

*British
Journal of
Undergraduate
Philosophy*



Editor: Robert Charleston
Open University

*British
Journal of
Undergraduate
Philosophy*

Journal of the British Undergraduate Philosophy Society

Editor: Robert Charleston, Open University
editor@bups.org

Sub-editor: Andrew Stephenson, Cardiff University
subeditor@bups.org

bjup@bups.org
www.bups.org/BJUP

ISSN 1748-9393

Many, many thanks to:

Professor Robin Attfield
Doctor Alessandra Tanesini
Doctor Richard Gray

And the soon to be doctors:

Andrew Thomas
Esa Diaz-Leon

who all help keep our high standards.

Contents

339. **Editorial: What are you doing?**
345. **Kripke, names, and the necessary a posteriori**
Lorna Finlayson
355. **Two types of certainty**
Steven Methven
363. **Endurantism and temporal gunk**
Andrew Bacon
372. **Evolutionary psychology and the green-eyed monster**
Andrew Goldfinch
385. **Kant on things-in-themselves: one world or two?**
Craig French
396. **Hegel's master and slave**
Ryan Dawson
411. **Russell and Kripke on names as definite descriptions**
Elaine Yeadon
418. **Dennett's stance: legs akimbo**
Alex Carruth
426. **Hume's causal reasoning**
Rhonda Grintuch
434. **Book reviews:**
Bernard Williams – On opera
Andrew Goldfinch
The Routledge philosophers series
Andrew Stephenson
449. **Upcoming BUPS events**
451. **Subscribing and submitting papers to the BJUP**

What are you doing?

Editorial

OK, put aside notions of intentional diffidence or postmodern irony. Do you have a strong reason or belief on *why* you are studying philosophy? Do you need one? What relation should this have to *how* you go about your work? I've been asking these questions a lot recently, and have particularly pestered a fair number of colleagues with the first question over the last few weeks. The answers I've received back have fallen into four distinct categories:

First, of course, you can study philosophy because you *enjoy the material*. That is, you like thinking about how and why language, mind and world work; or find argument structure, technique, and debate fascinating in themselves. Second, you might want to practise those skills, explore those areas of discussion because you believe they will be helpful in *the pursuit of something else* you are interested in, such as law, business or 'the greater good'. Third, studying philosophy might be your *niche* within academia. If you want to stay in education, you need a subject to study, and philosophy can fit the bill nicely if you tend to think in the way it requires. Finally, you may not want to claim any positive reason – plenty of people responded that they just 'fell into' philosophy as a subject.

Most people will naturally find that their motivation is a composite of several of these factors. I would expect that pretty much everyone reading this journal could immediately tick the first 'enjoy the content' reason. BUPS events and the BJUP are characterised by a notable enthusiasm and happiness to be involved in debate – and I think this has to point to people enjoying philosophy *as an activity in itself*. Many of our members go on to study other subjects or chase other, non-academic careers, though. So it also seems fair to assume that people have other goals and aims they want to pursue, and think they can bring philosophical techniques to bear on these other targets.

Some students – and I’ll gladly embarrass myself by admitting I’m one of them – want to do something ‘big’ or ‘worthwhile’ with philosophy. Sometimes I’m not sure what this really means. For instance, sometimes I want to argue that rationality, careful questioning and analysis are good things in themselves – so we do well simply to perpetuate and spread these habits. Other times this seems a little underwhelming, and I want to appeal to the usefulness of philosophy – suggesting that it helps us humans advance as thinkers, as knowledge-users, as technical ‘users’ of the world. Either way, if we pursue philosophy for such reasons we are attempting to make the world ‘a better place’, and it is a genuine worry to me that these thoughts read as naïve, emotive, even wrong-headed in our postmodern, academically separatist, anti-practical-action world. They were definitely serious concerns for some of the greatest writers in our subject – probably amongst the few things that thinkers as disparate as Aristotle, Locke, Kant, Marx and Gadamer have in common. Without wanting to be too *ad hominem*, perhaps we can take their seriousness about such intentions as a mark of at least *some* credibility. If those of us who want to better the state of the world through promoting philosophical thought, analysis, study and discussion are genuinely naïve, emotive and wrong-headed, we are at least in good company. In fact, if the alternative – speaking exclusively to a specialist audience, limiting ourselves to technical questions with no practical action implications, regarding philosophy as essentially inert in the wars, hunger and future development of the world – is the only sophisticated, intellectually proper, ‘right’ position, I think I would honestly rather be ‘wrong’. This may of course just prove my naivety, wrong-headedness etc. But feeling this way has important consequences, as I will argue in a moment.

But I think ‘personal enjoyment’ and ‘big, worthwhile’ considerations aside, we should probably acknowledge that most of us are also driven by the third category of reason to a far greater extent than we might want to admit. In these days of Amazon, external reader cards and Ingenta, pretty much anyone has access to the books and journals working academics do. The internet’s many lists and boards, open conferences and e-submission mean you can find informed readers and discussion if you are prepared to put in some work, no matter whether you are currently registered at a university or not. In fact, given

academia's reputation and standing in society at the moment, you may well find your opinion has greater mainstream credibility as a journalist, business consultant or even just popular blogger than as a tenured academic. In short, pursuing philosophy because you enjoy it, or because you think it is important for some other goal, has no necessary implication as to whether you study independently or within academia. In fact, there are reasons to stay away from university once you have reached a basic level of competence. The lack of fees leaves more money for books, journals and even tuition. Full-time student status looks less cost-effective when you consider than an hour of one-on-one tuition from a doctor of philosophy provided by a specialist agency currently costs about £25. A research MA with 24 hours of supervision through a year costs around £3,000 in tutorial fees. That's quite a mark-up.

So why do we prefer to study philosophy in a formal university setting? The libraries of most universities can be joined for a small fee per year. Tutorial time can be bought per hour from academics qualified at the same departments for less money. Discussion and publishing opportunities are available either way. Most national lectures and conferences are open access. Why do we persist with the formal system?

A few months ago I would not have had much of an answer to this question. But the extra delay forced by various Eastern European hackers attacking BUPS' production infrastructure has made a difference. During this time I have moved from the OU to Cambridge. And wow, it's fantastic. I'm hooked. I don't have to wake up early, go to the office, fix computers and wait until after work to start studying. I can get up at 10.30 if I want to. There are long holidays, no train rides to tutorials, great conversations every day and lots of lectures. I worked in shops and call centres immediately after I left school, but as a graduate I can walk into an office job. Even for those of us nominally 'self-funding', all of this is heavily subsidised by people who (still) have no practical chance to do the same. Make no mistake, academia and the attendant 'pieces of paper' offer status and privilege. I know students often doubt this – and it's true that this status has been diluted in the last twenty years – but it is still considerable and easily demonstrable. Even though I know I could continue studying philosophy using the resources I've mentioned above, I can (shamefacedly) admit that I really

would rather keep all these extra privileges that I've accrued, *even though* they have very little to do with the quality of the work I can output. Privilege is habit-forming. I prefer the extras.

I think it's better to acknowledge these preferences. Not least because they have consequences for us. My old headmaster used to say that the only possible justification for privilege is hard work. And I for one definitely need to work harder than I have recently. Again, this could easily be dismissed as naïve, wrong-headed and so on, but I think it's important that if you find yourself enjoying privileges you ought to at least try to live up to them. If we're honest, I think it's clear that choosing to study philosophy in a formal setting is not all about the subject material, or the selfless consideration of what philosophy can do. It's also about the status and privilege of the lifestyle, and I think that needs acknowledgement and a response.

Of course, a lot of this may be denied by anyone who claims to have *no* positive reason for studying philosophy – those who belong to the fourth 'I just fell into it' category. But I doubt whether many people actually inhabit this category. Philosophy is not an easy subject, and the competition for places and funding can be intense. Doing a good job as a philosophy student is a constant challenge, and slacking-off very noticeable. So the historical explanation 'I sort of fell into it' seems a little insufficient – it leaves the question of why, in the face of such difficulties, such a person *stayed* in the subject. These kinds of 'no real reason' answers are not really plausible, and they represent a missed opportunity for gaining a bit of insight into ourselves and our preferences.

I'm suggesting that if, like me, you sometimes worry about what you're doing studying philosophy, the best answers lie in being honest about your personal preferences, even if they seem not to be particularly justifiable, or even admirable. Philosophy covers enjoyable ground, and has – I believe – some practical, helpful importance in the world. But in the long run, there is a particular answer you need to be able to give if you are to be a success in the subject. Yet it is the answer people were most reluctant to give me when I asked. I think it's worth considering what this answer must involve.

There are professional philosophers who are genuinely unhappy, as their jobs do not even vaguely match their personal preferences. They have too high a teaching or organizational workload, when they are really in academia for their own research. Or they feel too much pressure to publish when they would really prefer to be teaching. Or any number of other incompatible desires. There are students who are genuinely unhappy because they honestly do not enjoy the day-to-day nature and routine of academic life. They do not want to work to such time or word deadlines, do not like to restrict themselves to the course modules, or do not like public discussion, or perhaps even the continual workload. But such unhappy people continue within the system. I've met quite a few over the last couple of years, and I realise the temptation is to blame the system for what it is, to grimace and carry on. But this is irrational. The system is not going to change its character any time soon. And there are many, many great opportunities for building your interests and pursuing what you enjoy *outside* the formal university system at the moment. Unless you are hopelessly addicted to the status and privilege of academic life, then you don't have to put up with it. There are other routes through. Of course, if you *are* hooked on the status and privilege, then being honest about what makes the difference for you should make it easier to feel happy in the face of the stuff you *don't* like.

But if – like me, I suspect – you do not particularly suit the official route, as there's too much you honestly don't enjoy, then please consider doing things *your way* instead. Find a route you'll enjoy. If it seems a little naïve or implausible to suggest such a course of action, I hope BUPS and the BJUP might provide an example of success. They were set up from scratch without mainstream departmental support, precisely because there were certain things that a few students wanted that were not really catered for within the university system. And, I'm glad to say, they've both been a big success. What's interesting is the extent to which these initially extra-mural ideas are now really part of 'the establishment' themselves – and can benefit from funding and sponsorship. You really can start something that works the way you like, do interesting stuff, and win some official recognition in the process. BUPS and the BJUP are proof of that. Without them, I for one

would not have found enough that matched my preferences to continue within the university system. Hopefully, they are a beginning for this kind of positive, developmental attitude, rather than an end. It would be great to see what people would like to do next.

So I apologise for the shameless soapboxing here, but this is rather my last opportunity. Those of you who go on in time to be lecturers, tutors, professors: please, if you have the chance, be open to the different ways of developing interesting, progressive thought. Take it seriously. Don't indulge in pointless snobbery or conformism about academic style. Notice the content. And to those of you who think, like me, that your future probably lies outside the formal academic system: please, don't lose touch or give up your studies. Philosophy needs continuous infusions of new perspectives, questions and approaches to develop true wisdom; and its techniques will help you in whatever new avenues you decide to follow outside academia. Non-academic philosophy is a decent, coherent, respectable notion with an active audience. Stay enthusiastic and optimistic, and if you cannot feel this way in your current context, *change things around*. I am convinced that the most important philosophical work in our lifetime will be done by those who can answer my title question with a simple: '*exactly what I want to*'.

Thank you for a great couple of years.

Robert Charleston

Founder & Chair, BUPS 2004-6

Founder & Editor, BJUP 2005-6

Cambridge, January 2007

Kripke, names, and the necessary *a posteriori*

Winner of the 2006 British Undergraduate Philosophy Conference Prize

Lorna Finlayson

Kings College, Cambridge

lf258@cam.ac.uk

Saul Kripke's *Naming and Necessity* (1972) is probably best known for its revision of the previously assumed link (even equivalence) between the terms '*a priori*' and 'necessary', and between '*a posteriori*' and 'contingent'. Apriority, Kripke points out, is an epistemological notion: something is knowable '*a priori*' if it can be known independently of experience. 'Necessary', for Kripke (and for my purposes too) is a term in modal logic, meaning 'true at every possible world'. These definitions are not – or at least not obviously – equivalent. The same goes for '*a posteriori*' (knowable from experience) and 'contingent' (not 'necessary'). I argue that although Kripke is correct to distinguish these notions, his attempt to weaken the link between them is an erroneous step too far.

Kripke claims to find examples of necessities that can only be known *a posteriori*, as well as *a priori* knowledge of contingent truths.¹ I will focus here on the necessary *a posteriori* – in particular, identity statements between names. I will sketch a view of names and necessity which accepts the Kripkean claim that identity is a logically necessary relation, but rejects the claim that identity statements between names can be necessary and *a posteriori*. I will also suggest that Kripke's view of names as rigid designators is over-simplified.

First, a quick remark about Kripke's view of names and modality. Kripke divides 'designators' (e.g. 'The King of France', 'Aristotle', 'the

¹ As for the contingent *a priori*, Albert Casullo constructs something similar to my view by applying Donnellan's distinction between attributive and referential usage to Kripke's example of the Paris metre bar. See 'Kripke on the A Priori and the Necessary', Albert Casullo, in (Moser, ed.) *A Priori Knowledge*, (OUP, 1987).

tallest man') into the 'rigid' and the 'non-rigid': rigid designators refer to the same object in every possible world; non-rigid designators have different referents depending on the possible world being considered. Kripke holds that proper names are rigid designators.² So, 'Aristotle' refers to the same man in every possible world, even those in which only some other man is *called* 'Aristotle' – this is because it is in *our* language, the language of the actual world, that 'Aristotle' rigidly designates that man. 'The tallest man', by contrast, is non-rigid: it refers to different men in different possible worlds. Kripke points out that if 'A=B' holds true, where 'A' and 'B' are both rigid designators, then 'A=B' expresses a necessary truth. Since 'A' refers to the same thing in every possible world and so does 'B', if they share a referent in *any* world then they share this referent in *every* world. So, the statement 'A=B' expresses a necessary truth. I turn now to Kripke's claim that there are identity statements of this form whose truth can only be known *a posteriori*.

'Hesperus' was the name given to the heavenly body which appeared in a certain position in the evenings, and 'Phosphorus' the name given to the body which appeared in that position in the mornings. For some time the ancient Babylonians thought that 'Hesperus' and 'Phosphorus' referred to different bodies, until they discovered that the two names shared a referent: the planet Venus. Kripke argues that the statement 'Hesperus is Phosphorus' expresses a necessary truth, but one which could only be learned through empirical investigation.³ It was this kind of case, where an identity appears to be discoverable only *a posteriori*, that led to the belief among philosophers that there must be some contingent identity relations between objects – remember that it was taken for granted that aposteriority goes hand-in-hand with contingency.⁴ But Kripke thinks it is simply obvious that if objects A and B are quantitatively identical, then the identity is necessary.⁵ This has been disputed, but only (as far as I know) because of the puzzle posed by Hesperus/Phosphorus-type cases.

² Kripke, S., *Naming and Necessity* (Blackwell, 1981), pp. 48-9.

³ *Naming and Necessity*, pp. 102-5.

⁴ Hume, Leibniz and Kant, despite their major differences, seem to agree on this.

⁵ Kripke's claim is about identities between *objects*, not between an object and a variable satisfier of a description – (variable according to the possible world). For

So, even if Kripke doesn't conclusively show that identity relations are logically necessary, he can strip the less intuitive opposing view of its only motivation if he can show that Hesperus/Phosphorus cases do not pose a problem for his rival view. And that is what he does. He argues that the intuition that Hesperus and Phosphorus might have turned out to be distinct bodies is the product of a misunderstanding. For Kripke, when people say that they can conceive of a possible world in which Hesperus is not Phosphorus, what they are really imagining is a world in which the names 'Hesperus' and 'Phosphorus' turn out to refer to two distinct objects, one of which is not Venus (and thus neither 'Hesperus' nor 'Phosphorus' in the language of the actual world). What they cannot possibly be imagining, says Kripke, is a possible world in which *that object* we called 'Hesperus' – imagine that Kripke points to it – is not identical with *that object* we called 'Phosphorus'. That would be to conceive of the inconceivable: a world in which Venus is not Venus. So, Kripke insists, the identity is necessary as well as knowable only *a posteriori*.

Kripke is challenging a powerful rationale here. The conventional view holds that a necessary truth is knowable *a priori* because we cannot conceive of a possible world in which it is false. We infer from this that there is no such possible world.⁶ Thus we don't need to

example, 'Fred is the tallest man' identifies an object – Fred – with the satisfier of the description, 'is the tallest man'. This kind of identity statement need not express a necessary truth. Of course, 'the tallest man' does designate a flesh-and-blood *object* – Fred; and it is necessarily true *of him* that he is himself and not another object. However, 'Fred is the tallest man' does not express a relation between objects, but attributes to an object the property of satisfying a description. Kripke relies on a very intuitive idea of 'objects' here: they are given *things*; we can point at them (abstract objects require different treatment, but that need not concern us here); and since each thing is necessarily itself, identities between things are necessary. If object A and object B are one identifiable thing, and I point first to A and then to B and say that A might not have been B, then I am pointing at the same thing twice, asserting that it might not have been itself. The absurdity of this is the basis of Kripke's conviction that statements expressing identities between objects always express necessary truths.

⁶ Kripke accepts the conventional rationale's assumption that a *conceivable* world is a *possible* world. I will follow Kripke in accepting this, both because I find the assumption plausible and because accepting it will allow me to argue against the necessary *a posteriori* on Kripke's own terms.

consult our senses in order to know a necessary truth – we can know it *a priori*. Kripke must give us sufficient reason to abandon this ‘conventional rationale’. As he does not attack the rationale directly, Kripke’s argument rests heavily on his examples. I will argue that the example of Hesperus and Phosphorus, far from establishing the Kripkean view, exposes its internal tensions.

Let’s reconsider the Hesperus/Phosphorus case. Those who thought that the identity relation was contingent thought so on the basis of the apparent conceivability of Hesperus not being Phosphorus. Kripke denies that this is conceivable. But in employing this argument to undermine the appeal of contingent identities, I believe that he also undermines his own conclusion. This is the structure of Kripke’s argument:

- (1) It is not conceivable that not-(H=P)⁷
- (2) It is necessary that (H=P) [from (1)]
- (3) It is knowable only *a posteriori* that (H=P)

Therefore, ‘H=P’ expresses a necessary truth knowable only *a posteriori*.

Premiss (1) seems correct. Kripke has given a convincing alternative explanation of the intuition that we can conceive of not-(H=P). But there is a tension between (1) and (3). Normally, if not-*p* is not conceivable, we can know a priori that *p*. For example, we cannot conceive of a married bachelor, and so we know a priori that all bachelors are not married. Yet Kripke seems to be suggesting that the whole of ancient Babylonian society failed to realise a truth of which the negation is inconceivable. He cannot scrap (1) without giving up

⁷ Technically, I should have used ‘H≠P’ to notate ‘Hesperus is not Phosphorus’, since ‘not-(H=P)’ means ‘it is not the case that Hesperus is Phosphorus’, which would arguably be true if Hesperus (or Phosphorus) did not exist. In *Naming and Necessity*, Kripke acknowledges this and is content to let ‘Hesperus is Phosphorus’ stand for ‘If Hesperus exists, then Hesperus is Phosphorus’, and so I use ‘H=P’ to mean ‘If H exists then H=P’. Given this, my use of ‘not-(H=P)’ is appropriate, as it is only true in worlds where Hesperus exists and is not Phosphorus – and Kripke thinks that there are no such conceivable or possible worlds.

either (2) or the thesis that ‘conceivable’ implies ‘possible’ – neither option is attractive.⁸ Both (2) and (3) are essential to his conclusion. He must therefore resolve the tension between (1) and (3) by explaining how, since it is inconceivable that not-(H=P), we cannot just work out in our heads that H=P.

One way of trying to explain this is to point to a gap between conceivability and the thinker’s *awareness* of conceivability. The thinker cannot in fact conceive of Hesperus and Phosphorus being distinct, but that won’t help her to know that they are identical if she still thinks she *can* conceive of them being distinct. But to talk of a conceivability/awareness-of-conceivability gap is misleading. It is rather that people report what they find conceivable in a way which has misled philosophers. This can be seen if we examine more closely the suggestion that people believed that Hesperus was not Phosphorus, since if they believed that this was true they must also have believed that it was conceivable.

