presents another set of serious flaws in his argument. Jackson states that "the details of the value function will not concern us; any reasonable ranking of outcomes of the usual agent neutral kind will serve our purposes in what follows."²⁷ This is simply not the case. His

formula calls for a value and presupposes that that value will be a definite number, and not merely a relational term such as a ranking. In order to evaluate the situation in the manner he prescribes, one needs to know much more than the rank of the value of the outcomes. It would often be the case that *how much more* valuable one outcome is than the next would be the determining factor in a calculation of expected utility.

However, as in the case of the subjective probability, any assignment of a numerical value to the value of an outcome would have to be completely arbitrary, and therefore meaningless. We assign values to our goals and desiderata qualitatively (except in the economic sphere), and translating these qualitative values into quantitative terms would require an arbitrary assignment of a numerical value to that occurrence (remember that merely ranking our desiderata is not sufficient). Are we to say that curing HIV/AIDS gets a value of five, while curing cancer is only a four, and juvenile diabetes only a three? These numbers tell us nothing about the actual value we place on these goals; they tell us no more than rank, and, as specified above, ranking does not provide sufficient information.

Finally, in order for Jackson's formula to provide the results he expects, one must be able to assign a negative value to certain outcomes. Consider Jackson's second drug example: If the doctor prescribes drug X, there is a 90 per cent chance of curing the patient, but a 10 per cent chance of death; while if she prescribes Y, there is a 50 per cent chance of a cure, but no possible side effects. So, even if we assign a value of 1×10^6 to a cure, and a value of zero to death, the result is that the expected utility of prescribing X is 9×10^5 , while that of Y is 5×10^5 . Therefore, the doctor should prescribe X, which is clearly not the result that Jackson was looking for; we would be horrified if she prescribed X. Of course, we would get the desired result if we could assign (arbitrarily, of course) a negative value to death. But, strictly speaking, the notion of a negative value of an independently evaluated result is nonsensical. How can something be worth less than nothing? Or, if we create a semantic space for negative values, we reach the problem that *no matter how minimal the likelihood*, anything with a negative value would be impermissible.

The only way that Jackson's theory could account for this necessity of assigning negative figures is if the value function were a comparison of the state of the world before and after the effects of each action. That way, death and the like could be assigned negative numbers because there would be a decrease in the valuable things in the world, such as human life. Jackson, on the other hand, only speaks of values to be assigned to the outcomes in and of themselves, and not as effects on the net value of things in the world. However, the fact that Jackson overlooked this crucial detail is irrelevant in the end. Because the values that we assign to empirical phenomena cannot be assigned quantitative values in the first place, reconstructing Jackson's theory so as to allow for negative values in the manner described would be entirely superfluous. I mention this problem only to buttress my point that the introduction of Decision

Theory cannot be of any assistance to the consequentialist. It merely adds its own problems to the mix.

Irrespective of the issues concerning Decision Theory, it might be argued that Jackson's theory presents a view of Consequentialism that makes it consistent with living in the way we think to be valuable and worthwhile. A consequentialist might argue that recasting the maxim of Consequentialism in terms of expected utility rather than objective utility, without the quantitative formula of Decision Theory, makes that ethical theory immune to the Nearest and Dearest Objection. However, this is not the case. Even if we grant that Jackson's account could be formulated without incorporating the specific failures of Decision Theory, it still fails to meet his own standards, namely that Consequentialism be practicable and that it is not made untenable by the Nearest and Dearest Objection.

Jackson's version of Consequentialism, even if construed qualitatively, is not a practicable theory. This point can be seen most clearly by first considering the nature of the examples he uses to demonstrate how the theory would be applied. In each of these examples the agent in question is presented with a specific goal and the only relevant question is how best to achieve that goal by maximising expected utility in that situation, whether it's treating a patient or controlling a crowd. The fact is that human life and the moral decisions that we all face are a lot more complicated than that. Perhaps our most important moral choices are those in which we have to decide what our primary goals will be. In those cases, it is impossible to estimate the expected utility of the ends that we will be able to achieve once our goals are well defined. Our ignorance of the situations with which we will be presented once we have defined our life goals makes it impossible to apply even a non-quantitative version of Jackson's moral maxim. Our epistemological limitations, our inability to predict the distant future, make it impossible to employ his formula when making these moral choices, perhaps the most important of all choices. Therefore, Jackson's theory is not practicable; it does not help us determine what we ought to do based on our values, as he believes that all moral theories must do.

A serious problem arises when one tries to determine whether Jackson has refuted the Nearest and Dearest Objection. The problem is that Jackson is never clear about whether his theory is a descriptive or a prescriptive (normative) moral theory. He is ambiguous about whether he is either describing human action and showing how it is consistent with Consequentialism or elaborating a theory that, if its prescriptions are followed, will allow us to live morally without giving up what we think makes life worth living.²⁸ Very often in the same breath he tells us that he is dealing with a theory that can prescribe which action will be the right one and then appeals to our common-sense intuitions to show us that this is how we think anyway. For example, "it is prescribing drug A which is the intuitively correct answer for [the doctor]...we have to see Consequentialism as containing as a constitutive part prescriptions for action."²⁹

If it is the case that Consequentialism prescribes the right action in any given circumstance, then any appeal to our intuitions is irrelevant. Conversely, if our moral intuitions yield answers that can be trusted, we have no use for a moral theory. Jackson's confusion on this issue probably stems from the fact that he is trying to incorporate a theory about how we *do* act (Decision Theory) into a theory about how we *ought* to act (Consequentialism). It is simply not clear whether he thinks that what *is* the case coincides with what *ought* to be the case, or whether he thinks that what *is* the case needs to be replaced with what *ought* to be the case. Clearly, we can assume that the latter is what Jackson has in mind, because if that were not the case, he would have no reason to bother with proposing an ethical theory. But, if Jackson has it in mind to replace what is with what ought, then his appeals to common-sense intuitions have an importance that is, at best, dubious.

Luckily, the success or failure of Jackson's theory is independent of this issue. It fails if we read it as descriptive, and it fails if we read it as prescriptive, though for different reasons. If he is being descriptive, his description is simply inaccurate. The fact that it is the case that we have a higher subjective probability of benefiting our nearest and dearest than we do of benefiting people we do not know is irrelevant as a justification for our doing so if it is also the case that we *do in fact* place a higher value on the well-being of this small group, thus failing to be impartial. Moreover, not only is it the case that we do value this group more than any other, it is the fact that we do so that makes them dear to us, which, in turn, is exactly what Jackson takes to be that which makes life valuable. Therefore, not only does his justification for our action fail because we do in fact have biased values, regardless of the likelihood of our benefiting these people, but it is also the fact that we *do* have these biased values that makes our lives valuable.

On the other hand, if Jackson is presenting a normative or prescriptive theory, he actually forbids us from making our lives worth living. That is, his theory tells us that though we cannot assign biased values to our nearest and dearest, we can pay them special attention because it happens to be the case that we are more *likely* to benefit that certain group. But if we obey his maxim, then these people are no longer dear to us; they do not have any special value for us. But these feelings of endearment were, *ex hypothesis*, what make life worth living. Therefore, if his theory is normative, it cannot refute the Nearest and Dearest Objection because it forbids us from *caring* more about our friends and family, thus removing that element of life that Jackson himself believes to be most valuable.²⁹

Jackson presented an interesting attempt to reconcile one of the most appealing of all modern ethical theories, Consequentialism, with a devastating objection. This objection is particularly powerful because it accuses Consequentialism of requiring that we live meaningless and unfulfilling lives. Jackson's attempt to show otherwise is a complete failure. His arguments not only inherited the fatal flaws of Decision

Theory, but even if we preserve the qualitative aspects of his theory while dropping the cancerous Decision-theoretic element, the theory still fails to meet the author's own requirements. The initial intuitive appeal of his theory might stem from the fact that he oscillates between prescription and description whenever it is convenient. However, this methodological confusion is of no help in the end, because Jackson's theory, even on the most charitable reading, fails in either case. Of course, this does not prove that the Nearest and Dearest Objection cannot be refuted from a consequentialist perspective, but it is nevertheless clear that Jackson's theory is of no assistance to the committed consequentialist.

Notes

¹ Frank Jackson, 'Decision-theoretic Consequentialism and the Nearest and Dearest Objection,' in *Ethics* 101:3, 1991, 462.

² It is important to note that Jackson is adapting standard Decision Theory for his purposes, rather than inserting it *verbatim*. He replaces the typical DT notion of a value "function," which is understood in individualistic (selfish) terms, with the consequentialist value function, which is agent-independent — that is, impartial. Thus, in attempting to maximise *expected* utility, one is trying to benefit others rather than secure the most goods for oneself.

³ *Ibid.*, 463.

⁴ *Ibid.*, 464.

⁵ *Ibid.*

⁶ *Ibid.*, 465.