What exactly did they believe? Arguably, they had a mysterious and illusory intuition that what I’ll call the ‘impossible world’ – where Hesperus is not Phosphorus, and Venus not Venus – was actual. But it seems that only a crackpot would believe *that* – the kind of crackpot who also believes in married bachelors. A more plausible story is that people believed that a conceivable and possible world where there are two bodies instead of one was actual, but that they reported this as believing that Hesperus was not Phosphorus. This story is supported by the effectiveness of one of Kripke’s arguments. He explains the intuition that we can conceive of Hesperus not being Phosphorus by pointing to genuinely possible worlds which we might mistake for worlds in which Hesperus is not Phosphorus. He would not do this (and it would not be so convincing) if people had not had these worlds in mind when they said they believed that Hesperus was not Phosphorus, or that they could conceive of Hesperus not being Phosphorus. If, instead, the ‘crackpot’ story was true, and they were

⁸ In case it is thought doubtful that Kripke is so attached to the link between conceivability and necessity, it’s worth pointing out that at the very end of *Naming and Necessity* he presents an argument for mind-body dualism which rests on the assumption that what is conceivable is also possible.

convinced of the actuality or conceivability of the 'impossible world', then Kripke's explanation would not wash. One could say, 'Sure, there are those possible worlds too, but what people believed in was the 'impossible world', in which Hesperus *really isn't* Phosphorus (and although we now know that the 'impossible world' is not actual, it is perfectly conceivable)'. But the ancient Babylonians were not crackpots. They didn't, strictly speaking, think that the 'impossible world' was actual, nor even that they could conceive of it. But it is quite possible that they *reported* the conceivability of worlds in which there are two bodies answering to the names 'Hesperus' and 'Phosphorus' by saying, 'Hesperus is not Phosphorus', and later, 'Hesperus might not have been Phosphorus.' If natural language is not always sensitive to the distinctions logicians want to make, this is not particularly surprising. However, if this is right, the names 'Hesperus' and 'Phosphorus' don't operate quite as Kripke tells us. I'll return to this point shortly.

If people know what they can and cannot conceive of, and the confusion comes only from the phrasing of their reports, then we cannot resolve the tension between (1) and (3) so easily. It seems likely that nobody thought they could conceive of the 'impossible world' in which Hesperus is not Phosphorus, and so we might expect people to have known *a priori* that Hesperus is Phosphorus. It seems that the only remaining option is to reject (3). Kripke doesn't argue for (3), seeming to inherit it unquestioningly from those philosophers who argued for contingent identities. But does scrapping (3) just trade one piece of strangeness for another? It seems obvious that the Babylonians didn't know (and *couldn't* have known without empirical evidence) that Hesperus was Phosphorus. I'll argue now, as I have hinted already, that the solution to this problem lies in the interpretation of language.

My suggestion is that when the ancient Babylonians announced, 'Hesperus is Phosphorus,' they used 'Hesperus' and 'Phosphorus' as non-rigid designators. In other words, their announcement expressed this contingent truth:

$(H=P)^{NR}$: There is some x such that x (appears in yonder position in the mornings and in yonder position in the evenings) and for all y , y appears in that position in the mornings or in the evenings if and only if $y=x$.

Now, it is unsurprising that they didn't put it quite like that, when they had the terms 'Hesperus' and 'Phosphorus' already in use to refer respectively to the body they saw in the evenings and in the mornings. Kripke would insist on a rigid reading of 'Hesperus is Phosphorus'. But then, I argue, we not only have a statement of a necessary truth, but a statement expressing a proposition of which the negation is inconceivable, which is equivalent to 'Hesperus is Hesperus', and which is technically knowable *a priori*. The fact that we don't say, 'Ah, but there's a sense in which the ancient Babylonians knew all along that Hesperus was Phosphorus!' suggests to me that Kripke's rigid reading of the statement is not the natural, appropriate one. Here, my criticism of Kripke's work on necessity spills over onto his work on naming. More details on this are given below, but I will first relate my suggestion to Kripke's argument as presented on page 3 of this essay.

We are now able to get rid of the tension between (1) and (3) by scrapping (3), and we can do so without introducing any philosophical strangeness. This is because we can say that whilst $H=P$ is technically knowable *a priori* (and thus (3) is false), the proposition expressed by the statement 'Hesperus is Phosphorus' is not $H=P$, but $(H=P)^{NR}$. It is $(H=P)^{NR}$ that is knowable *a posteriori*, and the failure to distinguish $(H=P)^{NR}$ from $H=P$ accounts for the initial intuitive appeal of (3). Once we make this distinction, we can detect two things that are wrong with the conclusion of Kripke's argument: First, 'H=P' does not (on this occasion – see below) express the proposition $H=P$, as the inference assumes, but $(H=P)^{NR}$ – a contingent truth; second, since we have scrapped (3), we must deny that $H=P$ is knowable only *a posteriori* (although it *is* necessary).

So, we can keep Kripke's premisses (1) and (2), but reject (3) and provide an explanation of why this is not hopelessly counterintuitive. This yields a view which does not suffer from the internal tensions of Kripke's, and is compatible with the 'conventional rationale' I identify

on page 3 above, but which retains Kripke's intuitive claim that identities between objects are necessary. I'll finish by outlining the modified account of names which makes this view possible.

My suggestion was that 'Hesperus is Phosphorus', as uttered by the Babylonians, expresses $(H=P)^{NR}$, because the names are being used as non-rigid designators. Kripke would not agree – his claim is that names are always rigid designators, and it seems that I am either denying this or denying that 'Hesperus' and 'Phosphorus' are both names. In fact, I'm rejecting his way of drawing the rigid/non-rigid distinction. He implies that designators are divided into the rigid and the non-rigid *types*, and says that this division matches that between names and definite descriptions. But when he says that names are rigid designators, he makes a claim about popular use of names. Names are only rigid designators to the extent that we use them to refer rigidly. This much I think Kripke would accept. The point on which I disagree is this: I see 'rigid' and 'non-rigid' not as types of designator but as alternative *levels* on which we can use any designator. For instance, 'Hesperus' can be used on the rigid level in the statement 'Hesperus is Phosphorus', expressing the proposition $H=P$; or on the non-rigid level (abbreviating, 'the body which appears in yonder positions in the evenings'), so that 'Hesperus is Phosphorus' expresses $(H=P)^{NR}$. To say that 'Hesperus' is a rigid designator is roughly correct, because it is *generally used* as a rigid designator. But where an identity is 'discovered' – and whenever else it suits us, in fact – we use names non-rigidly.

I cannot argue extensively for this view of the rigid/non-rigid distinction here, but I will state its two main advantages, and discuss a possible objection.

First, arguing that 'Hesperus' and 'Phosphorus' are (in the context) non-rigid allows me to explain Hesperus/Phosphorus cases in a way which has important advantages over Kripke's. It might be thought that a modified Kripkean 'two types' view of the rigid/non-rigid distinction could achieve this just as well as my 'two levels' view does – i.e. we could say that some names are non-rigid. But there would be occasions when these rogue non-rigid names were clearly being used

rigidly. For example, two ancient Babylonian speakers might simultaneously utter, 'Hesperus is not Phosphorus', and 'Hesperus might not have appeared in the evenings'. The first speaker uses the names non-rigidly, I've argued; but the second does not, since she is making a claim about what would be true of Hesperus in a possible world where the description which *in this world* abbreviates the non-rigid use of the name 'Hesperus' is false of Hesperus. My 'two levels' distinction allows us to give a better explanation of Hesperus/Phosphorus cases than Kripke gives, without running into the difficulty just mentioned for the modified 'two types' view.

The second advantage of the 'two levels' view is that it reflects the fluidity of natural language: there are rigid and non-rigid uses of the same designator by different speakers on different occasions. It strikes me as over-optimistic to assert simply that names are rigid designators. Our use of language is unlikely to conform precisely to such a simple rule, and as I've just argued, we run into difficulties if we try to refine Kripke's 'two types' view by positing exceptional non-rigid names. Note, however, that the 'two levels' view can allow that names are rigid designators, *in so far as they are characterised by a tendency to be used on the rigid level*.

Now for the objection. If the 'two levels' view is correct, how do we decide whether a designator is being used as rigid or as non-rigid on a given occasion? A critic could say that this decision is arbitrary, and that I am interpreting the names in Hesperus/Phosphorus cases as non-rigid for the sole reason that doing so allows me to uphold the conventional wisdom that necessary truths are knowable *a priori*. It is true that the 'two levels' view of names was designed to complement my account of Hesperus/Phosphorus cases. But since the appeal of that account was independent of the appeal of the 'two levels' view of names, there is no circularity. Furthermore, I've argued that the 'two levels' view also enjoys some independent intuitive support. However, there remains the question of how we determine the level on which a given speaker is using a designator. A satisfactory answer to this would require further work, but the main consideration is that we should always assume that the speaker is rational. Thus the decision is not entirely subordinate to the preservation of the 'conventional rationale'.

I have sketched a view which preserves the advantages of Kripke's but eliminates some of its strangeness. Kripke argues that not-(H=P) is inconceivable and therefore that H=P is necessary, but then it becomes obscure why H=P cannot be known a priori. My answer is that it can, but that this doesn't strike us because the statement 'Hesperus is Phosphorus' doesn't always assert only that H=P. It was the truth of the proposition expressed by the *non-rigid* reading of that statement – (H=P)^{NR} – that had to be discovered empirically. By reworking Kripke's distinction between rigid and non-rigid designators, we can construct a view that reflects this. It rests on abandoning the strict equation of names with rigid designators, which on reflection seems too good to be true anyway. Once we see rigid and non-rigid designation as two levels on which statements can operate according to speaker intent, we can interpret Hesperus/Phosphorus cases so as to accommodate both the necessity of identicals and the connection between necessity and apriority. We can also account (as Kripke does) for the illusion that the identity of Hesperus and Phosphorus can be known only *a posteriori*, whilst avoiding the conclusion that the identity must be discovered empirically even though its negation is inconceivable. The 'two levels' view thus enjoys the best of the Kripkean and pre-Kripkean worlds.

Bibliography

Dummett, M., *Frege: The Philosophy of Language*, 2nd Edn. (Duckworth, 1973, 1981)

Kripke, S., *Naming and Necessity* (Blackwell, 1981)

McCulloch, G., *The Game of the Name* (OUP 1989)

Two types of certainty

Steven Methven

Birkbeck College, University of London

steven.methven@gmail.com

Peter Unger seeks to establish universal scepticism by showing that because it is never all right for anyone to be certain of anything, nobody can ever know anything.¹ This argument is valid, but the conclusion that nobody can ever know anything is false. I seek to show which of the premises is false.

Unger's argument is as follows:²

- (1) If someone knows something to be so, then it is all right for the person to be absolutely certain that it is so.
- (2) It is never all right for anyone to be absolutely certain that anything is so.
- (C) Nobody ever knows that anything is so.

(1) is true; if one knows that p then, if one were absolutely certain that p , then that would be all right. One's knowledge that p warrants one having a particular attitude about p , namely that one is certain that p . But being certain is the sort of attitude that may become detached from knowledge, for we encounter people who are absolutely certain that p when in fact not- p obtains. In such cases, it is not all right for that person to be absolutely certain that p , for they are wrong. Also, although knowledge warrants the adoption of the attitude, it need not guarantee it. So there will be cases where S knows that p , but where S is not certain that p , e.g. the schoolboy who knows the answer to the

¹ 'An Argument for Skepticism', *Philosophical Exchange* 1, 4 (1974); reprinted in *Epistemology: An Anthology*, Eds. Sosa, E. and Kim, J. (Blackwell, 2000). All page references are to the Blackwell volume.

² Unger, 1974, p. 42-43.

question, but whose confidence is eroded when questioned aggressively.³ Even in this case, though he would not claim it himself, we may say *of* him that he knows that p and, as such, it would be all right for him to be absolutely certain that p (and he might have been so had the teacher been meeker).

It will be helpful to identify two kinds of certainty. The first we shall call *subjective certainty*, and say of it that:

x is subjectively certain that p iff x has an attitude of certainty that p.

The second we shall call *actual certainty*:

x is actually certain that p iff x is subjectively certain that p and x knows that p.

The use of these terms allows us to see why statements like the following may be true without contradiction:

(Q) Jack was absolutely certain that Everest is in Tibet, even though he didn't know that Everest is in Tibet, which it isn't.

It is true that Jack was absolutely certain about the location of Everest because he was absolutely *subjectively* certain, while it was false that he was absolutely *actually* certain because the proposition that Everest is in Tibet is false.

Some account must be given of how Jack comes to be subjectively certain. If Jack is subjectively certain that p , then it goes without saying that he believes that p . Further, I think that he must also believe to a sufficiently high degree that he is justified in believing that p . Imagine that Jack's uncle is Edmund Hillary, and that the mountaineer dislikes his nephew. Being irritated by him, Edmund informs him, during the course of a story where all other statements are true, that Everest is in Tibet. Jack believes this proposition up to the point of a certain

³ Dancy, J., *Introduction to Contemporary Epistemology* (Blackwell, 1985), p. 24.

encounter, at which the person that he met, call her Jill, made the statement Q above. It is even plausible that, at the encounter, if Jack's belief were expressed with absolute subjective certainty, it could cause doubt in the mind of Jill who holds a true belief about the location of Everest, but who is not subjectively certain of this belief because her true belief does not seem to her to be as justified as Jack's. Jill, for instance, may have learned that Everest is in Nepal while at school, in some half-remembered lesson that she cannot recall. Jack's subjective certainty that Everest is in Tibet could shake Jill's uncertain, albeit true, belief because (a) Jack seems to Jill to be better placed than she to have knowledge, and (b) Jack expresses his belief with absolute subjective certainty.

This description of subjective certainty characterises it as a higher-order state of mind, an attitude that one will have where one has a belief about the level of justification that one has for a particular belief that p . What conditions must obtain for Jack's subjective certainty to be actual certainty? Simply, he must know that p while also being subjectively certain that p . John is Jack's twin, who Edmund is fond of. Edmund tells John the same story, only this time he states that Everest is in Nepal, which is true. John believes as much, and believes that he is justified in believing this because the information is from the best possible source.

The states of Jack and John with regard to their respective certainty can be put as follows:

Jack	John
(A) $\neg K_a(p)$	(A') $K_b(p)$
(B) $B_a(p)$	(B') $B_b(p)$
(C) $B_a(HJB_a(p))$	(C') $B_b(HJB_b(p))$

Here let K =knows, B =believes, HJB =believes with a high level of justification. The meaning of 'a high level of justification' should become clear shortly. Also, I have stipulated that John knows that p , and I do so by allowing that whatever conditions we have for knowledge are satisfied. I don't believe that anything turns on what

these conditions are, for if one substituted for (A') a conjunction of these conditions, it would only trivially affect the description of John's state with regard to certainty. His belief, for example, might be included in the conjunction, in which case this would merely obviate a need to have (B') as a separate condition. The only case I can think of where substitution of knowledge conditions for (A') would have a non-trivial effect would be if one held that subjective certainty was a condition for knowledge; this position, however, is untenable.

John epitomises a case of actual certainty, and Jack a case of subjective certainty. Note that this analysis detaches subjective certainty from actual justification, for to be subjectively certain S need only meet his own evidential standard, which may be at variance with any actual standard. This allows for cases where it is false that S is justified in believing that *p*, but true that S believes that he is highly justified in believing that *p*; cases, for example, of S being subjectively certain that *p* where he has been told that *p* by an alien voice, or read it in the cards. It also allows for cases where S has very high evidential standards, cases where it might be true that S is justified in believing that *p* but false that S believes that he is highly justified in believing that *p*.

Unger's second premise is:

- (2) It is never all right for anyone to be absolutely certain that anything is so.

In order to reach his conclusion from (1) and (2), Unger must show why it is never all right to be absolutely certain that anything is so, and cannot do so by claiming that it is because we can never know that anything is so, for this would beg the question. I think, then, that Unger is pressing on what I have identified as the other conjunct in the necessary condition for actual certainty i.e., subjective certainty. Thus, I think that Unger must argue that it is never all right for anyone to be absolutely *actually* certain that *p* *because* it is never all right for anyone to be absolutely *subjectively* certain that *p*. Under what conditions is it not all right for someone to be subjectively certain that *p*? Only if either it is not all right for them to believe that *p*, or if it is not all right for them to believe that their belief that *p* is highly justified. The former

condition may be immediately rejected for, if we hold that knowledge entails belief, to say that it is not all right for someone to believe that p is to disallow knowledge, which is question-begging. And even if we don't believe that knowledge entails belief, or wish to reserve judgement, to say that it is never all right to believe that anything is so seems to be making a demand that, even assuming it could be convincingly argued for, could not be met. Thus we must look to the latter condition.

If I am right, then the argument will be that it is never all right for Jack to believe that his belief that p is highly justified. Undeniably it will be all right for Jack to believe that his belief that p is highly justified if it is the case that his belief that p is highly justified. And this is precisely the line that Unger takes; for Unger, it is not all right to have this attitude because our beliefs are never *really* highly justified. Why? Because there are a number of possibilities that, were they to obtain, would show that our most certain beliefs are not justified⁴. By this argument, one is only ever barely justified in holding a belief because the number of possibilities that might obtain *against* the belief outstrip any amount of evidence that one has *in favour* of the belief.

For Unger, any claim of certainty is an expression of dogmatism in that a person who purports to be certain that p has excluded from their considerations the possibilities that would cause one to doubt that p . Certainty is then characterised as an attitude that it is never all right to have, for to have it is to deny that these possibilities are in fact possibilities; to be certain is to be dogmatically closed-minded:

One thing which must be entirely absent [for certainty], and which is, I think, implicit in the absence of all doubt, is this: any openness on the part of the man to consider new experience or information as seriously relevant to the truth or falsity of the thing. In other words, if S is certain that p , then it follows that S is not at all open to consider any new experience or information as relevant to his thinking in the matter of whether p .⁵

⁴ Unger, 1974, p. 49-51.

⁵ *Ibid.*, p. 47.

On my view, Unger is here referring to subjective certainty, for if S is actually certain that p then, if being certain does indeed entail not being at all open to consider any new information as relevant to one's thinking in the matter of whether p , it *is* all right for S to have this attitude in the matter of whether p , for S *knows* that p . This is simply Unger's first premise. But does being certain, subjectively or actually, entail this attitude?

Yes, but this is not the end of the matter. Subjective certainty precludes doubt about the proposition in question, for it is inconsistent to say: 'I am absolutely certain that p , though I have my doubts'. On the other hand, it is not inconsistent to say: 'I am absolutely certain that p , though I intend to keep an open mind'. But Unger thinks that being certain that p precludes being open-minded *about whether p* , and in this he is correct. This, however, does not mean that being certain that p precludes being open minded about whether one's belief that p is highly justified or not. Thus when S says consistently: 'I am absolutely certain that p , though I intend to keep an open mind', S is indicating an open-mindedness about his standard of justification that is consistent with his being absolutely certain that p .

Let us characterise doubts as beliefs that serve as evidence to the contrary of a currently held belief that p , i.e. evidence that $\neg p$. Thus Jack may believe that Everest is in Tibet, but also hold doubts about this proposition as when he has come across some evidence to the contrary e.g. someone else has asserted that Everest is in Nepal. If such conditions obtain, we may say that though Jack currently believes that Everest is in Tibet, he is not certain about it because he also believes that there is some evidence to the contrary.

If S is certain that p , S will not be able to get into Jack's state as described above. Unger, quoting Malcolm, writes that certainty that p is such that 'there is nothing whatever that could happen in the next moment that would by me be called evidence that $\neg p$ '⁶. So, if S is certain that p , his attitude is such that he is not disposed to accept

⁶ *Ibid.*, p. 44.

evidence that $\neg p$ as evidence that $\neg p$ under any circumstances because he takes himself to know that p irrefutably. If such a strong attitude obtains, then S will not believe that there is any evidence that $\neg p$ at all. In order for any evidence that $\neg p$ to be efficacious on S's subjective certainty that p , S's attitude would first have to be one that would at least recognise the evidence as evidence. Since S is subjectively certain that p , this is not the case. On such a view, it is easy to see why Unger thinks this attitude dogmatic i.e. indicative of the absence of an open mind. But such dogmatism need not be pathological, for open-mindedness may also be characterised as a second-order attitude about our standards of justification, which is consistent with subjective certainty.

Jack is subjectively certain that p , and is thus not disposed to consider as evidence any evidence to the effect that $\neg p$. But this does not mean that he will not consider as evidence any evidence to the effect that his belief that p is not highly justified. Edmund's attitude towards him might come to light, in which case he would no longer believe that his belief that p was highly justified. In such a case he would no longer be subjectively certain that p .

What about Unger's examples? Is it never all right to believe that our beliefs are highly justified because there are always a number of possibilities which, were they to occur, would be evidence against our belief that p ? No. For why should a possibility impact upon my standard of justification? It seems as though I would not be justified in allowing simply the possibility that $\neg p$ as evidence against the belief that p , unless the possibility that $\neg p$ had evidence in *its* favour. A mere possibility is not, in itself, evidence in favour of anything, and as such it seems that it would be *irrational* to allow a proposition for which there is no evidence at all to have an impact on my belief in propositions for which I believe myself to be highly justified.