⁷ *Ibid.*, 466.

⁸ It is important to note that by introducing the notion of *expected* utility, the Decision Theorist is not forcing us to be irrational utility maximisers. If we grant them their assumptions, their theory does make it more "rational" to play it safe in a wide variety of circumstances, such as Jackson's example. Their point, as far as I understand it, is that because of the epistemological limitations of our subjective probability function, insofar as it is subjective, we are forced to "play it safe" most of the time. This follows simply because the risks of attempting to achieve the maximum possible (as opposed to expected) utility are often too large, and we simply cannot be certain that attempting to gain the maximum possible utility will not be disastrous. However, this superficial merit of the theory quickly disintegrates into sheer confusion as soon as we inquire into the nature of the subjective probability and value functions. See below.

⁹ *Ibid.*, 467.

¹⁰ *Ibid.*, 468.

¹¹ *Ibid.*, 470.

¹² *Ibid.*, 471.

¹³ Michael Stocker, 'The Schizophrenia of Modern Ethical Theories,' in *Virtue Ethics* (New York: Oxford University Press, 1997), 68 *et passim*.

¹⁴ Jackson, 'Decision-theoretic Consequentialism,' 472.

¹⁵ *Ibid.*

¹⁶ *Ibid.*, 473.

¹⁷ *Ibid.*

¹⁸ *Ibid.*, 474.

¹⁹ *Ibid.*

²⁰ *Ibid.*, 475.

²¹ Bernard Williams, 'Consequentialism and Integrity,' in *Consequentialism and Its Critics* (New York: Oxford University Press, 1988), *passim*.

²² Jackson, 'Decision-theoretic Consequentialism,' 476.

²³ Peter Railton, 'Alienation, Consequentialism, and the Demands of Morality,' in *Consequentialism and Its Critics* (New York: Oxford University Press, 1988), 120.

²⁴ Jackson, 'Decision-theoretic Consequentialism,' 479.

²⁵ *Ibid.*, 478.

²⁶ *Ibid.*, 480.

²⁷ *Ibid.*, 462.

²⁸ *Ibid.*, 466.

²⁹ Though it is well beyond the scope of my current project, it is interesting to note that Jackson is dealing with a theory about how we ought to act — an ethical theory — which is often taken to mean that he is dealing with a theory about how to live well. But the contentious issue is whether or not Consequentialism can be construed so as to allow us to act ethically while continuing to live what we currently take to be the good life, as if living ethically and living well were independent issues.

Sources

Jackson, Frank. "Decision-theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101:3, 1991.

Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Consequentialism and Its Critics*. Ed. Samuel Scheffler. New York: Oxford University Press, 1988.

Stocker, Michael. "The Schizophrenia of Modern Ethical Theories." *Virtue Ethics*. Eds. Roger Crisp and Michael Slote. New York: Oxford University Press, 1997.

Williams, Bernard. "Consequentialism and Integrity." *Consequentialism and Its Critics*. Ed. Samuel Scheffler. New York: Oxford University Press, 1988.

Contributors

Tara Mitchell is graduating this year from Trent University with a Bachelor of Arts degree in philosophy. Her interests are metaphysics, music, art, and madness.

Michael Brent recently graduated from the University of Toronto with an Honours Bachelor of Arts degree in philosophy. After travelling abroad, he will pursue a Master of Arts in philosophy.

Carl Joakim Gagnon is finishing an undergraduate degree in philosophy and European history at McGill University. He is not convinced that he has had the last word on the topic of 'authentic' music, but hopes one day to provide some theoretical underpinnings of the very real artistic achievements of the authenticist movement in classical music. On other days, he is more interested in how aesthetic evaluations relate to their moral and political counterparts. He plans to study these things next year in the University of Chicago's Committee on Social Thought.

Mark Sentesy is a philosopher and poet by disposition and a student by profession. He studies at the College of the Humanities at Carleton University, under whose aegis he attended the Institute of Philosophy at the Katholieke Universiteit Leuven in Belgium. Literature, photography, travel, and companionship are other bright lights for his spirit. Upon graduation, he will continue his education for a year or more outside of academia, perhaps working in South America before returning to graduate studies in continental philosophy.

Jared Will is in the final year of a Bachelor of Arts degree in philosophy at McGill University. He is completing an undergraduate thesis on the political thought of J.G. Fichte and G.W.F. Hegel, focusing specifically on the role of the individual in the state. He will be pursuing studies in law, with a primary interest in human rights and international law.