(2) is false: sometimes it is all right to be certain that something is so, for to be certain that something is so is *not* to have the dogmatic attitude that Unger supposes. We therefore refuse Unger his conclusion.

Bibliography

Dancy, J., *Introduction to Contemporary Epistemology* (Blackwell, 1985)

Malcolm, N., 'Knowledge and Belief' in *Knowledge and Certainty* (Englewood Cliffs, New Jersey, 1963)

Unger, P., *An Argument for Skepticism*, (1974) reprinted in *Epistemology: An Anthology*, Eds. Sosa, E. & Kim, J. (Blackwell, 2000)

Endurantism and temporal gunk

Andrew Bacon

Lady Margaret Hall, Oxford

andrew.bacon@lmh.ox.ac.uk

Endurantism is the thesis that at any given instant of time physical objects exist wholly in the sense that they, and all their parts, exist at that time. Endurantists will typically hold that physical objects are three dimensional, only have spatial parts and will believe in diachronic identity in the strictest sense. In particular they will deny that physical objects are extended in the temporal dimension or that they have temporal parts.

It is important here to know what the endurantist means by a temporal part. For her opponent, the perdurantist, objects are taken to be four dimensional regions of space-time, or the material contents thereof. A temporal part of an object is therefore taken to mean a particular four dimensional region of space-time, or its material contents – one which is a sub-region of the intended object, is extended in the temporal dimension, and fills the spatial regions the intended object fills, i.e. reaches its spatial boundaries. A temporal part could also be a time slice of the object, an instantaneous 3D snapshot of the object at a time, in which case it would not be extended along the temporal dimension since it is only one instant wide. We shall discuss whether objects have time slices as parts later. The endurantist on the other hand, does not believe in temporal parts in this sense. She will concede to an object having different parts at different times, but it is only under this weak interpretation that she accepts temporal parts. For example, to describe the loss of a hand, an endurantist might say: at time t object o has part p , and at time $t' > t$, o does not have part p . However, endurantists deny that I have any temporal parts *now*. For an endurantist it *will be the case* that I have different spatial parts, but it is not the case that I *now* have a temporal part with the corresponding spatial difference. If the endurantist is said to believe in any kind of temporal part, it would be the three dimensional time slice that exists now.

Perhaps the way to express the difference between perdurantism and endurantism is that one and not the other posits four dimensional objects and four dimensional parts of objects. This might be too restrictive for the endurantist, for she might still believe in the possibility of space with more than three dimensions, perhaps even infinite dimensional space. It might even be possible for an endurantist to make sense of the notion of many temporal dimensions. Either way, we can make the distinction that perdurantists believe in objects which are extended in the temporal dimension(s) and endurantists do not.

If it is true at all, endurantism must be necessarily true. It appears to be making a fundamental claim about the *nature* of physical objects and time, that such and such a thesis is a feature of a particular ontological category. If something is a characteristic of an ontological category, then how can an entity of this category exist without this feature? There are various attacks on endurantism, some involving the indiscernability of identicals, others approaching the problem from modern physics. In this paper I shall consider possible worlds in which endurantism cannot be true. If such worlds exist, then endurantism, as a necessary claim about the nature of objects and how they persist, is false.

Mereology

So far we have talked about parts (temporal or otherwise) intuitively - it is now time to get bit more rigorous. The area of formal metaphysics dealing with the notion of parthood is called *mereology*. Mereology refers to a collection of axiomatic systems for capturing various inferences about the relation of part to whole. To get a taste, I shall outline a fairly weak mereology. Our language is first order¹ with the only non-logical symbol being ' \leq ' to be read as 'is a part of'. For

¹ We can have second order mereologies if we wish. This is useful if we want to formulate unrestricted fusion without an axiom schema. The problem with using a first order language here is that, in a mereology with atoms, we expect the size of the universe to be 2^κ for some cardinal κ . If κ is finite so is the domain, and if κ is infinite, the domain is uncountable, so either way the domain is never countably infinite. If the mereology is gunky then the universe is always uncountable. But for first order languages there are always countable models if there are infinite models (due to the Löwenheim-Skolem theorem), so first order mereology will always have unintended models.

convenience, we also introduce the defined terms ‘ \bullet ’, ‘ \perp ’ and ‘ $<$ ’ to be read as ‘overlaps with’, ‘is disjoint from’ and ‘is a proper part of’:

<i>Overlap</i>	$x \bullet y \leftrightarrow \exists z[z \leq x \wedge z \leq y]$
<i>Disjoint</i>	$x \perp y \leftrightarrow \neg x \bullet y$
<i>Proper Part</i>	$x < y \leftrightarrow [x \leq y \wedge \neg x = y]$

The axioms are:

<i>Reflexivity</i>	$\forall x x \leq x$
<i>Anti-symmetry</i>	$\forall x \forall y [[x \leq y \wedge y \leq x] \rightarrow x = y]$
<i>Transitivity</i>	$\forall x \forall y \forall z [[x \leq y \wedge y \leq z] \rightarrow x \leq z]$
<i>Supplementation</i>	$\forall x \forall y [\neg y \leq x \rightarrow \exists z[z \leq y \wedge z \perp x]]$
<i>Product</i>	$\forall x \forall y [x \bullet y \rightarrow \exists z \forall w [w \leq z \leftrightarrow [w \leq x \wedge w \leq y]]]$
<i>Sum</i>	$\forall x \forall y [\exists z [x \leq z \wedge y \leq z] \rightarrow \exists z \forall w [w \bullet z \leftrightarrow [w \bullet x \vee w \bullet y]]]$

The first three axioms say that parthood is a partial order. That is:

<i>Reflexivity</i>	Everything is a part of itself
<i>Anti-symmetry</i>	If x and y are parts of each other, they are the same
<i>Transitivity</i>	If x is a part of y and y a part of z , then x is a part of z

Conjoined with the assumption that there is something which is a part of everything, and something of which everything is a part, the next three axioms show we are dealing with a Boolean Algebra². Of course these assumptions are controversial, but Boolean Algebras are useful when you are trying to give models of Mereology, as Tarski noted.

<i>Supplementation</i>	If y isn’t a part of x , then there is a remainder: a part of y disjoint from x
<i>Product</i>	If x and y overlap, there is an object where they intersect (the biggest object which is part of both)

² In the presence of these assumptions we can drop the antecedent of *Product* and *Sum* since every pair of objects overlap and underlap.

Sum If x and y are part of something bigger
 (they underlap), there is an object which
 consists of x and y only

We can make our mereology even stronger if we add the axiom of unrestricted fusion, which allows us to collect arbitrary objects into one object (this is actually an axiom schema). An even stronger axiom says that this fusion is unique:

Unrestricted Fusion $[\exists x\phi \rightarrow \exists!y\forall z[z \bullet y \leftrightarrow \exists x[\phi \wedge x \bullet z]]]$
 For any well formed formula ϕ with no
 free occurrences of y or z

As I have already mentioned, to see that these axioms are consistent, we need only take just about any Boolean Algebra as a model – the one most appealing for our purposes would be to consider sets of points in 3D or 4D Euclidean space and interpret \leq as \subseteq (subset). One of the reasons for being concerned with the formal notion of parthood, is that it not only allows us to be rigorous with our reasoning, but allows us to formalise otherwise slippery notions. David Lewis has coined the term ‘atomless gunk’ for one such notion. We say that an object is gunky iff all its parts have proper parts, and we say an object is an atom iff it has no proper parts. Given this we can consistently add one of two mutually incompatible axioms to our mereology:

Atoms $\forall x\exists y[y \leq x \wedge \neg\exists z[z < y]]$
Gunk $\forall x\forall y[y \leq x \rightarrow \exists z[z < y]]$

One says that every thing is made up of atoms: simples which have no proper parts. The other says that there are no atoms, everything is made up of gunk, which in turn is made up of more gunk and so on and so forth - turtles all the way down. Each but not both of these axioms can be consistently added to the preceding axioms of mereology, i.e. they are independent. To see this simply consider two models: for *Atoms* take sets of points in Euclidean 3-space and subset as parthood, our

atoms will then be the singleton sets. For *Gunk* take the non-empty regular open³ sets in Euclidean 3-space and subset as parthood.

Gunk worlds

The possibility of a gunky universe is truly an exotic possibility, but it is exactly this possibility I wish to consider. The notion of gunk certainly seems to be *logically* consistent from the above argument. It is also conceivable and physically plausible (for example, Leibniz thought our universe was gunky). Gunk therefore passes three very strong tests for possibility.

One way in which a universe could contain gunk would be if its space-time was gunky. In fact Russell's co-author on the 'Principia Mathematica', Alfred North Whitehead, thought space and time was gunky. However it is possible that physical objects are gunky even in a space-time which is made of atoms (space-time points). Much like our model for *Gunk* we can just stipulate that no physical object has parts which are not extended in one or more dimensions. When we specified our model for *Gunk*, the underlying set, Euclidean space, was made of points, or atoms. For example this rules out having time slices as parts as I mentioned earlier. Similarly if you are a relationist about space or time, i.e. you think that space is reducible to objects and their spatial relations, then there is no reason to think that it is impossible for objects to be made of gunk and space of atoms.

Let us focus now on spatio-temporally gunky objects. If physical objects are spatially and temporally gunky then there are no three dimensional objects. According to endurantism all physical objects are three dimensional. Therefore, in these worlds there can be no physical objects. We must therefore conclude that endurantism is false. To summarise, in a world where physical objects are gunky the following holds:

1. Physical objects have spatial parts

³ A set X in 3-space is said to be open iff for all $x \in X$, we can find a $\varepsilon \in \mathbb{R}$, such that $B_\varepsilon(x) \subseteq X$ where $B_\varepsilon(x)$ is a sphere of radius $\varepsilon > 0$ with centre x , called the epsilon ball around x .

2. If physical objects have spatial parts, they have spatio-temporal parts. (Spatial parts *are* spatio-temporal parts)
3. Physical objects are spatio-temporally gunky
4. If physical objects are spatio-temporally gunky and have spatio-temporal parts, there are no three dimensional objects⁴
5. If endurantism is true all physical objects are three dimensional
6. Therefore, if endurantism is true, there are no physical objects

Although this argument appears to be valid, it could be construed as question begging. To say that an object is spatio-temporally gunky implies that all its temporal parts have proper parts. This sits consistently with physical objects having no temporal parts, but if we want to allow them to have spatial parts at a time slice we have an inconsistency, since endurantism would imply that the object occupied an instant wide time slice, and nothing more. It could be argued that we have sneaked the existence of temporal parts into the third premise. Furthermore, the third premise could well be false in the actual world - I only introduced it as a possibility. This is a problem because endurantism is necessary and gunky objects are possible, but of course the endurantist could accept the soundness of the argument and simply claim that worlds which have these spatio-temporally gunky things are simply devoid of physical objects. Whatever these gunky entities are, they aren't physical objects.

For these reasons, arguments from spatio-temporally gunky objects are unconvincing. Let us now turn to Whitehead's idea that space-time itself is gunky. This leads to a much more powerful objection to endurantism. If space-time is gunky all its parts are four-dimensional, since each region of space-time will contain a four-dimensional ball (i.e. is open). This conjoined with the commonly accepted fact that objects have spatial parts⁵ quickly contradicts endurantism. In any world in which space-time is gunky:

⁴ To see this it is helpful to think of our model for *Gunk*. Every physical object corresponds to a non-empty open subset of \mathbb{R}^4 . They are non-empty because they have spatio-temporal parts. Let X be the set of points corresponding to a physical object, then since X is non-empty it contains an element x . By the openness property there is a 4D epsilon ball around x inside X , therefore X has a 4D subset (i.e. a 4D part, which means X itself is 4D).

⁵ See [7] for an alternative view.

1. Physical objects have spatial parts
2. If physical objects have spatial parts they overlap with space-time
3. All parts of space-time are four-dimensional
4. Therefore physical objects will have four-dimensional parts

So far we have reasoned from the fact that objects have spatial parts to the conclusion that they have temporal parts in gunk worlds. To talk about objects being spatio-temporally gunky or space-time being gunky involves treating space-time as some kind of connected entity. However, treating space-time in this way seems to be at odds with the endurantist view of space and time in which time is fundamentally different from space. Our discussion of spatio-temporal gunk has required a minimal amount of four dimensional talk. If the endurantist rejects the coherency of such talk altogether then there is little hope of formulating the argument. However, I feel that a theory that does not have the vocabulary for four-dimensional talk, whether or not it has a four dimensional metaphysics, is an untenable theory. It is in this vocabulary that most modern physics is phrased, and if the endurantist world view does not account for this then it is so much the worse for it.

Let us try to respect the endurantist view of time, and briefly note an account in which time is gunky, but the spatial dimensions may or may not be gunky. For a good account of temporal gunk see [5] and [8] especially chapters I.3 and I.4. For a physical object, x , the endurantist claims that at any instant t , x is wholly present at t . Here a more general difficulty presents itself: if time were gunky there would be no instants, so how would the endurantist phrase her position? It is tempting to say that at any interval of time, (a, b) , x is wholly present at (a, b) . For time with instants, Merricks [4] suggests we analyse 'x is wholly present at t' as 'all of x's parts exist when time t is present'.⁶ The equivalent analysis with intervals either commits us to temporal parts, or the notion of time indexed properties breaks down altogether for the endurantist.

⁶ There are slight problems with the notion of 'present' in gunky time. If the interval (a, b) is present, why is it that intervals contained in or containing (a, b) are not present. Perhaps there is no unique present? These matters seem to pose problems for presentism or even the possibility of gunk time depending on your sympathies.

The possibility of spatio-temporal gunk poses some puzzles for the endurantist. First and foremost, there is the problem of coherently stating the endurantist position when time is gunky. Secondly there is the problem of squaring spatio-temporal gunk with a disbelief in temporal parts and a belief in spatial parts. Two strategies suggest themselves naturally. One would be to deny the possibility of temporal gunk. Another interesting response would be to deny that physical objects even have spatial parts, so that physical objects are simple yet extended in space. If physical objects were completely disjoint from space-time the arguments above would fail. Simons discusses the possibility of extended simples in [7]. This idea that objects and space-time do not overlap also seems to be entailed by relationism about space-time. Whether or not such theories will stand up to scrutiny is an open question, however, for those of us who cannot stomach this rather radical notion of extended simples, there are few options but to accept temporal parts.

Bibliography

Bacon, A, 2006, 'First Order Mereology and Unintended Models', <http://users.ox.ac.uk/~lady1900/writings.htm>, unpublished

Boolos, G., 1984, 'To Be Is To Be the Value of a Variable (or To Be Some Values of Some Variables)', *Journal of Philosophy* 81: 430-449

Casati, R., and Varzi, A., 1999. *Parts and Places: the structures of spatial representation*. MIT Press

Clarke, Bowman, 1981, "A Calculus of Individuals Based on 'Connection'", *Notre Dame Journal of Formal Logic* 22:3, 204-217

Goodman, N, 1977 (1951). *The Structure of Appearance*. Kluwer

Hawley, Katharine, 2004, 'Temporal Parts', *Stanford Encyclopedia Of Philosophy*

Heller, Mark, 1984, 'Temporal Parts of Four-Dimensional Objects', *Philosophical Studies*, 46: 323-334

Leonard, H.S., and Nelson Goodman, 1940, "The calculus of individuals and its uses," *Journal of Symbolic Logic* 5: 45-55

Leśniewski, Stanisław, 1992. *Collected Works*. Surma, S.J., Srzednicki, J.T., Barnett, D.I., and Rickey, F.V., eds. and trans. Kluwer

Lewis, D. K., 1991, *Parts of Classes*, Oxford: Blackwell

Merricks, Trenton, 1994, 'Endurance and Indiscernibility', *Journal of Philosophy*, 91: 165-84

Needham P., 1981, 'Temporal Intervals and Temporal Order', *Logique et Analyse* 24: 49-64

Simons, Peter, 1987, *Parts: A Study in Ontology*, Oxford: Clarendon

Simons, Peter, 2004, 'Extended Simples: A Third Way Between Atoms and Gunk', *The Monist* 87: 371-84

Tarski, Alfred, 1984 (1956), "Foundations of the Geometry of Solids" in his *Logic, Semantics, Metamathematics: Papers 1923-38*

van Benthem, J, 1983, *The Logic of Time*, Dordrecht: Reidel (1991)

Varzi, A, 2003, "Mereology", *Stanford Encyclopedia Of Philosophy*

Varzi, A, 2004, "Boundary", *Stanford Encyclopedia Of Philosophy*

Whitehead A.N., 1929, *Process and Reality*, New York: Macmillan

Whitehead, A.N, 1919. *An Enquiry Concerning the Principles of Natural Knowledge*. Cambridge Uni. Press. 2nd ed., 1925

Woodger, J., and Corcoran, J., eds. and trans. Hackett

Evolutionary psychology and the green-eyed monster

Andrew Goldfinch

London School of Economics

a.m.goldfinch@lse.ac.uk

Evolutionary psychologists hypothesize that jealousy is a psychological mechanism – an adaptive solution to pressures and conditions faced by our ancestors in the Pleistocene epoch. From this explanation, evolutionary psychologists have made predictions concerning sex differences in jealousy. The aim of this paper is to identify and evaluate the concepts and methods used to derive this hypothesis and to evaluate its explanatory and predictive power. Sections 1 and 2 outline the concepts and methods used in the explanation; how they are deployed to provide an adaptationist explanation of jealousy is explored in section 3. Sections 4 to 5 then offer a critical analysis of the concepts, methods, and explanation.

1. The conceptual apparatus of evolutionary psychology

Evolutionary psychology is an adaptationist programme: it seeks to identify traits in humans that are psychological adaptations. An adaptation¹ is an inherited trait selected by natural selection because it resulted in the increased reproductive success (fitness) of the individual possessing it in the environment in which it was selected. These naturally selected traits or adaptations increased fitness on average, relative to other possible traits obtainable by natural selection.

Evolutionary psychologists hold that adaptations are designed by natural selection to perform the function of solving adaptive problems. More precisely, to say that the function of adaptation *X* is to solve

¹ Adaptations must be distinguished from their by-products. By-products are traits that do not solve adaptive problems; they are present simply because they are connected to traits that were selected for.

adaptive problem Y is to claim (a) X is there because it does Y and (b) Y is a consequence of X 's being there (Wright 1973). An adaptationist explanation is thus a functional explanation.² Broadly, functional explanations seek to explain the existence or prevalence of something in terms of its beneficial consequences. A functional explanation in our context is an explanation in which the beneficial effects of the adaptation explain the existence of the adaptation. The causal mechanism that links the beneficial effects of the adaptation with its existence is natural selection. By natural selection, selected traits spread through a population.

The environment of evolutionary adaptedness (EEA) refers to the conditions in which an adaptation evolved. More precisely, the EEA is the 'statistical aggregate of selection pressures over a particular period of time that are responsible for the emergence of an adaptation' (Buss et al. 1998, p. 536). The EEA of the structure of the human mind is the set of selection pressures and conditions that prevailed during the Pleistocene epoch (approximately 1.8 million to 10,000 years ago). Evolutionary psychologists hold that the Pleistocene selection pressures and conditions that constitute the EEA of the adaptations in the human mind include the following: nomadic or semi-nomadic hunter-gatherer lifestyle, low population density, competition for mates and resources, high infant mortality, and low life expectancy relative to modern standards. Hunter-gatherer society constitutes 99% of humanity's evolutionary past. Agriculture has only been around for approximately 10,000 years, insufficient time, according to evolutionary psychologists, for significant changes to occur in our gene pool (Symons 1979).

Evolutionary psychologists posit that different psychological adaptations – also called 'psychological mechanisms' and 'modules' – were designed by natural selection for solving different adaptive problems. A psychological adaptation is held to be a distinct, domain-specific, functionally autonomous and specialised information processor devoted to solving a specific adaptive problem in its EEA. Our ancestors faced a great many adaptive problems in the EEA; hence

² There is a debate as to whether teleological terms in evolutionary explanation can be eliminated. For a good summary of teleological notions in biology, see Allen (2004).

evolutionary psychologists hold that the mind contains many psychological mechanisms – hundreds or even thousands. This is known as the ‘massive modularity’ thesis of the mind.

2. The methods of evolutionary psychology

Having outlined the conceptual framework of evolutionary psychology, we can now look at how evolutionary psychologists provide an explanation using these concepts. Two related methods used by evolutionary psychologists can be distinguished: the explanatory approach³ and the predictive approach (Barkow et al. 1992). Broadly, the explanatory approach is characterised by reasoning from the present to the EEA, whilst the predictive approach reasons from the EEA to the present.

The explanatory method. We can formulate this method into three stages. First, one identifies a particular universal behaviour of interest that is capable of being inherited (the explanandum). Second, one looks back to the EEA and theorizes how developing the behaviour would increase the fitness in the EEA. If it would increase fitness by solving an adaptive problem then it would probably have been selected. Third, one hypothesizes a psychological adaptation that could generate the behaviour. This psychological mechanism adapted in the EEA explains the behaviour in the present.

The predictive method. First, one constructs a model of the EEA. Second, one predicts behavioural traits that could increase fitness in the EEA model. Third, one hypothesizes a psychological adaptation that could generate the predicted behaviour. Fourth, one tests for the predicted behaviour. If the predicted behaviour is confirmed then the hypothesized psychological adaptation is held to be the explanation of the predicted behaviour.

³ Also known as ‘reverse engineering’.

3. Evolutionary psychology's explanation of jealousy

Having outlined the conceptual resources and methods available to evolutionary psychologists, we shall now examine how they are deployed to provide an explanation of jealousy. Jealousy is defined as an emotional 'state that is aroused by a perceived threat to a valued relationship or position and motivates behaviour aimed at countering the threat.'⁴ (Daly et al. 1982⁴) In a nutshell, evolutionary psychologists, such as Buss, claim that 'jealousy is an evolved adaptation, activated by threats to a valuable relationship, functioning to protect it from partial or total loss.'⁵

Applying the explanatory method to the explanandum of jealousy yielded the following reasoning. In the EEA, the fitness of an individual would have been under constant threat by the possibility of a partner's infidelity. Those who possessed the trait of jealousy in the EEA would have enjoyed an advantage over those who didn't. The trait would have motivated behaviour that would have discouraged, limited, or prevented infidelity. It is easy to conceive what kind of behaviour this would consist of. Jealousy can motivate an individual to give more attention to their partner; it can also motivate an individual to inflict threats, intimidation and violence on the partner or the rival (or both). Natural selection would select jealousy to alert individuals to infidelity threats, being so designed that it would be activated when threats are detected.

Analysing the selection pressures in the EEA led evolutionary psychologists to a further hypothesis: that there are sex differences in jealousy. They predict that:

Men and women differ psychologically in the weighting given to sexual and emotional cues that trigger jealousy, such that (i) men more than women become upset at signals of sexual infidelity [...]

⁴ As quoted in Buss et al. 1992, p. 251.

⁵ Buss and Haselton 2005, p. 506.

*and (ii) women more than men become upset at signals of a partner's emotional infidelity*⁶

Sociobiology and Evolutionary psychology were the only social science programme to predict sex differences in jealousy (Buss et al. 1996). The reasoning that generated this prediction was as follows. Men and women faced different adaptive problems with regard to infidelity in the EEA. In an age before DNA tests, a man could never be absolutely certain that the offspring of his partner was his. A woman's infidelity would reduce a man's fitness in at least two ways: it would deprive him of opportunity to reproduce for a period of time (especially harmful when life expectancy was limited) and, possibly more damaging, for years he would be investing his resources in a genetically unrelated child. In contrast, the problem for women in the EEA was not uncertainty about the genetic relatedness of children that they gave birth to (every child a woman gave birth to was genetically related to its mother). Rather, the problem facing women was that they were heavily dependent on the resources of their partner and it was always possible for their partner to become emotionally attached to other women. If emotional infidelity occurred then vital resources and protection could be diverted away from her and her offspring to the rival woman and any offspring resulting from the infidelity, and so reduce her fitness.

So sexual and emotional infidelity threatened the reproductive success of men and women in different ways. These different pressures affecting the sexes led to different solutions, namely sex differences in the design of the jealousy mechanism that looks out for potential infidelity. Signs of sexual and emotional infidelity will, of course, upset both men and women. However, the jealousy mechanism for men is so designed that it will be 'particularly activated'⁷ by signs of sexual infidelity; whereas for women, the jealousy mechanism is so designed that it will be 'particularly activated' by signs of emotional infidelity (Schützwohl 2005, p. 289).

⁶ *Ibid*, p. 506, original italics.

⁷ Or 'triggered primarily' as Buss puts it (Buss 1994, p. 128).

Thus evolutionary psychologists utilise both methods outlined in the previous section to (i) provide an explanation of jealousy as a psychological mechanism selected to solve the problem of infidelity in the EEA and (ii) to predict differences in jealousy, providing an explanation of this as being caused by sex differences in the design of the jealousy mechanism due to different pressures in the EEA.

4. Problems with the EEA and methodology

The EEA premise does a lot of explanatory work, but how sound is it? Evolutionary psychologists use the EEA of the Pleistocene epoch as the sole set of conditions and pressures that shaped the human mind. Referring to the 10,000 years since the end of the Pleistocene period, Symons held that 'insufficient time has elapsed since the invention of agriculture 10,000 years ago for significant change to have occurred in the human gene pools' (1979, p. 35). Tooby and Cosmides echo this: 'it is unlikely that new complex designs [...] could evolve in so few generations' (Barkow et al. 1992, p. 5). However, latest research reveals that over 700 genes have been selected during the past 10,000 years⁸. Conceivably, 700 genes are together sufficient to engineer or adapt neural pathways into a psychological mechanism to generate a behaviour. These new findings should make us cautious about assuming that all adaptations in the human mind are a product of the EEA of the Pleistocene epoch⁹. However, in all likelihood, if jealousy is an adaptation, then it is probably a product of pressures that obtained during the Pleistocene epoch rather than any that have obtained since the recent agricultural revolution. Certainly it is difficult to conceive why jealousy would arise during the past 10,000 years rather than in the Pleistocene epoch.

⁸ New Scientist, 11th March 2006, pp. 5 & 30-33.

⁹ Evolutionary psychology's predecessor - sociobiology - didn't assume this. Indeed, Wilson warned precisely against this assumption: 'The theory of population genetics and experiments on other organisms show that substantial changes can occur in the span of less than 100 generations, which for man reaches back only to the time of the Roman empire...it would be false to assume that modern civilizations have been built entirely on capital accumulated during the long haul of the Pleistocene' (Wilson 1975, p. 569). In my view, any new wave of evolutionary explanation of social norms and behaviour that succeeds evolutionary psychology will need to abandon this assumption.

Another objection that can be levelled against the EEA premise is that many details of the events and pressures faced by our ancestors during the Pleistocene epoch are unknown, that indeed the pressures probably varied considerably during that epoch. Whilst this might be problematic for hypothesized adaptations that depend on finely balanced configurations of the EEA that we don't know for certain, it doesn't present a significant problem for the jealousy mechanism hypothesis. This is because its rationale largely stems from sex differences regarding the nature of human fertilization that we know for certain existed in the EEA (i.e. fertilisation takes place internally and this has always been the case for our species).

More problematic is the methodology. The methodology can be used to generate all sorts of hypotheses about observed behaviour in the present that involves adaptations selected for in the EEA. The EEA is sufficiently elastic to allow for just about any observed universal behaviour to have an adaptation hypothesized for it. It is even conceivable that one could reason from the EEA to explain or predict both behaviour B and its opposite $\neg B$.

That the method allows arbitrary narratives to be generated requires that any adaptationist hypothesis be scrutinised to determine whether it is more than a just-so story¹⁰. In order for an adaptationist hypothesis to be more than a just-so story, it needs to satisfy a number of requirements. First, it must be internally consistent and consistent with known science. Second, where possible, it must be able to make precise predictions. As Popper would recommend, the more novel and unexpected these predictions the better. Third, those predictions must be successfully confirmed.

We know that the first two criteria have been satisfied: the adaptationist account appears to be internally consistent and consistent with evolutionary theory. It has also made precise predictions about sex differences in jealousy. Have these predicted differences been found? The adaptationist hypothesis initiated research into whether sex

¹⁰ This is based on Gould's criticism of Sociobiology being panadaptationist.

differences exist in jealousy, and generally these have been confirmed.¹¹ In Buss et al. (1992) participants completed questionnaires asking whether imagining their partner 'trying different sexual positions' or 'falling in love' with another person was the most distressing. 60% of men reported sexual infidelity the most upsetting. In contrast, 83% of women reported emotional infidelity the most upsetting. This does offer support for the hypothesis: presented with the forced choice between imagining both types of infidelity, more men reported sexual infidelity the most upsetting, and more women emotional infidelity the most upsetting. The results have been replicated in other countries and also by using other measurements. A second experiment recorded by Buss et al. (1992), this time measuring physiological responses rather than using self-reporting, also confirmed sex differences. In cross-cultural experiments, Buunk et al. (1996) applied a similar forced choice dilemma to participants in USA, Germany, and the Netherlands. The Netherlands in particular was selected as Buunk et al. considered it to have fairly relaxed views on infidelity relative to the USA and Germany. Sex differences were confirmed in all three countries, being large in USA and medium in Germany and the Netherlands. In addition to the USA and Europe, sex differences in jealousy have also been confirmed in China (Geary et al. 1995), Japan and South Korea (Buss et al. 1999). Schützwohl and Koch (2004) found that men were better able to remember signs of sexual infidelity and that women were better able to remember signs of emotional infidelity.

5. Evaluating explanatory power

The above is strong evidence that sex differences exist in jealousy; it seems that the predictions have been confirmed. At the very least, we can conclude that the hypothesis is not merely a just-so story. But how

¹¹ An exception is homosexual men. A majority of homosexual men report that emotional infidelity is more upsetting to them than sexual infidelity. An explanation sometimes offered is that homosexuals are "cross-gendered" psychologically (Buller 2005, p. 320). More generally, the interpretation of the research findings can be challenged (see, for example, Buller 2005). Unfortunately, the word limit prevents me from offering more than a summary of the research undertaken in this area.

powerful is this hypothesis? The explanatory power of this hypothesis faces problems.

Hempel (1959) showed that functional explanations face the problem that there can be several ways to accomplish the same task or fulfil the same need. In our context, for any trait T which is evoked to solve an adaptive problem, there is another trait T' such that it could also solve the adaptive problem. For example, some animals use lungs to breathe, others use gills. We could conceive others traits that could also solve the problem of infidelity. Take, for example, the trait of possessiveness, the characteristic of desiring to own and control things. Instead of attacking rivals or threatening his partner, a man could exert possession over his partner by restricting her freedom of movement, confining her to the dwelling and so on. In turn, women could develop and exert emotional possession, manipulating their partner emotionally to ensure resources go to them as maximally as possible regardless of their partner's sexual fidelity. So it is conceivable that another trait could have been selected by natural selection to solve the problem of infidelity. The explanatory power of the hypothesis that jealousy is an adaptation will be somewhat limited if it cannot explain why other functionally equivalent traits weren't selected for instead.

We could claim that jealousy is the best trait to solve the adaptive problem of infidelity and that this is the reason why it was selected for instead of other traits. However, selected adaptations aren't always the optimal solution: sometimes, different designs can solve the adaptive problem more successively. There is a reply to this. The reason why optimal designs are not always selected for is because there are constraints on what can be selected. For example, what designs are available to natural selection is dependent on the degree of variation in the gene pool of the population. Another constraint is that any design selected must be compatible with existing mechanisms. Knowing what the constraints in the EEA were could explain why the trait of jealousy was selected for instead of other designs. In such conditions, jealousy might have been the only possible solution – a necessary one – or the best solution available given the constraints and conditions. However, the problem raised in the previous section - that many features of the EEA are unknown – comes into force here. Although this problem

doesn't challenge the rationale for why there was an adaptive problem of infidelity and why the problem was different for the sexes (because the rationale depends on one of the few things we do know for certain about the EEA, i.e. that internal fertilisation took place), the EEA being largely unknown does present challenges when trying to explain why jealousy was selected for instead of other traits. How much variety was in gene pool? What existing mechanisms did our ancestors already have in the Pleistocene epoch? What were the costs of each design relative to the environment? We cannot adequately explain why jealousy was selected for instead of other traits until we have answers to these and other questions. At present, we don't know the EEA well enough to be able to provide adequate answers to these questions; indeed, many features of the EEA might prove to be unknowable.

Another related problem is that different mechanisms could bring about jealousy. The evidence summarised in the previous section is broadly consistent with the existence of a jealousy mechanism with sex differences in its design features. However, the evidence is also consistent with other functionally equivalent mechanisms. One could, like Buller, posit 'an evolved emotional alarm specific to threats to relationships in which one has invested one's reproductive effort', but account for sex differences as being 'due to differences in what the sexes learn about one another' rather than due to innate design differences between the sexes (2005, p. 333). Alternatively, one could posit a mind structured by a few domain-general modules carrying out general computations. Here, jealousy would be a by-product of these general computation modules (jealousy as a behavioural strategy given certain conditions and circumstances) and sex differences in jealousy could be accounted for as different values and beliefs about gender roles being computed. The evidence discussed in the previous section is consistent with all these hypotheses – the evidence does not entail a domain-specific module with innate sex differences in design.

One could respond by saying that the evolutionary psychologists' hypothesis, unlike the other two hypotheses just mentioned, brought about specific predictions which have been confirmed, thereby expanding our field of knowledge. Surely this must make it more likely to be the actual cause of the explanandum? That the predictions of sex

differences in jealousy have been broadly confirmed certainly proves that the hypothesis enjoys heuristic value¹². However, this doesn't entail that the hypothesis of a jealousy mechanism with innate sex defences enjoys greater explanatory power over the other two hypothesized mechanisms. For Hempel, functional explanations have heuristic value but little or no explanatory value; they help in the discovery of new phenomena but fail to adequately explain them. In addition, different hypotheses - such as the one explaining sex differences in jealousy as being a by-product of a few domain-general modules computing different perceived values and gender-roles - can also yield predictions of sex differences in jealousy.

Cognitive neuroscience is a new science, and ultimately we need that science to mature before we can really judge the explanatory power of the hypothesis that jealousy is a specialised mechanism with sex differences in design. After all, if the massive modularity thesis of the mind turns out to be false, then the explanatory power of the hypothesis will be nil (as it would be inconsistent with known facts). It is not sufficient to simply observe a correlation between actual sex differences in jealousy and the hypothesised jealousy mechanism with its innate differences in design. At the very least, we need to know whether the mind is massively modular, let alone whether one of the modules is a jealousy mechanism, and neuroscience has yet to tell us whether this is the case.

Conclusion

The hypothesis that jealousy is an evolved psychological mechanism designed to solve the adaptive problem of infidelity is more than a mere just-so story. It has expanded the explanandum of jealousy to include sex differences. As such, it enjoys considerable heuristic value and predictive strength. However, its explanatory power is limited in two ways. First, an adequate explanation of why jealousy was selected for instead of other functionally equivalent traits is unlikely due to much of the EEA being unknown. Second, the hypothesis depends on the truth of the massive modularity thesis of the mind, a thesis that we cannot

¹² That is, valuable as a guide for research.

judge to be true or false until cognitive neuroscience matures. The available evidence appears to be broadly consistent with the hypothesis but does not entail it.

Bibliography

Allen, C., 'Teleological Notions in Biology', *The Stanford Encyclopedia of Philosophy (Summer 2004 Edition)*, Edward N. Zalta (ed.), URL=<<http://plato.stanford.edu/archives/sum2004/entries/teleology-biology/>>

Barkow, J. H., Cosmides L. & Tooby J. (1992), *The Adapted Mind*. New York: Oxford University Press

Buller, D. J. (2005), *Adapting Minds*. Cambridge, Mass: MIT Press

Buunk, B. P., Angleitner, A., Oubaid, V., & Buss, D. M. (1996), 'Sex differences in jealousy in evolutionary and cultural perspective'. *Psychological Science* 7: 359-363

Buss, D. M. (1994), *The Evolution of Desire*. New York: Basic Books

— and Larsen, R. J., Westen, D., & Semmelroth, J. (1992), 'Sex differences in jealousy: Evolution, physiology, and psychology' *Psychological Science* 3: 251-255

— and Larsen, R. J., & Westen, D. (1996), 'Sex differences in jealousy: Not gone, not forgotten, and not explained by alternative hypotheses', *Psychological Science* 7: 373-375

— and Haselton, M. G., Shackelford, T.K., Bleske, A. L. & Wakefield, J. C. (1998), *Adaptations, Exaptations, and Spandrels*. *American Psychologist* 53: 533-548

— and Shackelford, T. K., Kirkpatrick, L. A., Choe, J., Hasegawa, M., Hasegawa, T., & Bennett, K. (1999), 'Jealousy and the nature of beliefs about infidelity'. *Personal Relationships* 6: 125-150

— and Haselton, M. (2005), 'The evolution of jealousy', *Trends in Cognitive Sciences* 9: 506-507

Daly, M., Wilson M., & Weghorst, S. J. (1982), 'Male sexual jealousy'. *Ethology and Sociobiology* 3: 11-27

Geary, D. C., Rumsey, M., Bow-Thomas, C. C., & Hoard, M. (1995), 'Sexual jealousy as a facultative trait', *Ethology and Sociobiology* 16: 355-384

Hempel, C. G. (1959), 'The Logic of Functional Analysis', in Martin, M. and McIntyre L. C. (1994), *Readings in the Philosophy of Social Science*. Cambridge, Mass: MIT Press

Schützwohl, A. (2005), 'Sex differences in jealousy', *Evolution and Human Behavior* 26: 288–299

— and Koch, S. (2004), 'Sex differences in jealousy'. *Evolution and Human Behavior* 25: 249–257

Symons, D. (1979), *The Evolution of Human Sexuality*. Oxford: Oxford University Press

Wilson, E. (1975), *Sociobiology: The New Synthesis*. Cambridge, Mass: Belknap Press

Wright, L. (1973), 'Functions', *Philosophical Review* 82: 139-68

Kant on things-in-themselves: one world or two?

Winner of the 2006 BUPC Prize for Philosophical Communication

Craig French

Heythrop College, University of London

craig.french@bups.org

In the *Critique of Pure Reason* (henceforth the *Critique*) Kant makes a distinction between things as they appear and things as they are in themselves. This paper is largely concerned with the philosophical significance of this distinction, rather than the correct interpretation of Kant. With that in mind I follow, without going into heavy exegetical detail, a sympathetic ‘one-world’ interpretation of Kant’s distinction offered by Rae Langton (1998). In section (I) the key interpretive points of Langton’s understanding of Kant will be outlined. In section (II) it will be argued that a proper appreciation of the consequences of Kant’s claims (as interpreted by Langton) shows that, *pace* Langton, Kant is committed to the *possibility* of ‘two-worlds’; i.e. to the possibility of there being two-types of non-overlapping objects: one knowable (spatial) type and one unknowable (non-spatial) type. In conclusion I argue that the result of this possibility is, alas, epistemological scepticism.

For Kant, an object-as-appearance is an object of experience.¹ Appearances for Kant are *empirically real* – not mere ‘seemings’ or ideas. As for the thing-in-itself, Kant says ‘if the senses represent to us something merely *as it appears*, this something must also in itself be a thing...’ (A249). Kant urges that an object-in-itself should be understood as the *ground* of that which appears because if there were no such objects-in-themselves we ‘should be landed in the absurd conclusion that there can be appearance without anything that appears’ (Bxxvii). The existence of things-in-themselves, then, is a necessary

¹ Kant often uses ‘appearance’ to mean ‘phenomenon’; that is, ‘sensible entity’ (B306).

condition of experience which confers objectivity on the objects of experience (appearances).² In this sense, the thing-in-itself is necessary for one of Kant's aims, that of avoiding Berkeleyian subjective idealism.³

Langton (1998) offers a plausible reading of Kant's distinction between appearances and things-in-themselves ('the distinction' henceforth).⁴ Two of Langton's key interpretative points are as follows:

1. Kant's distinction does not admit of a 'two-world' (two-object) interpretation. It is a 'one-world' distinction, i.e. pertaining to *a* thing which can be *considered* as it is in itself (noumenon), and the *same* thing which can be considered as appearance (phenomenon). Such passages as 'this object as *appearance* is to be distinguished from itself as object *in itself*' (B69) support this reading.⁵
2. The distinction, nonetheless, is a *metaphysical* distinction between two non-overlapping classes of properties: the intrinsic and extrinsic properties of things.

So when Kant distinguishes a thing-in-itself from appearance, he is merely distinguishing the intrinsic properties of *a* thing from the extrinsic properties of *that* thing. As Langton says:

² Here, in taking Kant to be concerned with the 'conditions of the possibility of experience', I follow the 'analytic' interpretation of Kant; Strawson (1966), disputed by Gardner (1999).

³ Ellis (2005), p. 58, notes how Kant's understanding of Berkeley fails to take account of Berkeley's role for God, which can be understood along the lines that Kant intends for his 'thing-in-itself' in so far as the thing-in-itself for Kant is the *ground* of appearance.

⁴ This interpretation is different to that of Allison (2004), who also endorses a 'one-world' reading of Kant, but takes Kant to be making a *meta-epistemological* distinction, not a metaphysical distinction as Langton claims. Langton's reading is also different to that advocated by Bennett (1966, 1974) who reads Kant as a kind of phenomenalist, where objects are logical constructs created out of actual or possible sensory states. Van Cleve (1999) also advances a kind of phenomenalist reading of Kant, where objects of appearance are said to be 'virtual objects'. McDowell (1994) offers a 'two-world' view of Kant, but offers a differing view in his (1998).

⁵ Although see Van Cleve (1999) Chapter 10, for reservations.

*An object in itself is a thing that exists independently of its relations to other things [it is] a substance, which has intrinsic properties. A phenomenon is an object in relation to something else. The same object can be described both as phenomenon and as object in itself, precisely because the same object that has relations to other things also has an 'intrinsic nature' ... [Phenomena] are relational properties of substance s...*⁶

Objects need some intrinsic-nature – they must have some intrinsic properties. To say that a thing-in-itself is a thing *qua* thing-with-intrinsic-properties is just to say that a thing-in-itself is an object, the existence of which is *independent* of other things (A284/B340). As Langton says, 'the intrinsic properties are those which do not imply coexistence with any other thing.'⁷ So the existence of a thing-in-itself is an existence compatible with loneliness.⁸ Intrinsic properties do not depend for their existence on any other properties or objects, whereas extrinsic properties do.

Kant is keen to emphasise how we have no knowledge of things as they are in themselves (Bxx.). It follows from (1) and (2) above that Kant's insistence that we can't know things-in-themselves amounts to the claim that we can't know the intrinsic properties of things. For Kant, knowledge is receptive – it requires that something is *given* to what he calls 'sensibility'.⁹ So Kant's crucial claim, which Langton calls 'Kantian Humility', is that for any object O, sensibility is *receptive* to its extrinsic properties, but *not* receptive to its intrinsic properties. Knowledge only comes from extrinsic-properties/relations, not intrinsic properties (A277/B333). Kant says that it is 'illegitimate and unreasonable' to expect knowledge of the intrinsic nature of things, '[f]or what is demanded is that we should be able to know things, and therefore to

⁶ Langton (1998), pp. 19-20.

⁷ Ibid, p. 18.

⁸ In formulating this understanding of intrinsic properties Langton borrows from David Lewis and Jaegwon Kim, see Langton, (1998), pp. 17-18, footnotes. 3 & 6. David Lewis however admits that some dispositional properties are intrinsic properties. No such admission could be made by Kant though. For Kant, if Langton's reading is correct, all dispositional properties are extrinsic/relational.

⁹ See the 'Transcendental Aesthetic' of the *Critique* for Kant's epistemology.

intuit them, without senses...’ (ibid). What is unreasonable according to Kant is the wish that our knowledge were not receptive.¹⁰ Langton takes Kant’s epistemic humility as signifying that there are inevitable constraints on that with which we can become acquainted, as such it is *neither* idealism *nor* scepticism.¹¹

That intrinsic-properties (and hence things-in-themselves) cannot be given to the senses, and hence that we can have no knowledge of them, is the reason why we cannot give a single example of an intrinsic-property. This is a departure from traditional conceptions of intrinsic properties where spatial-properties, shape-properties or even certain dispositions were thought to be intrinsic. For Kant, these are all extrinsic/relational properties. Kant’s claim regarding intrinsic properties is radical; we must give up any attempt to state what the intrinsic properties of things might be.

In summary, the following theses can be attributed to Kant:

- a) The distinction is a distinction between the intrinsic and the extrinsic properties of a *single* thing
- b) Knowledge is receptive, and intrinsic properties are not given to receptivity, therefore we can have no knowledge of things-in-themselves (humility).
- c) We can therefore give no example of an intrinsic-property.
- d) We can know that intrinsic properties exist because they are the necessary grounds of extrinsic properties (i.e. appearances).

The further claim is:

- e) Knowing *that* there are intrinsic properties is consistent with not knowing *what* those intrinsic properties are.

¹⁰ Langton (1998), p. 43.

¹¹ Ibid, p. 2.

As we have seen, Kant should be understood as positing two types of *properties* – intrinsic unknowable properties, and extrinsic knowable properties. Hence, there are no unknowable *objects* (as the ‘two-world’ view has it), just unknowable *features* of objects. In this section I intend to argue, through two dilemmas, that if we stick to this understanding of Kant, granting epistemic humility, we get the following results:

1. Along with Kant’s distinction between knowable and unknowable *properties* comes the *possibility* of a realm of unknowable *objects*
2. Hence Kant’s one-world/one-object view entails the *possibility* of the same ontological commitments as the two-world/two-object view.
3. The crucial claims with which I shall conclude are that the mere possibility of the division of reality into a realm of unknowable and knowable objects is (a) a form of epistemological scepticism, which undermines what we think we know about supposedly *knowable* objects and (b) a possibility we have *no reason* to accept is possible.

The following question arises:

For any object O is it possible for O to exist without having any extrinsic properties? In other words, can there be an object that has *only* intrinsic properties?

This question forces a dilemma because the two available answers – ‘yes’ and ‘no’ – are both unattractive.

Horn 1: If the answer is ‘no’, then it seems that *having some extrinsic property* is an *intrinsic* feature of O, since O’s existence depends on it

having *some* extrinsic property or other. But by expressing this intrinsic feature *Kantian Humility is violated*. So Kant cannot answer 'no'.¹²

Horn 2: If the answer is 'yes', then the result is that there could be in existence *objects* (not just properties) that we have no access to at all. This follows from Kantian Humility. If we can only know objects through their extrinsic properties, then any object that has no extrinsic properties will be unknown. This second answer – the one that Kant must accept – reveals that with Kant's distinction between appearance and unknowable-thing comes the *possibility* of there being two 'realms' of objects. That is, the possibility of there being two types of objects: objects-with-intrinsic-*and*-extrinsic-properties constituting one type of object, and objects-with-*just*-intrinsic-properties constituting the other.

Thus far, the second horn of the dilemma reveals that there is the possibility of a distinction between *knowable* objects and *unknown* objects. This, of course, should not worry us, if an object from the realm of *unknown* objects is nonetheless a *knowable* object. If it is possible for one of those objects that has only intrinsic-properties (and therefore is *unknown*) to (somehow) have or acquire an *extrinsic* property, then that object will become *known*. If this is a genuine possibility, then the set of unknown objects is included in the set of

¹² It might be objected that *having some extrinsic feature* is itself an *extrinsic* property of O, since the feature of *having some extrinsic feature* implies the existence of other objects/properties. But isn't this a confusion? Surely what implies the existence of other objects or properties are *some* extrinsic features, not the feature of *having some extrinsic property*. But this is a matter of debate. Perhaps this point can be pushed: to answer 'no' to the question is to say that any object *must* have *some* extrinsic property; hence the feature of *having some extrinsic property* is a necessary feature of any object. If one were to contend that this necessary feature is an *extrinsic* feature (a feature that implies the existence of other things) then we would have to ascribe to Kant the view that *no object* can exist alone. But this seems to fly in the face of Langton's insistence that objects-in-themselves are objects the existence of which is compatible with loneliness. Surely all this means is that we can consider (and it is possible for there to be) objects that *don't have any extrinsic properties*. So it seems that answering 'no' to the question is unavailable to Kant. To put it another way, how can we make sense of Langton's definition of things-in-themselves if it is a necessary feature of every object that it cannot exist alone? Furthermore, as Langton says elsewhere (Langton & Lewis (1998), p. 340), any necessary property is *intrinsic*.

knowable objects (but not vice versa), and as such, there is no 'two-world' commitment. So if the second horn is to be a genuine horn, it must entail that the objects in the set of *unknown* objects are *unknowable*.¹³ To establish this, I aim to show how Kant is caught by a further dilemma which entails that the objects in the set of unknown objects are unknowable.

The following question arises:

Can any object O from the 'realm' of *unknown* objects possibly have an extrinsic feature and thereby be a known object?

This question forces a dilemma because the two available answers 'yes' and 'no' are both unattractive.

Horn 1: To answer 'no' is to impale oneself on the horn of scepticism: it is to say that for any O from the realm of unknown objects, it is not the case that O could possibly have an extrinsic property and thereby be a known object. The sceptical thesis follows: it is possible that there is a 'realm' of *unknowable* objects.

Horn 2a: To answer 'yes' to the question is to say that *it is* a feature of any unknown object O that it *could possibly be known*. Call this modal feature F. If F is an *intrinsic-feature*, then *Kantian Humility is violated*. Hence, the answer 'yes' in this sense is not available to Kant. Alternatively, there is another sense in which one can answer 'yes', but this is just as problematic.

¹³ Regardless of whether we want to talk about two-types or one-type of object, the damaging point is that we end up with two 'realms of facts': facts about objects-*qua*-only-intrinsic-properties, and facts about objects-*qua*-extrinsic-and-intrinsic-properties. McDowell (1998) says something similar when complaining about Allison's 'one-world' interpretation of Kant. He says 'The picture still involves two realms of fact, one knowable by us and one unknowable by us; it does not undermine the damage this does to say that the same objects figure in both' p. 469, fn. 23. I intend to show how with Langton's reading comes the even more damaging *possibility* of two-realms of *objects*.

Horn 2b: To answer ‘yes’ to the question is to say that *it is* a feature of any unknown object O that it *could possibly be known*. Call this modal feature F. If F is an *extrinsic-feature* then since we know that objects can exist without extrinsic features (as the first dilemma shows) we know that it is possible for any object O to lack feature F. But to say that it is possible that there are objects that lack F is just to say that it is possible that there are objects such that they *couldn't possibly be known* – since F is the feature that makes it possible for objects to be known. In other words, it is to admit the possibility of unknowable objects.

The argument of the two dilemmas is as follows:

1. Assuming Kantian Humility (for *reductio*). Therefore,
2. *The First Dilemma*: Either (1) we violate humility or (2) countenance the possibility of a realm of *unknown* objects.
3. We countenance the possibility of a realm of *unknown* objects. Therefore,
4. *The Second Dilemma*: Either (1)/(2b) we countenance the possibility of a realm of *unknowable* objects or (2a) we violate humility.

At this point two conclusions are available. Firstly, the conclusion that logically follows from Kant’s premises is the sceptical one

5. We countenance the possibility of a realm of *unknowable* objects.

Alternatively is the conclusion that treats the above argument as an (informal) *reductio ad absurdum* of Kantian Humility (since it entails ‘absurd’ scepticism), in which case

6. We violate (reject) Kantian Humility.

The purpose of the two dilemmas is to illustrate how, from the very premises of Kant’s ‘one-world’ philosophy, comes the possibility of the

ontological and epistemological commitments of a 'two-world' philosophy, and that this entails epistemological scepticism: Kant is committed to the possibility of a *necessarily unsurpassable gap* between the conceptual realm of experience and the realm of things-in-themselves (the possibility of a realm of *unknowable* objects).¹⁴ This is a possibility not envisaged by Langton, and a possibility which we have no reason to believe is possible.

But is it enough to note this counter-intuitive possibility of an unknowable realm? Perhaps the possibility *isn't* a problem for Kant; after all, what is revealed in experience is the 'physical world' and is 'empirically real'. So, it is argued, the possibility isn't really a sceptical possibility since the possibility of an unknowable realm does not undermine empirical knowledge in the sense that the possibility of an evil demon undermines empirical knowledge for the Cartesian sceptic. The problem with this is that if we respect Kantian Humility, the proper conclusion is that *for all we know* the unknowable realm *might* undermine empirical knowledge. Who knows? The counter-intuitive nature of the possibility suddenly seems more striking.

It should be remembered that the (possible) objects of the unknowable realm are *non-spatial*, since for Kant spatial-properties are extrinsic. It is not clear how one could conceive of unknowable/non-spatial objects. Perhaps the only option is to conceive of such objects in a purely transcendental fashion, that is, as objects that are necessary conditions of experience and knowledge. So one might conceive of them, somehow, by the function that they have. After all, Kant introduced the thing-in-itself in order to *underpin* the realm of experiences in this sense; i.e. in order to underpin extrinsic-properties. What we must

¹⁴ The idea that Kant separates concepts from reality, two concepts which belong together, comes from Ellis (2005). Ellis situates Kant as being in the grip of 'The Syndrome' (the defining framework of which is that 'things in themselves are to be dualistically opposed to the realm revealed in our concepts, the conceptual realm comprising the things which are available to experience' p. 1). It should be noted that although Ellis concludes negatively against Kant, in her chapter on Kant, she does examine those passages in which a positive interpretation of Kant can be arrived at – one where he isn't caught in the grip of the syndrome. However, there are tensions in Kant's own writings which show him wavering between positions, and as such, the negative conclusion follows.

realise, however, is that the possible realm of *unknowable* objects considered here is a realm of objects that couldn't possibly have extrinsic properties, and therefore couldn't possibly have the transcendental function that makes conceiving of such objects vaguely plausible.¹⁵

It seems that the most fruitful option, for those that want to maintain things-in-themselves on transcendental grounds, is to drop Kantian Humility. As such, things-with-intrinsic-properties can be said to be the ground of *those same things-with-extrinsic-properties*, without us having to countenance the possibility of a further mysterious unknowable/non-spatial realm, because we have imposed no restrictions on knowing things-with-intrinsic-properties. A further virtue of this is that we could allow intrinsic properties to be spatial, and hence not be stuck with the strange idea that non-spatial properties support spatial extrinsic properties.

Not only is it an unargued premise of Kant's that we can't know things-in-themselves because our senses are not receptive to them (Strawson (1966), p. 250),¹⁶ but also, 'humility' gives rise to the counter-intuitive possibilities outlined above. Therefore, Kantian Humility should be rejected.¹⁷

¹⁵ I leave the conclusion here open. Perhaps one could argue that non-spatial, unknowable objects are not conceivable, and therefore not possible. Or perhaps one could argue that unknowable, non-spatial objects are not conceivable. Yet they must be if the consequences of Kant's philosophy are followed through, and therefore we must reject something from Kant's philosophy; namely the something that makes Kant think that non-spatial, unknowable objects are conceivable.

¹⁶ Langton (1998) agrees, but claims to find, in Kant's pre-critical work, an alternative argument for humility. Langton's argument is criticised both philosophically and exegetically by Falkenstein (2001), and Langton (2001) has a reply to Falkenstein.

¹⁷ I am grateful to Akosua Bonsu, Fiona Ellis, and George Reynolds, who all at some point have helped me enormously through discussions and comments on drafts of this paper. I am also grateful to the blind referee for comments on a first draft.

Bibliography

- Allison, H.E. (2004). *Kant's Transcendental Idealism An Interpretation and Defense* (2nd edn). New Haven: Yale University Press
- Bennett, J. (1966). *Kant's Analytic*. Cambridge: CUP
- (1974). *Kant's Dialectic*. CUP
- Ellis, F. (2005). *Concepts and Reality in the History of Philosophy: Tracing an Error from Locke to Bradley*. London: Routledge
- Falkenstein, Lorne. (2001). 'Langton on Things in Themselves: Critique of Kantian Humility'. *Kantian Review* vol. 5
- Gardner, S. (1999) *Kant and the Critique of Pure Reason*. London: Routledge
- Kant, I. (1781/1787). *Critique of Pure Reason*. [Trans. Norman Kemp Smith, 1929, Macmillan]
- Langton, R. & Lewis, D. (1998). 'Defining 'Intrinsic'', *Philosophy and Phenomenological Research*, 58
- Langton, R. (1998). *Kantian Humility Our Ignorance of Things in Themselves*. Oxford: Oxford University Press
- (2001). 'Reply to Lorne Falkenstein'. *Kantian Review* vol. 5
- McDowell, J. (1994). *Mind and World*. Cambridge, Mass.: Harvard University Press
- (1998). 'Having the World in View: Sellars, Kant, and Intentionality', *Woodbridge Lectures, Journal of Philosophy*, vol. XCV, no. 9
- Strawson, P.F. (1966). *The Bounds of Sense An Essay on Kant's Critique of Pure Reason*. London: Routledge
- Van Cleve, J. (1999). *Problems from Kant*. Oxford: Oxford University Press.

Hegel's master and slave

Ryan Dawson

Selwyn College, Cambridge

ryan.dawson@ntlworld.com

The master-slave discussion is found in the section of the *PbG* titled 'The Truth of Self-Certainty'. As this section begins, we see Consciousness fight through conceiving of itself within an abstract realm to a conception of itself within time. It is at this stage of Consciousness's journey, the stage at which it is thrown into the concrete world, that the Master-Slave Dialectic takes place. In entering this world Consciousness first finds a sense of itself as distinct from what it perceives – it becomes Self-Consciousness.

Every stage of Consciousness's journey involves embracing and overcoming contradictions. At this stage the central tension that Consciousness must deal with is that between a conception of itself as a spiritual entity and a conception of itself as located in the natural world, or 'Life'. Consciousness must find unity with Life. As Hegel says, 'Life points to something other than itself, viz. to consciousness, for which Life exists as this unity, or as genus.'¹ The transition to Self-Consciousness is one out of 'the colourful show of the sensuous here-and-now and the nightlike void of the supersensible beyond' and into 'the spiritual daylight of the present.'² It is Desire that brings Consciousness into this spiritual daylight. Consciousness Desires objects and in doing so it must see them as external and independent.

'Self-Consciousness exists in and for itself when, and by the fact that, it so exists for another; that is, it exists only in being acknowledged.' The need for Recognition (acknowledgement) drives us both to develop personalities and to understand one another. Paragraph §180 gives a profound analysis of how relationships work. We project ourselves onto

¹ *Hegel's Phenomenology of Spirit*, Miller (trans.) (Oxford University Press, 1977), §172.

² *Ibid.*, §177.

others in order to understand them. We also redefine ourselves so as to be distinct from others. Thus we overcome the other by overcoming ourself in a never-ending process. Only in moments of self-overcoming can we see the other as distinct and let them 'go free'. At these same moments we attain a Self free of Otherness. 'Each is for the other the middle term, through which each mediates itself with itself and unites with itself.' This process of redefinition is of no value unless the other does likewise.

But Self-Consciousness does not start with well-functioning relationships. It begins by revelling in the denial of everything external. 'Each is indeed certain of its own self, but not of the other, and therefore its own self-certainty still has no truth.' Self-Consciousness can find truth for its self-certainty only through Recognition.

Upon encountering another Self-Consciousness, Self-Consciousness undertakes to show that it is not attached to Life in order to prove itself as Self-Consciousness. It does so by showing that it is not attached to life. It risks its own life to show it is not part of Nature. It does so by trying to kill the other Self-Consciousness. 'They must engage in this struggle, for they must raise their certainty of being for themselves to truth, both in the case of the other and in their own case.'

'This trial by death, however, does away with the truth which was supposed to issue from it.' In killing the other, the victor loses the middle term. He gains no Recognition. From this experience, Self-Consciousness learns that it needs life. This gives rise to a new form of Consciousness. This is Consciousness that is both for-itself and for another – the Master and Slave.

The Master is primarily for-itself. For him, the Slave is as much thing as man. The Slave gave up his independence in the struggle so that he could retain his life. The Master uses the Slave so that he no longer needs to deal with objects. He simply consumes objects, leaving the slave to work on them for him. The Master is able, through the Slave, to forget the independence of objects. He deals with them only in their dependent aspect.

This arrangement, however, is not satisfactory. The Master is unsatisfied because he wants to dominate an independent Consciousness. But the Slave is only a dependent Consciousness. The relationship is unequal and is deficient in this respect.

Consciousness moves to the next level through the Slave. He has experienced a fear of death – of becoming a thing. He rids himself of this through learning to control objects and thus become independent of them. Hence ‘the fear of the Lord is indeed the beginning of wisdom.’ For the Slave, work becomes the middle term. He is able to recognise the independence of the object and thus attains a more permanent consciousness. The Slave comes to identify with his work. As such, he becomes conscious of himself.

Once we conceive of objects as independent we can think of them in their own right. We rise above the ‘picture-thoughts’ of sense-certainty. Thinking is a new freedom. ‘In thinking, I am free, because I am not in an other but remain simply and solely in communion with myself.’ This freedom is celebrated by Stoicism – the next form of Self-Consciousness.

In this essay I shall argue that Kojève distorts Hegel’s dialectic of the Master and Slave. I shall not be interested in whether this distortion is deliberate. My interest is in how the distortion comes about and what it reveals about Hegel’s text. My argument will proceed in several broad stages. First, Kojève’s subject negates the world, whereas Hegel’s subject is unified with the world. Second, Kojève views the subject as developing from within, whilst Hegel’s develops with and in response to the world. Third, Kojève’s subject develops according to the narrow logic of desire. The development of Hegel’s subject is according to a multiply suggestive logic. Fourth, Kojève’s subject is located in historical time. Hegel’s is located in fictive time.

From this, I argue that Kojève gives the Master-Slave dialectic the character of a thought-experiment. He presents it as though the master and slave could be placed together and we could use logic to deduce what will ensue. This, in itself, is true to Hegel. However, Hegel’s discussion is better seen as a thought-experiment that appeals more as a

*parable*³. The great appeal of the Master-Slave is in its suggestions and unification of many disparate areas of life. As Butler says, there is a mismatch between Hegel's project (heavily logical) and his methods (surprisingly literary).⁴ I take this conclusion to be evidence for Taylor's view that the historical dialectics of the *Phenomenology* purport to be logical but in fact convince only by their ability to bring together an interpretation.⁵

Subject's negation of the world

For Hegel, human action is the embodiment and expression of a larger metaphysical order. As Butler points out, Kojève reverses the order of significance of this.⁶ He views human action as constituting and creating metaphysical concepts. As he says,

*to say that there is Totality, or Mediation, or dialectical Overcoming, is to say that in addition to given-Being, there is also creative Action which ends in a Product.*⁷

These concepts are left to Kojève's subject to create because they are not to be otherwise found in the world. The world, for Kojève, is 'a natural World that has no beyond.'⁸ Kojève's subject is faced with an inhuman natural world. Its response is to carve out a human (historical) world in opposition to the natural world. This is why desire is directed toward the Other. The subject is bound with the objects of his desire – he takes on the same status as those things he desires. Hence only by directing

³ Seeing the Master-Slave as a parable is something I take from Stern in *Hegel and the Phenomenology of Spirit* (London: Routledge, 2002).

⁴ In *Subjects of Desire* (New York: Columbia University Press, 1987) Butler argues that Hegel's sentences defy their literal meanings, p. 18, and that the transitions of the *Phenomenology* purport to be logical but are really narrative devices, p. 21.

⁵ Historical dialectics concern purposes and their accomplishment in reality. Ontological dialectics concern standards and their realisation. Taylor, *Hegel* (Cambridge University Press, 1975) p. 218, 4th paragraph.

⁶ *Subjects of Desire*, p. 65.

⁷ Kojève, *Introduction to the Reading of Hegel*, Nichols (trans.), Bloom (ed.) (New York: Basic Books, 1969), p. 259.

⁸ *Ibid.*, p. 258.

his desire toward another desire can he separate himself from the natural world.

In Kojève's analysis of the Master and Slave, the inhuman status of the natural world is reflected in the positive emphasis on negation. Kojève's Slave triumphs by negating the world. The Slave becomes Master of Nature. The Slave's work 'creates a real objective World, which is a non-natural World, cultural, historical, human World'⁹. This newly created world is the escape from the natural world, which represents the threat of death.¹⁰

I do not think that any of these claims is wrong as an interpretation of Hegel. They distort only insofar as they are presented as constituting a final, all-embracing interpretation. The Hegelian dialectic of self-consciousness progresses after the Master and Slave, but Kojève ends the dialectic of self-consciousness with the Slave's triumph. It is the 'slavish Consciousness that in the end realizes and reveals the ideal of autonomous Self-Consciousness and is thus its "truth"'. But the Hegelian dialectic continues, eventually to end with *Geist*. *Geist* is aware of and comfortable with its physical embodiment – its presence in the natural world. This transition is achieved by self-consciousness surrendering its particular will to a universal will. Man's definitive satisfaction lies not in recognition of particularity, as Kojève says it does¹¹, but in recognition of unity.

Kojève emphasises the negating side of the Slave's triumph. Hegel's Slave comes to recognise the reality of the world through work. But Kojève's Slave comes to construct an alternative reality. The natural world remains negated for Kojève's Slave. Hegel's Slave does not achieve satisfaction in this way. Hegel's Slave must come to acknowledge the reality of the world in order to gain satisfaction by overcoming it.

⁹ *Ibid.*, p. 26.

¹⁰ *Ibid.*, p. 27.

¹¹ *Ibid.*, p. 258.

Subject's development from within

So Kojève's subject is located in a world in which he cannot see himself reflected. Hence he is driven to create his own world within it. So the motor of dialectical development for Kojève is entirely internal. All change comes from within the subject. Hegelian change, however, comes about in response to and in accord with the world. The Hegelian subject is led by forces not entirely internal to itself.

This is a point which Butler emphasises and from which she draws significance.¹² To put it as she does, it is a vital part of the *Phenomenology's* narrative that the subject be 'comic'. The Hegelian subject is never aware of why he is at the place that he is on his journey. Each point is the subject's best efforts to respond to his circumstances. He tries to create a universal perspective that will deal with his situation. Yet each effort is thwarted by something that has escaped his attention. The reader is led to identify with each effort and with each failure. No failure is ultimate because the subject's goals are not fixed. Goals change in response to the world.

Hegel's subject must be comic in order for the *Phenomenology* to reflect a view of reason as an attempt to construct universality out of the particularity of the world. But Kojève's subject fixes his goals for himself. His problem is only how to realise his goals in the world – he does not change his goals in response to the world.

Understanding Kojève's distortion of the powers of the Hegelian subject thus allows us to understand the importance of the narrative structure of the *Phenomenology*. We might normally call this narrative structure a literary device. But we can see that Hegel uses narrative structure for philosophical purposes.¹³

¹² *Subjects of Desire*, p. 76.

¹³ The use of narrative for philosophical purposes is a result of Hegel's phenomenological approach. His philosophical insight has a phenomenological character and hence is not expressed in strict argumentation. Phenomenological writings are necessarily descriptive. Since the subject of Hegel's insight progresses with time, he must use a descriptive narrative.

Narrow logic of development

Kojève's subject is located within a natural context. His is much like the vision of man from the end of the *Nicomachean Ethics* – man as part of nature but containing a higher element within him.¹⁴ Hence it is important for Kojève that this higher element not be mysterious. He must give an account of its origin that is compatible with the subject's being located in the natural world. Kojève does this through his highly original analysis of the concept of 'desire'. This is worth reviewing briefly.

Desire, for Kojève, is the first stage in the development of personality.¹⁵ It is a progression from passive knowing because passive knowing only yields absorption in the object concerned.¹⁶ Desire is a transforming action. Desire says no to the world that it finds and replaces it with a different world. But animal desires cannot move us to a new level of consciousness. We are as we make the world. In making the world we express ourselves. To just direct desire towards natural objects is to be continuous with nature.

But consciousness needs to go beyond the natural world. So desire must be directed towards the desires of other consciousnesses, as these are the only things that go beyond the natural world. This need means that humans must seek each other out. A man in a state of nature has a subjective certainty of his humanity. This can only be verified by his directing his desire toward the desire of another. Thus desire is recognition.

This account of desire gives the master-slave dialectic a narrower logic than Hegel intends. It does so because desire is seen as developing purely from within the subject. In Hegel, the line between consciousness and the world is blurred and constantly shifting as it slowly dissolves. Ultimately, Hegel wants to unite consciousness and the

¹⁴ *Nicomachean Ethics*, Book X.

¹⁵ Personality is used here to mean the way we differentiate ourselves from one another.

¹⁶ Of course, for Hegel the problem with passive knowing is that it does not fully acknowledge the reality of the object. This again indicates how Kojève regards desire as opposing nature, whereas Hegel regards it as unifying.

world. In Kojève, the line is clear. For example, Hegel presents the development of desire as the next stage of consciousness after the attempts of the understanding to deal with objects in the abstract. Consciousness has failed to find stability in this and is forced to recognise objects in more concrete terms. Desire will enable consciousness to do so. For Kojève, no such problem exists. Desire is simply part of man.

Hegel's master and slave are partly driven by forces outside of themselves. His slave is not to know that work will transform the object and force him to recognise its reality. In Kojève, this recognition of the object has a different character. The object is more real because it has been worked upon and thus reflects humanity. In Hegel, the slave's shaping the object to reflect himself is no doubt important. But the recognition of the reality of the object is also recognition of its *independence*. It furthers the theme from the dialectic of consciousness that objects must slowly be acknowledged as external to consciousness. The Hegelian world is an active player in the dialectic. Kojève's world is passive.

The narrowness of Kojève's logic is revealed by an argument from George Armstrong Kelly.¹⁷ Kelly argues that there are three equally valid ways to interpret the Master-Slave relationship. One is a purely external reading – the Master and Slave are simply people in the world that interact. At the other extreme there is an internal reading – the Master and Slave are not people but representations of forces within a single mind. To use Kelly's terms, we might say that an external reading is largely political and that an internal reading is psychological. There is also a possibility of a mixed reading whereby the Master and Slave are both people and representations of mental forces.

Kelly says that Kojève does address both exterior and interior *consequences* but Kelly accuses Kojève of seeing the Master-Slave relationship as a purely external *confrontation*. In order to see what

¹⁷ Kelly, 'Notes on Hegel's "Lordship and Bondage"' in *Hegel: A Collection of Critical Essays*, MacIntyre (ed.) (Notre Dame: University of Notre Dame Press, 1972).

Kelly is getting at here, we need to understand more of how an internal reading would work.

The essence of the internal reading seems to be that the Master represents 'Master-principle' and the Slave represents 'Slave-principle'. These are different drives within the mind. The theme that is being tapped into is familiar from Plato – external relationships are being used as a metaphor for internal ones. The Platonic influence is particularly important for understanding why this theme should come up in Hegel. It was a popular concern of Hegel's era that the society be restructured via a revolution in the individual's mind. This came about through the interest in Plato of Enlightenment thinkers, inspired by the climate of the French Revolution. By putting Reason in charge of an individual mind, social injustice could be eliminated.

It was very important for Enlightenment thinkers to have a view of the self as so divided. Otherwise social division had to be otherwise explained and the only obvious explanation was that some people were master-like and some were slavish. Social division just reflected actual difference in people. By seeing the self as divided, Enlightenment thinkers could argue that social division came about through the faults of the supposedly master-like people.

Hegel himself was not very motivated by concerns of social justice. But the state-soul metaphor remained a useful one for him to exploit. Hegel wanted to assert the unity of mankind. This is blocked by the differences between individuals. But this can be got around if we argue that the mind of each man works in a certain way and that this working is revealed in the workings of society.

So what is Hegel getting at in the Master-Slave on an internal reading? A Humean reading might have some appeal. The Master is passion and the Slave is the reason that is put to the service of passion. In their resolution, Hegel shows that the mind must attain a state whereby it is not divided. At the end of history there need be no master-principle or slave-principle within the mind.

The difficulty with developing this reading is that there is an equally plausible, yet incompatible reading. The Master's affirming of autonomy from the world and from the Slave seems much more Kantian than Humean. We might also think that the Master represents Kantian reason got out of control – reason trying to affirm itself as independent of the rest of the mind. The failing of the Master-Slave relationship reveals reason to be dependent on passion and vice versa. Thus the unity of the mind is asserted in the opposite direction.

This difficulty, however, does not rule out an internal reading. The Master-principle can be both Humean and Kantian at once. It can be ambiguous between the two. It just means that an internal reading would not have a very specific message. We can see the Master as a kind of amalgam of Humean passion and Kantian reason. His attempt to assert his autonomy is Kantian but his negation of objects by Desire is more Humean. The Master-Slave relationship thus taps into both themes. In doing so, it tells a tale of how the mind must be united. The Master is then a place-holder for *any dominant internal force*. This ambiguity is part of why an internal reading is so hard to pin down.

But there are serious problems for an internal reading that cannot be overcome. One is that much of the text seems superfluous on such a reading. What could 'work' represent? What could 'objects' represent? Much of the language of the passage is external in a way that has no real internal suggestions to it. An internal reading would render the *Phenomenology* too solipsistic. It would eliminate the intersubjectivity that is clearly central to Hegel's agenda in explaining the development of *Geist*.

Having said that an internal reading is untenable, I do not want to dismiss its value entirely. There is insight that can be gained from it. The Master *is* an amalgam of Humean passion and Kantian reason. He is a placeholder for a dominant internal force. But he is not explicitly so. He merely carries suggestions. The parallel between the mind's coming to an internal equilibrium in which there is no dominant force and human relationships developing to equilibrium is used by Hegel to contribute to the sense of destiny. Hegel's world is conspiring towards *Geist*. Pointing to an internal reading is pointing to a literary device.

So Kelly's criticism does not entirely miss the point. The internal, external and mixed readings are not equally valid. An internal reading would be too solipsistic. A purely external reading would deprive the text of the device just discussed, therefore reducing the sense that the world is developing. A mixed reading is not just an equal option but is the least reductive of the three. However, Kelly is right that Kojève distorts this internal aspect of the Master-Slave. This is because the internal aspect is not so much a reading as a set of suggestions that Hegel exploits. Hegel's master and slave are being driven by forces outside of themselves. Conflicts from different dialectics recur and take new forms in later dialectics. It is not just that the master and slave are working out a conflict – it is that *Geist* is working out a path for itself. This allows Hegel to trade off different suggestions. Hegel's text has an ambiguity to it that Kojève's reading restricts. This shall be further revealed in the next section.

Historical time

The second criticism that Kelly makes of Kojève is that he makes the master and slave into historical persons, rather than ideal types. A similar point is made by Butler. She claims that the *Phenomenology* takes place within fictive time, whereas Kojève places it within historical time.¹⁸ These points are closely related. The fictive time is partly constituted by the master and slave being ideal types and vice versa. Further, this fictive time is vital for the ambiguity of the text.

Kelly stresses that his second point should be taken in conjunction with the first. This makes sense since if we believe that master and slave may represent internal principles just as much as historical persons, then we must accept that the text doesn't pin the master and slave down more firmly than as ideal types. We have rejected the possibility of an internal reading but Kelly's second point can still stand. In fact, this essay can be understood as a strengthening and remodelling of Kelly's argument.

¹⁸ *Introduction to the Reading of Hegel*, p. 72.

We are forced to see the master and slave as ideal types by the way that the text appeals to the reader. Charles Taylor's point is important here:

If we look at Hegel's most successful historical dialectics, the ones which are most illuminating and convincing, we find that in fact they convince the way any good historical account does, because they 'fit' well as an interpretation.¹⁹

No matter how Hegel might conceive of it, the appeal of the master and slave lies in its apparent applicability to situations across a broad range of experience. It is most obviously applicable to the actual historical development and fall of slavery. But we might also apply it across a range of political situations. We might look at some societies and see masters guided by Humean passion. And look at other societies and see masters led by Kantian reason. From a more psychological angle, we might see the brutish husband and the dominated housewife. Or we might see parent and child. Or teacher and student.

I do not pretend that the master and slave presents us with a model that fits any of the above situations perfectly. But I think it could reveal something of value about each. The appeal of the passage lies in its ability to unite such disparate features of experience.

The Master-Slave could not work in this way if it were located within historical time. This is why the fictive time of the *Phenomenology* is important. But Kojève historicises the narrative.

That a fictive account of time is in play in the Master-Slave should be clear from the discussion that introduces self-consciousness. Here consciousness contracts itself into an external world by conceiving of itself as within time. Time is structured by consciousness as a succession of unique shapes, each a for-itself. Out of these shapes consciousness constructs a unity in which each is an in-itself. The concept of time requires retaining these shapes as each in-itself and for-itself.

¹⁹Hegel, p. 217.

Kojève sees time as being created by desire.²⁰ Desire aims to bring about a new state of affairs in the future. Thus it requires a concept of time. It is by changing the world that we create time. Thus it is the slave's work that is the sole realisation of historical progress.²¹ The location of the dialectic within historical time is an important part of Kojève's philosophy. But it heavily prejudices us towards reading the Master-Slave as applying to political situations, as opposed to personal situations.

Master-slave as thought-experiment

From what I have argued, the Master-Slave appeals to us as a parable.²² It has a narrative that hangs together but not by a strict, explicit logic. It grips us because it presents a model that is applicable across a vast range of situations. And it has *normative force*. In each situation to which we apply it, the parable will give us some sort of resolution.

However, the Master-Slave is presented by Hegel as more than a parable. The *Phenomenology* is supposed to reveal the stages that consciousness must, of necessity, go through in order to reach its completion in *Geist*. This necessity means that we should be able to locate each stage in some concrete form. The stages of consciousness are properly embodied by historical eras. So we can reasonably expect to find the Master-Slave embodied at some early stage of the development of civilisation. This view is supported by a remark by Hegel in the ensuing section on stoicism that seems to pin it to a specific era:

As a universal form of the World-Spirit, Stoicism could only appear on the scene in a time of universal fear and bondage, but

²⁰ *Introduction to the Reading of Hegel*, p. 134.

²¹ Although the life and death struggle is also the work of desire so presumably that falls within historical time too. *Ibid.*, p. 23.

²² Verene, in *Hegel's Recollection* (Albany: State University of New York Press, 1985), uses a similar term. He has a view of the *Phenomenology* as a set of 'metaphors'. We could call the master-slave as a metaphor for relationships but I prefer parable.

*also a time of universal culture which had raised itself to the level of thought.*²³

One presumes that Hegel is referring either to ancient Greece, or possibly Rome.

Under this view, the dialectics of master and slave and stoicism etc. take on the character of thought-experiments. Hegel possesses a perception of all of the routes that are open to consciousness. The dialectics result from a postulation of 'if consciousness were at this stage, then it would develop like so'. I suspect that this is part of Kojève's motivation in trying to give the master and slave a tighter logic. There is a feeling that we should be able to set up the initial conditions of the master and slave and then deduce how things will run.

But if history is seen as developing according a strict logic then we must see the Master-Slave as a stage that has been and gone. This is disheartening since the appeal of the Master-Slave lies in being able to see the master and slave as pointing towards *live options*. Perhaps this is not excluded by Hegel. It might be that *Geist* has, in our time, overcome the master-slave but it might still be possible for individual consciousnesses to lapse into earlier forms. Or it might be that the master-slave continues to manifest itself insofar as it has been negated and *retained*. But neither reply gives the contemporary significance we would want for the Master and Slave.

It is a tension in Hegel that the nature of Dialectic undermines the appeal of the individual historical Dialectics. By giving the Master-Slave a stricter logic than Hegel, Kojève more significantly undermines the master and slave as pointing towards live options. His model allows us to see the industrialist and the worker and little else.

23 *Hegel's Phenomenology of Spirit*, Miller (trans.) (Oxford University Press, 1977), §199.

Bibliography

Aristotle, *The Nicomachean Ethics*, Ross (trans.), Ackrill and Urmson (rev.) (Oxford University Press, [1925] 1998)

Butler, Judith, *Subjects of Desire* (New York: Columbia University Press, 1987)

Engelhardt and Pinkard (eds.), *Hegel Reconsidered*, (Dordrecht: Kluwer Academic Publishers, 1994)

Hegel, Georg, *Hegel's Phenomenology of Spirit*, Miller (trans.) (Oxford University Press, 1977)

Kelly, George Armstrong, 'Notes on Hegel's 'Lordship and Bondage'', in *Hegel: A Collection of Critical Essays*, MacIntyre (ed.) (Notre Dame: University of Notre Dame Press, 1972)

Kojève, Alexandre, *Introduction to the Reading of Hegel*, Nichols (trans.), Bloom (ed.) (New York: Basic Books, 1969)

Stern, Robert, *Hegel and the Phenomenology of Spirit* (London: Routledge, 2002)

Taylor, Charles, *Hegel* (Cambridge University Press, 1975)

Verene, Donald Philip, *Hegel's Recollection* (Albany: State University of New York Press, 1985)

Yar, Majid, 'Recognition and the Politics of Human(e) Desire', *Theory, Culture & Society* (SAGE, London, Thousand Oaks and New Delhi, 2001), Vol. 18(2-3), p. 57-76

Russell and Kripke on names as definite descriptions

Elaine Yeadon

University of Sheffield

elaine.yeadon@bups.org

Bertrand Russell's infamous theory that names are disguised definite descriptions has fallen from favour. It has been beaten down by counterexamples which, using names as definite descriptions, yield confusion over counterfactual circumstances. I begin my paper by explaining the theory and will look at the motivation for it. I will then introduce Kripke's modal argument against it. From this it should be evident that there is weakness in Russell's argument in that it does not index between meanings in the actual world and other possible worlds. I will argue that a strengthened version of Russell's theory which introduces world indexed references for names and definite descriptions meets Kripke's criticism and retains the conclusion that names are disguised definite descriptions. Crucially for the Kripkean theorist, this still allows for names to remain as rigid designators.

Russell's theory that names are disguised definite descriptions is a response to such problems as that of the non-existent referent with regard to the law of bivalence. Briefly, this is that we want to hold that, for example, the statement 'The King of France is bald' is neither true nor false¹ because the referent of the definite description 'the King of France' does not exist. Yet this conclusion seems to compel us to reject a standard law of logic, namely the law of bivalence, which states that a proposition is either true or false. Another problem is that of Informative Identity Statements. If someone makes a claim in the form of ' $x = y$ ' such as the often quoted example 'Aristotle = the teacher of Alexander the Great', it would only be possible to understand

¹ A sentence with the form 'a is F', where 'a' is a referring term, is true if the referent expressed by 'a' has the property expressed by 'F' and false if the referent does not have the property expressed by 'F'.

completely what the statement means by being already aware of its truth. However we would like to say that many sentences of this form are more informative than this interpretation suggests.

Russell's response was that the propositions have a hidden underlying logical form which displays the actual non-contradictory meaning of the speaker. In the case of 'the King of France is bald', the logical form would be:

$$\exists x (Fx \ \& \ \forall y (Fy \rightarrow x = y) \ \& \ G x)^2$$

Here, F is the property of being king of France and G is the property of being bald. This sentence simply turns out to be false on analysis rather than neither true nor false. Furthermore, because names are disguised definite descriptions, the 'Aristotle = the teacher of Alexander the Great' case is solved in the same form. Simply substitute 'F' for 'has the property of being Aristotle' and 'G' for 'has the property of being a teacher of Alexander the Great' and one has a statement about two properties of a single entity rather than a mere identity assertion.

If Russell's theory is correct, when an agent uses a name they are in fact thinking of a definite description which picks out one unique object. If one interprets Russell as meaning that every name corresponds to a single definite description it may be difficult to give an account of how someone could use a name meaningfully without knowing the complete definite description. I do not think this poses a problem for Russell as there are at least two possible ways of avoiding it. One might say that an agent does not need to think of every aspect of a fixed definite description in order to think and speak meaningfully using the name; that is that they simply rely on others knowing, or being able to know, the full description. But this implies that names are reduced to possibly knowable unknowns (that's things we *could* know, but actually don't know). Alternatively, one might say that the agent's use of a name is equivalent to the definite description that they have for an object at the time that they use the name. This description can then change

² Loosely translated, this reads: 'there's exactly one thing which is the King of France and that thing is bald'.

depending on what knowledge the agent has regarding the referent of the thought or utterance. I prefer this latter option as I think it is closer to the way people naturally use names in specific situations without thinking about what others might know about the referent.

A more challenging criticism of Russell's theory comes from Kripke³. Kripke's response to Russell, in what is now known as the Modal Argument, is that if we are to accept Russell's theory, we ought to say that the descriptions relating to the name in use are only contingently true, that is to say that they could have been false if the world were different. Suppose, for example, that Aristotle instead of becoming the teacher of Alexander the Great had become a farmer. It makes sense for us to propose that 'Aristotle might not have been the teacher of Alexander the Great' yet it seems contradictory to say that 'Aristotle might not have been Aristotle'. This difference stems, Kripke argues, from the fact that definite descriptions are not rigid designators⁴ but names are.

One can happily suggest that Aristotle could have turned down the offer to teach Alexander the Great without changing the referent for the term 'Aristotle'. However, to suggest that the term 'Aristotle' might not correspond to the object Aristotle, as does the second sentence: 'Aristotle might not have been Aristotle', *would* require that we change the referent of the word rendering the complete sentence nonsensical. If Kripke's modal argument is correct, we can infer that the nature of thought about a name is not identical to the nature of thought about a definite description. When an agent uses a name in a proposition he has thoughts of contingent truths or falsities. This allows the agent to identify the same object in all possible worlds without involving contradictions in the propositions constructed.

My view is that, as Russell's theory stands, Kripke makes a strong criticism of it. He shows that we use names and definite descriptions in different ways which does imply that the name has a different meaning

³ Kripke S., *Naming and Necessity*, revised edition (Oxford: Basil Blackwell 1980).

⁴ Kripke states: 'Let's call something a *rigid designator* if in every possible world it designates the same object, a *nonrigid* or *accidental designator* if that is not the case.' Kripke S., *Naming and Necessity*, revised edition (Oxford: Basil Blackwell 1980) p. 48.

to the description, so Russell must be mistaken about them being the same. However, I believe that Russell's position can be amended to avoid Kripke's criticism. In its present form, Russell's theory does not provide a specific enough account of how speakers use names. Kripke has introduced the idea of possible worlds into the debate and as such *his* theory contains a level of precision which Russell's theory is lacking. Using the name 'Aristotle' as Russell's theory suggests one might is too ambiguous, for it does not specify whether one is speaking about an object in the actual world or in one of many possible worlds. Russell's assumption is that when one refers to an object, they speak of it, unless they specify otherwise, as we have access to it in the actual world. Kripke exploits this assumption by playing on the ambiguity between actual and possible referents. I think that Russell's theory would be strengthened against Kripke if it stated that the meaning of names needed to be indexed to a specific world.

This need for greater precision in a Russellian theory can be demonstrated through logical analysis. As Ahmed⁵ points out, Russell's original position results in the problematic sentence 'Aristotle might not have taught Alexander the Great' turning out to be a logical impossibility:

$$\diamond (\exists x (\forall y (Ty \leftrightarrow y=x) \ \& \ \neg Tx))$$
⁶

whereas we know that it *is* logically possible that Aristotle might not have become the teacher of Alexander the Great. However, if we narrow the scope of the diamond so that it is only relevant to x having the property T (as opposed to the whole statement including the identity of x and y) then the above contradiction does not occur:

$$(\exists x (\forall y (Ty \leftrightarrow y=x) \ \& \ \diamond \neg Tx))$$

⁵ http://www.phil.cam.ac.uk/teaching_staff/ahmed/kripke2.pdf.

⁶ Adapted from Ahmed. 'T' is 'has the property of being a teacher of Alexander the Great. The statement amounts to the utterly counterintuitive claim that x can be both T and not-T.

Gareth Evans⁷ argues that, given that we can make a distinction between statements in the actual world and a possible world, we can distinguish between the propositions 'Aristotle is entity x ' and 'the teacher of Alexander the Great is entity x ' without needing to alter the equivalence⁸ of the two. The first, he argues, requires that 'Aristotle' in the real world refers to entity x but allows that in a possible world it might not refer to that entity. The second requires that in a possible world the teacher of Alexander the Great is entity x . As they stand the two propositions cannot be combined to reach the conclusion that Aristotle was the teacher of Alexander the Great, but if we index both propositions to the actual world we end up with:⁹

1) Aristotle = x

2) The teacher of Alexander the Great = x

which do combine to conclude that Aristotle = the teacher of Alexander the Great.

Therefore, the two sentences, while they do not express the same propositions, can be shown to be equivalent, which was the claim that Kripke denied in his criticism of Russell.

So Evans would say that some distinct propositions use different names or definite descriptions but are equivalent in the actual world. If Evans is correct, and I think he is, it would imply that an agent can equate a name with a definite description in their thoughts and meaningfully express them in utterances. These utterances are duly understood by others as uniquely describing an object within the actual world.

However, the crux of Kripke's criticism is that names are rigid designators – that they identify the *same* object across *all* possible worlds. So far, the analyses do not attempt to specify how names rigidly

⁷ Evans, G., 'Reference and Contingency', *Collected papers*, (Oxford: OUP, 1985).

⁸ Evans states that if two statements are equivalent they would be verified by the exact same state of affairs. Evans, G., 'Reference and Contingency', *Collected papers*, (Oxford: OUP, 1985) p. 205.

⁹ The alpha subscript means 'in the actual world'.

designate in other possible world scenarios and so do not fully meet his criticism. Nevertheless, I believe that people *do* use the word 'Aristotle' as a rigid designator, although leaving it ambiguous as to *which* rigid designator it is supposed to be, and it is this recognition which I think braces the Russellian position against Kripke's attack. When we say 'Aristotle' usually, we refer to Aristotle in the real world. Suppose the real world is W1 and that other possible worlds are numbered W2, W3, W4, etc. When we say 'Aristotle' in the real world we speak of Aristotle_{w1}. Similarly, if someone from W2 were to say 'Aristotle' meaning 'Aristotle in my world (W2)', that is 'Aristotle_{w2}', they would be referring to Aristotle_{w2}. If someone from W3 were to say 'Aristotle_{w2}' it would have the same referent as when the person from W2 said it; that is, they would be talking about the Aristotle in W2. If they intended to speak of Aristotle in W3 but said 'Aristotle_{w2}' they would be using the wrong name. Thus, 'Aristotle_{w2}' is a rigid designator just as 'Aristotle_{w1}' or 'Aristotle_{w3}' are.

Our typical failure to index our utterance meaning explicitly is, I think, indicative of the fact that the issues raised by this debate are more to do with the meaning of the agent's thoughts than with the meaning of the utterances themselves. Consideration of the theories has shown that most people do not analyse how to index their propositions appropriately unless problems arise such as the one highlighted by Kripke's criticism. However I believe that agents do perform some natural indexing, if only to allow their thoughts and utterances of names to default to referents in the actual world.

In conclusion, I have argued with Evans and Ahmed that Kripke's modal argument, while it defeats Russell's original non-specific theory, fails to defeat a strengthened, world indexed one. The result of this is that we can retain the claim that when we think about a name it has the same meaning for us as a definite description for the object to which it refers.

Bibliography

Evans, G., 'Reference and Contingency', *Collected papers*, (Oxford: OUP, 1985)

Kripke S., *Naming and Necessity*, revised edition (Oxford: Basil Blackwell, 1980)

Martinich, ed., *Philosophy of Language*, 4th edition, (New York: OUP, 2000)

Ahmed, A. (2005) *Notes for Lecture 2, Kripke on Names*, Philosophy Part 1B, http://www.phil.cam.ac.uk/teaching_staff/ahmed/kripke2.pdf

Dennett's stance: legs akimbo

Alex Carruth

Durham University

a.d.carruth@dur.ac.uk

In his papers *Intentional Systems* and *Real Patterns*, Daniel Dennett offers an account of intentionality which purports to successfully accommodate such phenomena in a naturalistic account of mind. It is my intention in this paper to demonstrate that Dennett's position cannot reasonably be maintained. I shall challenge Dennett on two fronts, through the exposure of an inconsistency in his account which I do not believe he can avoid and by challenging an underlying assumption upon which Dennett's approach rests. In the final section of this paper I intend to introduce an alternative account of intentionality.

Intentionality is the property of 'being about something' that is associated with various mental states and processes. Mental states such as beliefs, desires and fears are all identified as intentional states; certain of our beliefs are about some specific object or set of objects, we tend not to desire in general; rather we are desirous of something. When faced with a rapidly approaching bull, on the day we chose to wear our favourite red t-shirt, the sense of dread and overwhelming fear is not dislocated or abstract, it is strongly associated with the presence of the bull, our spatio-temporal position in regards to the bull, our beliefs about what the outcome of the imminent collision between ourselves and the bull will be and our desires to avoid these outcomes. The object of an intentional state is its mental content. As exemplified above, the mental content of a single intentional state can incorporate several objects, and these can be other mental states as well as features of the external world. We constantly engage in mental behaviour which appears to be intentional, and upon introspection it seems very clear that many of our everyday mental states, processes and phenomena are

about objects external to themselves, that is to say possess mental content, and are in fact intentional.

Intentionality is seen as problematic for just this reason; intuitively it seems to be something real, but it cannot easily be reconciled with a scientific account of the world. One committed to the project of naturalism must ask just how is it that an internal mental process can be actually related to an external object or set of objects, what the nature of such a connection might be. Several attempts have been made to answer these questions, notably the causal and teleological accounts of intentionality. It falls outside of the scope of this paper to explore these fully, but I agree with Lowe when he claims that whilst:

Both theories provide quite plausible accounts of certain types of naturally occurring representational states ... neither seems straightforwardly applicable to the kind of mental representation involved in the attitudinal states of intelligent subjects of experience.¹

Whilst the claims made by Lowe regarding these two positions cannot be considered completely uncontroversial, I believe they carry considerable weight, as such they shall be assumed for the sake of this paper. If one is to maintain a naturalistic outlook without banishing intentional states and processes from our theory of mind, which, given the weight of intuition regarding the existence of such phenomena, would be somewhat unpalatable, then an alternative account must be sought.

It is Dennett's claim to provide just such an account. He claims that we engage with and understand the external world through the employment of three interpretational stances, by adopting these various stances we make the world intelligible. These three stances he orders hierarchically. The most primary of these, that which yields an

¹ Lowe identifies major difficulties with each theory. His primary charge against a causal theory of intentional states is that it seems to endow abstract entities with causal efficacy. The teleological approach implies that only individuals with an evolutionary history are capable of intentional mental behaviour, which Lowe finds problematic. For a full account, see Lowe, E J; *An Introduction to the Philosophy of Mind* pp. 89-101.

interpretation of the world which is closest to the actual way the world is. Dennett calls the physical stance, concerned with the exact physical make up of an object we wish to understand. Dennett analogises this with the bitmap file format of a computer, which, by storing the information of the exact position of each bit of data in an image, provides a completely accurate but computationally cumbersome storage medium. Ideally we would always interpret the world under the physical stance, were it not for the ungainly nature of this approach. Next is the design stance, which we adopt when attempting to predict the behaviour of a certain object by an assumption about its proper function, we most regularly adopt this stance when considering machines and other human artefacts. Dennett calls his third stance, that which is lowest in the hierarchy, the intentional stance.

To consider the world under the intentional stance is to treat the entities, or a set of the entities, within that world, *as if* they were really possessing of intentional mental states and processes such as beliefs, desires, fears and so on. The adoption of this stance, Dennett claims, is as a predictive and explanatory tool. If we can successfully predict and understand the behaviour of an entity when we consider it under the intentional stance, in a way which we cannot through the physical or design stance (be this due to a lack of information required to adopt a more primary stance, or the practical difficulty of doing so) then we are licensed to consider that entity as actually engaging in intentional mental states and processes. As Dennett puts it:

...all there is to being a true believer is being a system whose behaviour is reliably predictable via the intentional strategy, and hence all there is to really and truly believing that p... is being an intentional system for which p occurs as a belief in the best (most predictive) interpretation²

Under such an account, we can consider intentional states as existing in an entity just in virtue of our adoption of the intentional stance towards

² Dennett, D; 'True Believers: The Intentional Strategy and Why It Works' in Heath, A. F. (ed.); *Scientific Explanation* p. 68.

that entity, that is to say, just in virtue of how useful a predictive tool the intentional stance is in regards to the behaviour of that entity.

Dennett claims that when we adopt the intentional stance we ascribe intentional states to an entity in order to predict its behaviour. To do so it is clear that we must have an understanding of intentional states, if we did not it would be hard to explain how it is we ascribe them to the entity in question. Yet, if intentional states exist just in virtue of the success of adopting the intentional stance, as Dennett's account claims, then how could it come about that any interpreting agent could have knowledge of intentional states prior to the very first adoption of the intentional stance? If in order to have adopted the intentional stance in the first instance we must have had a firm grasp of the concepts of intentional states, then intentional state concepts must exist independently of our predictive exercises. It is hard to see how one might come to acquire such concepts without intentional states existing independently of the stance we adopt. If this is the case, then Dennett's approach is weakened, as it seems intentional states derive existence from some source prior to our adoption of the intentional stance, and he faces the same problems as initially; namely, how can we explain these prior-to-stance-adoption intentional states?

Furthermore, what is it to adopt the intentional stance? If we are to hold that intentional states are dependent for their existence on the successful adoption of the intentional stance, then how are we to specify what it is to adopt such a stance whilst avoiding intentional states inherent therein? The adoption of a particular stance towards an object or set of objects is to hold a certain set of beliefs about those very objects, or at the very least to hypothetically hold to obtain a certain set of beliefs, attitudes and assumptions about those objects. To so do, to adopt a stance, is to engage in an intentional process. So, as Slors points out, this amounts 'to characterizing the intentional stance in terms of intentional states.'³

³ Slors, M; 'Why Dennett cannot explain what it is to adopt the intentional stance' in *The Philosophical Quarterly* 46 (1996) p. 95.

Dennett's account relies on the engagement of intentional mental behaviour in order to adopt the so called intentional stance, and necessitates knowledge in terms of intentionality prior to being able to adopt such a stance. This renders Dennett's account of intentional states and processes as existing just in virtue of making successful predictions about entities through the adoption of the intentional stance inconsistent and untenable. Having shown Dennett's position to be technically flawed, I now wish to challenge the underlying assumptions which motivate Dennett's approach.

Dennett maintains that intentional states are merely the products of an interpretation and prediction method we use to understand day to day events in the world, essentially a rudimentary form of quick computation and estimation, confused by 'noise'. His position regarding the design stance is similar, it is a useful way of interpreting and predicting the behaviour of objects, but essentially confers no realism itself. However, he claims that to consider an object via the physical stance, in regards to temporal and spatial location, and relation to physical laws, is consider an object really, that is to say, the observations one might make when considering an object through the physical stance carry ontological weight. The physical features of objects are taken to be stance-independent. However, it is unclear why the physical stance is given this privileged position. Dennett is guilty of an unwarranted prejudice in favour of interpreting the world through the mechanisms of modern science under the prevailing paradigm. This prejudice leads him to relegate intentional interpretation of the world to the lowest position in his hierarchy, and motivates his analysis of intentional states as mere products of this interpretation. Whilst I in no measure am suggesting modern science is without use, I can see no rational reason for the primacy it is so often given, not least by Dennett, in regards to the ontological weight scientific explanation confers. This may be motivated by the primacy physical science justifiably enjoys in terms of explanatory efficacy, however, to transfer this to the realm of ontology is without justification. What Dennett sees as inherent 'noise' (resulting in inaccuracy and therefore a lower place in the hierarchy of stances) in the system of intentional prediction seems in fact to be lack of information, similar to that which inhibited early science (and still inhibits science today, we don't have a GUT yet!).

Dennett's account, whilst of some use in understanding the ways in which we make our world intelligible, fails to provide a plausible explanation for the intentional phenomena which occur within the mental behaviour of intelligent subjects of experience. This given, we must uncover some other explanation of the participation of intentional phenomena in our mental behaviour.

Intentionality must be subject dependent; it cannot be an independent feature of the external world. Without some entity to engage in apparently intentional processes, there could be no intentional phenomena. Nor can we consider intentionality to be an essential feature of particular mental states or processes, as we can experience what are usually intentional mental states non-intentionally; fear without any particular object, an unsettling desire for something we know not what. Perhaps intentionality is explicable as a relationship in which mental states and processes stand to objects external to themselves. Intuitively such an explanation seems attractive, however, we regularly engage in intentional mental behaviour without any existing object, or we are mistaken about the proper object of that apparently intentional state; if we are to see intentionality as a relational property, this must be explained. Furthermore, even if one did establish grounds for understanding intentionality in this way, one would be left with the question of what exactly such a relationship would be like, and I find it hard to see how this might be answered in non-intentional terms.

I wish to suggest that intentionality is to be understood in terms of the qualia produced in a subject of experience upon the co-consideration of a certain proposition or set of propositions and some object or concept external to that proposition or set of propositions. Intentionality cannot be considered independently of the specific subject of experience for whom the particular example of intentional phenomena occurs. For a mental state or process to be intentional *is* for the entity for whom that state occurs to feel it is so, just in virtue of the presence of these intentional qualia can we consider intentionality to exist. We engage with and understand the world around us through the filter of our basic emotional responses to experience, without emotion we would be left

with an alienating, dispassionate and largely unintelligible world. Our emotional framework provides the basis for making these associations which underlie intentional phenomena, which whilst being mostly non-rational, are none the less of great pragmatic value in our engagement with the world. The very property of aboutness which identifies intentional mental phenomena in intelligent subjects of experience is conferred by these responses. Intentional phenomena are quasi-emotional phenomenologically defined responses to our experiences of the external world.

The notion of unconscious intentional mental phenomena seems to pose a problem for such an account. If a subject is on every level unaware of an occurring intentional phenomena, then it follows that there can be no accompanying intentional qualia, as this would constitute an awareness of a kind, and so an account of intentionality in terms of intentional qualia as provided above collapses. However, if such a counter-example is to be invoked, then the onus is on those who wish to invoke it to provide a compelling account of an occurring intentional state or process of which the subject is entirely unaware. An account of the kind given above is not merely epistemological, that is to say an account of how we come to know about intentional states. Rather it treats what it is for a given mental phenomena to *be* intentional, namely that the co-consideration of that state or process with some feature external to that phenomena is accompanied by intentional qualia.

Understood in this way, I do not see that the existence of intentionality continues to pose a specific problem for those who are committed to the project of naturalism. It is beyond the scope of this paper to explore all aspects of the above account, such as the exact nature of intentional qualia, what mechanisms produce them, what causes underlie their production and so on, and in all these areas further work is necessary.

Bibliography

Dennett, D; 'Intentional Systems' in *The Journal of Philosophy* 68 (1971); pp. 87-106

Dennett, D; 'Real Patterns' in *The Journal of Philosophy* 88 (1991); pp. 27-51

Heath, A F (ed.); *Scientific Explanation*; Oxford; OUP 1981

Lowe, E J; *An Introduction to the Philosophy of Mind*; Cambridge; CUP 2005

Slors, M; 'Why Dennett cannot explain what it is to adopt the intentional stance' in *The Philosophical Quarterly* 46 (1996) pp. 93-98

Hume's causal reasoning

Rhonda Grintuch

University of Toronto

Rhonda.grintuch@utoronto.ca

'No element of Hume's philosophy has had a greater and more lasting influence than his theory of causality. It has been frequently attacked, and frequently misunderstood' (Ayer, p.55). Hume indeed took his theory of causality to be fundamental to his philosophy, claiming that, 'all reasonings concerning matter of fact seem to be founded on the relation of *cause and effect*' (*Enquiry*, 323). The present paper will not focus on Hume's theory of causal reasoning but more specifically, it will assess Hume's definition of 'cause,' which underlies his important theory. Hume advances two definitions of 'cause.' Although they might appear to play but a minor role in his theory of causality, they serve as the foundation for his famous theory of causal reasoning and, moreover, are in themselves a contentious issue. When evaluating Hume's definitions of 'cause', this essay will focus in particular on the claims, mainly advanced by Robinson, that Hume's definitions are inconsistent. The evidence will show that Hume's definitions are not in fact inconsistent, but are the subject of misinterpretations.

There are two potential difficulties with Hume's theory of causality that must be discussed from the start, so as to avoid any inconsistencies in our later discussion of his theory. It is important to note that when Hume speaks of the relation of cause and effect, he means something quite different from the meaning in current use. We are accustomed to distinguishing, say, between causal and statistical laws; or of drawing a distinction between the events that are related because they are the cause and effect of a single occurrence or, alternatively, are the common effects of a single cause. In contrast, Hume more generally discusses the 'causal connection' as encompassing any law-like connection between matters of fact.

Another potentially problematic issue in the literature on Hume's theory concerns Hume's point of reference for his theory of causality. Hume's *Treatise* is often read as discussing the relation of causality as holding between events, as opposed to between objects, despite the fact that he *does* define causality as a relation holding between objects (see definition (1) below) (Ayer, 56). This seemingly inaccurate reading of Hume seeks justification in the fact that he refers to mental elements, like feelings and volitions, as causes and effects. A more accurate reading of Hume's theory of causality, suggested by Ayer, would interpret the relation as holding between matters of fact, such that it includes objects, events, actions and passions, and physical and mental states and processes. This interpretation is consistent with Hume's enlargement of the scope of the causal relation, which, as mentioned earlier, refers to matters of fact. This reading would also encompass both definitions of 'cause' put forth by Hume.

The specific focus of this essay, an evaluation of Hume's definition of 'cause', will now be discussed. An integral part of Hume's philosophy seems to involve analyzing and clarifying such concepts as 'cause'. J.A. Robinson accuses Hume of 'engaging in a twofold treatment of the causal relation,' empirical psychology and philosophical analysis, such that 'it culminates in *two* different definitions of the term 'cause' (Robinson, 'Cause', p.130). What is more serious is Robinson's claim that the two definitions that Hume proposes are, both intensionally and extensionally, at odds. Let us now turn to a discussion of Hume's definitions of 'cause' and discuss their potential problems.

The crucial passages in which Hume defines 'cause' are the following. In one case, Hume claims (1) 'we may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second' (*Enquiry*, 362). However, on that same page, Hume states (2) 'we may...form another definition of cause, and call it, an object followed by another, and whose appearance always conveys the thought to that other' (*Enquiry*, 362). These are considered 'natural' and 'philosophical' definitions respectively, due to the fact that in the first case we attend only to the phenomena in which the relation is exhibited and in the second case, we consider in addition the way in which we view the phenomena (Ayer, 67).

It is clear that there are some defects in Hume's two definitions of cause, aside from the major criticism of their inconsistency, which Robinson wages against Hume and which will later be considered. Ayer believes it is problematic in the first place to 'include a reference to the mind's propensity in what was supposed to be a definition of causality' (Ayer, 68). Ayer's reason for this accusation does not seem to warrant his claim that Hume makes 'a venial mistake' (Ayer, 68) on this point; his argument is, nonetheless, worth considering. Ayer claims that when making causal judgments, we are as a result expressing our 'mental habits' of inferring causal connections; but when asserting these causal judgments, we are not also stating that we have such mental habits. Thus, we assert causal judgments, but 'that is not to say that when we attribute causal properties to some objects or events, we make an assertion about ourselves' (Ayer, 68). As a result, Ayer criticizes Hume for stating a definition of cause that holds the assertion of our mental attitudes as integral to causal relations.

I cannot agree that Ayer makes a legitimate criticism of Hume on this point. It is true that to attribute a relation of causality is a mental habit, and this habit is implicit in any such attributions. From this, it does not seem correct to claim, as Ayer does without further justification, that attributions of causality 'are not making an assertion about ourselves' (Ayer, 69). This seems to me to require a further discussion of whether or not, in making a judgment of causality, we are attributing the 'causal' property to the objects. It seems that a case can be made for why the causal relation would be asserted as holding between our ideas of the objects or events, and not between the events themselves. One could argue that we can only discuss our perception of events and objects and, thus, only attribute the property of causality to our ideas, as a connection we perceive and not as one that we would assert as holding in the world. Moreover, Ayer's criticism might even require a discussion of the existence of a mind-independent world, and it seems that he is required to give a justification for this notion in light of his belief that an expression of a causal relation says something only about the causally-related objects and not anything about ourselves. Thus, I cannot agree with Ayer that Hume is mistaken to assert that mental attitudes are integral to his definition of 'cause'.

Another potential problem with Hume's theory involves, as Robinson tries to show, a lack of equivalences between the two definitions of 'cause.' To demonstrate his point, Robinson claims that one definition implies a class that contains different members than those contained in the second class, as implied by the second definition. Thus, the class of objects that are true of the first definition does not define the same class of objects that satisfy the second definition. His argument proceeds as follows. Robinson argues that the (aforementioned) definition (1) determines a class of ordered pairs (x,y) of *particular occurrences*. Thus 'C(x,y)' denotes the relationship dictated by (1). Furthermore, in the relationship between objects as outlined by definition (1), it is not required that the objects x and y *be observed by anyone as having occurred*. Robinson further points out that according to (1), even if someone does observe both x or y or both, 'it would not be necessary for him to be aware that he had witnessed an instance of general uniformity' in order for C(x,y) to obtain (Robinson, 'Cause', 131).

At this point, it seems to me that Robinson's argument is problematic, for how could someone witness something and not be aware that he had witnessed it? As an example, we might consider a case where I am present during a shooting and had been watching the shooter and the victim as the shooting took place. However, I happened to be daydreaming and in fact, I was not aware of the shooting that took place. In this (however unlikely) case, I would not be a worthwhile witness for the police. I had not, in fact, witnessed the shooting.

In any case, however problematic, Robinson goes on to claim that a witness need not be aware of the causal relationship between x and y , such that y follows x whenever x appears. As an additional implication of (1), Robinson holds that the causal relationship existing between objects 'x' and 'y' depends on much more than the circumstances immediately surrounding their particular occurrences.

Robinson wants to argue that in contrast to definition (1), definition (2) has very different, even incompatible, concomitants. In the case of definition (2), the causal relation refers to a class of ordered pairs a,b , which can be denoted as D(a,b). These members are particular

occurrences and belong to the class in virtue of a certain property they share, a property 'which is defined quite essentially in terms of a certain *mental* phenomena' (Robinson, 'Cause,' 131). According to Robinson, the relation $D(a,b)$ is a *definite* relation. Thus, the relation implies that some human observer has observed either a or b or both and, as a result, would be compelled to assent to both the view that the idea of an occurrence of 'a' will be followed by the idea of the occurrence of 'b', as well as the fact that an observation of the occurrence of 'a' will result in an expectation of an observation of the occurrence of 'b'. Additionally, Robinson claims that knowing the circumstances that immediately surround the occurrences of 'a' and 'b', it would be possible to determine whether or not the relation $D(a,b)$ holds true. For example, consider the case that I happen to know the exact circumstances through which positively charged electrons build up near the top of clouds and negative ones at the bottom of clouds, and the fact that the negative charge seeks a path to the positively charged ground, etc. Then, upon observing 'a' (lightning) and 'b' (thunder) I can determine whether the relation 'the interaction of positively and negatively charged electrons in clouds causes (lightning, thunder)' holds true.

Given the above, Robinson claims that (1) and (2) differ in both their meaning and their extension. But is Robinson correct? One might first look to Hume's own writing to evaluate Robinson's claim. It is clear that Hume *deliberately* put forth two definitions of 'cause,' and this is evident from the following excerpt:

There may two definitions be given of this relation [cause], which are only different, by their presenting a different view of the same object, and making us consider it either as a philosophical or as a natural relation; either as a comparison of two ideas, or as an association betwixt them (Hume, Treatise, III, XIV).

Thus Hume intends for his definitions to denote the same class of members and to differ only in that one is a philosophical definition and the other is a natural relation. As mentioned above, according to Hume, a philosophical relation is a comparison between ideas and a natural relation is an association between ideas.

Robinson attempts to reconcile Hume's supposedly inconsistent definitions by concluding that Hume's definition (1) is a true definition of 'cause' and definition (2) is actually an empirical psychological theory concerning the relation expressed by (1). Richards, however, rejects Robinson's claim that (2) is an empirical comment on (1) in favour of the more likely and compelling view that Hume presents two definitions of 'cause'. On this point, Richards's view seems consistent with Hume's own intention of introducing (1) and (2) as *definitions of cause*. Richards is essentially critical of Robinson's view that Hume's definition (2) can be reconceptualized so as to be a restatement of (1) such that the causal relation is a natural relation. On this point, Richards states, 'it must not be thought of that Hume anywhere argues that all causal relations are natural' (Richards, 155). It is worth looking more critically at what Robinson takes 'naturalness' to be, and why he claims that causality is a natural relation. He defines 'naturalness' as:

simply the property of any relation R between a thing or event A and a thing or event B (not between the idea of A and the idea of B) whereby the observation of A and B standing to each other in relation R is enough to induce an association between the idea of A and the idea of B (Robinson, 136-137).

Robinson also asserts that there are three relations between things or events that possess the property of being natural: resemblance, close spatial or temporal proximity, and *the cause-effect relation* (Robinson, 137).

Robinson's definition of naturalness and inclusion of 'resemblance' in natural relations is at odds with what Hume himself states, which suggests that Robinson's conception of naturalness and statement that (2) is an empirical claim on (1) is flawed. Hume claims that resemblance 'is a relation, without which no philosophical relation can exist ... But tho' resemblance be necessary to all philosophical relation, *it does not follow, that it always produces a connexion or association of ideas*' (italics mine. Hume, *Treatise*, I, 5). This, however, is inconsistent with Robinson's aforementioned definition of naturalness.

Given that Robinson includes resemblance under the category of 'naturalness', and that at the same time that resemblance is inconsistent with the characteristics of naturalness, it seems unlikely that he presents a correct interpretation of Hume. Rather, it is more plausible that, as Hume clearly states, he indeed puts forth two definitions of 'cause', both of which present a different view of the same aspect. This is what Richards concludes, claiming that Hume is 'presenting a different view of the same object' (Richards, 151) and thus Hume is presenting the same relation, *not the same individuals related*. This further undermines Robinson's criticism that Hume's two definitions are inconsistent because they pick out members of two different classes

Our discussion has shown that Hume's text often allows for different interpretations, but does not equally support them. Philosophers might be led to advance unjust criticisms of Hume based on the fact that his definition of causality incorporates claims about mental attitudes. Although it might at first seem bizarre to include a statement about mental attitudes in a definition of cause and effect, this discussion has shown that it is not so problematic. Furthermore, it might at first seem bizarre to advance two definitions of 'cause', but the definitions are in fact consistent. It can be concluded that Robinson is mistaken to criticize Hume's definitions as inconsistent, and his attempt to reconcile these definitions is untenable.

Bibliography

Ayer, A.J., *Hume*, New York: Oxford University Press, 1980

Flew, Antony, *Hume's Philosophy of Belief*. New York: The Humanities, Press, 1961

Hume, David, *A Treatise of Human Nature*, Carl Mickelsen, McMaster University, Archive. March 14th, 2005:
<<http://www.class.uidaho.edu/mickelsen/ToC/hume%20treatise%20ToC.htm>>

Hume, David, *An Enquiry Concerning Human Understanding*, New York: Random House, 1990

MacNabb, D.G.C., *David Hume: His Theory of Knowledge and Morality*, Connecticut: Archeon Books, 1966

Richards, Thomas J., 'Hume's Two Definitions of 'Cause.'" In V.C. Chappell, *Hume*, New York: Anchor Books, 1966, pp.148-161

Robinson, J.A., 'Hume's Two Definition of 'Cause.'" In V.C. Chappell, *Hume*, New York: Anchor Books, 1966, pp.129-147

Robinson, J.A., 'Hume's Two Definitions of 'Cause' Reconsidered.' In V.C. Chappell, *Hume*. New York: Anchor Books, 1966, pp. 162-168

Waxman, Wayne, *Hume's Theory of Consciousness*, Cambridge University Press, 1994

Wright, John P., *The Sceptical Realism of David Hume*. Manchester University Press, 1983

Book reviews

On opera

Bernard Williams, *Yale University Press*

Andrew Goldfinch

London School of Economics

a.m.goldfinch@lse.ac.uk

Bernard Williams was not only a lover of wisdom but also a lover of opera. He served on the Board of the English National Opera and over the years wrote entries for books and opera programmes. *On Opera* collects these scattered entries into one volume. The collected pieces – sixteen in total – range from three pages to twenty pages in length. Like his other posthumous works, the book is prefaced by his widow and introduced by one of his friends (in this case, Michael Tanner).

Perhaps the most well known example of opera inspiring Bernard Williams is Janáček's *The Makropulos Case*. According to his widow, it was a performance of this opera that encouraged Williams to reflect on the tediousness of immortality. This is only mentioned in passing in the editorial preface but it's worth elaborating a little. The opera focuses on Elina Makropulos, a woman who has maintained her age, and character, for three hundred years. She has become bored because she has exhaustively experienced all the things that could happen to a woman of her age. Her life has become joyless, a mere life of boredom. For Williams, this is not the result of an accident of her character; it's the inevitable consequence of immortality. Williams's argument for this conclusion can be found in his 1973 paper 'The Makropoulos Case: Reflections on the Tedium of Immortality'.

On Opera does not have a main thesis or argument or proposition to summarise, analyse, and discuss. Instead, the book contains numerous philosophical reflections on specific operas by Mozart, Wagner, Verdi, Puccini, Debussy, Janáček, and Tippett. Most of the chapters contain

multiple philosophical reflections and observations about their target opera. In what follows, I shall review and discuss certain chapters in more depth than others, leaving some with no more than a description. I hope this will strike the right balance between discussing some philosophical reflections contained in the book and providing a survey of the entire book.

The first chapter – ‘The Nature of Opera’ – begins by defining opera. Opera is a form of dramatic art. In order for a dramatic art to be an opera, it must satisfy at least two necessary conditions: it must be staged and it must be sung. Issues relating to satisfying these conditions are identified. Williams notes, for example, that there are verbal styles that challenge the distinction between speech and song.

Operas are distinguished from oratorios (which do not have scenery or acting), operettas (like operas but generally perceived to be less serious), and musicals. However, the distinctions are to some extent arbitrary. There are performances that can be classified as either an opera or an operetta and there are performances that can be classified as either an operetta or a musical (to cite a popular example – *Jerry Springer: The Opera* can be classified as either an operetta or a musical). This is particularly the case with the distinction between opera and operetta. Before I read *On Opera*, I was under the impression that operas were defined as being entirely sung whereas operettas were partly spoken, partly sung. However, Williams notes that in both the French and German traditions, the existence of speech in a dramatic art does not prevent it being an opera. Rather, what seems to distinguish operas and operettas is ‘musical ambition and dramatic content’. Thus, Williams says that the term ‘opera’ is ‘to some extent an evaluative term, used to refer to sung drama which is either ‘serious’ enough, or traditional enough in form and technique, to be stage in an opera house’ (p. 4). Pointing out that the term ‘opera’ has an evaluate element, as well as denoting differences in singing and performance style, seems to me to be accurate. There are often stylistic differences between what we term operas and operettas. However, since opera itself contains a wide range of musical and performance styles, sometimes the classification of a performance as ‘opera’ or ‘operetta’ is largely evaluative. However, the distinction between operas and musicals is more substantive: there are

important differences in the style of both performance and singing. Williams points out that attempts by opera companies to perform musicals have generally failed.

After discussing some difficulties and limitations of *opera seria* (a style of Italian opera prevalent in 18th century), Williams turns his attention to the resources of opera as a medium. Williams claims that the single most powerful resource of opera is its capacity to use music to deepen the action in time. The orchestra and singers can powerfully represent the inner dimension of action – motivation, mood etc. In Mozart and Da Ponte's operas, this resource is used to distinguish characters in various subtle ways. For example, in *Don Giovanni*, characters sing about themselves – all except Giovanni. In Wagner's operas, 'every aspect of the characters' experience is expressed and nourished by the unified flow of musical material' (p. 13). The chapter ends with remarks on performance and repetition.

The second chapter – 'Mozart's Comedies and the Sense of an Ending' – is the first of four consecutive pieces primarily devoted to Mozart. As the title implies, this piece is concerned with the endings of Mozart's comedies, specifically *The Marriage of Figaro* (hereafter abbreviated as *Figaro*), *Don Giovanni*, *Così fan tutte*, and *The Magic Flute*. Since the next three chapters in the book are individually on *Figaro*, *Don Giovanni*, and *Così fan tutte*, it would seem more logical or natural to have placed this piece after these three chapters.

'Mozart's *Figaro*: A Question of Class?' (chapter 3) explores the question of whether Da Ponte and Mozart's *Figaro* is a revolutionary socio-political critique. Mozart's *Figaro* is based on Beaumarchais's *Le mariage de Figaro*, a politically radical play that was initially banned in Vienna. The traditional understanding is that De Ponte and Mozart changed the story into a brilliant work of sensuality and character, but one that is free of socio-political content. Williams points out in order for something to be a radical socio-political critique, it need not explicitly identify great problems on a macro scale, to explicitly demand action to remedy it or propose a solution. Social critique can be radical and yet subtle: it can display human emotions and relations between humans in various social contexts, and show how those social contexts

form and change these emotions and relations. For example, we can explore the feelings of loneliness that are caused by possessing a social rank. By representing humans in realistic situations, opera can offer socio-political critiques without being explicit.

For me, one of the advantages of the dramatic arts is that they are able to model agents with various predispositions, tendencies, degrees of freedom, motivations, intentions, ethics; they can model various relations between these agents; and they are able to model domains of various conditions and structures. Indeed, they can even model conflicts within individual agents. This can be used for socio-political critique but also for exploring moral motivation, testing the consequences of norms and ideas when applied to various situations, exploring the effect of society on natural dispositions, and so on. In 'Don Giovanni as an Idea' (chapter 4), Williams discusses Giovanni as an idea since Mozart. *Don Giovanni* is a story about a seducer; indeed, he is the embodiment of sensual desire. He is also portrayed as a hero. What drives Giovanni? Is he a hero in the existential sense? Don Giovanni is a good example of modelling an agent to embody a principle and then observing that principle in action on the stage. Of course, a thought experiment can also model agents. However, opera has an advantage over thought experiments in the form of the resource Williams mentioned in his first chapter: the ability, by musical means, to more fully represent inner states and so deepen action. As Nietzsche observed, 'By means of music the passions enjoy themselves' (*Beyond Good and Evil*, § 106). Music can affect our psychological state; the music in operas can help shed light on the inner dimensions of social events.

Chapter 5 – 'Passion and Cynicism' – looks at Mozart's most problematic opera (*Così fan tutte*), the problem being the emotional power that Mozart supplies to an artificial narrative structure. Verdi's *Don Carlos* is discussed in chapter 6 ('Rather Red than Black'). Chapters 7 to 9 are pieces on Wagner. 'Tristan and Time' (chapter 7) is essentially a summary of *Tristan und Isolde* with some observations. Williams observes that in the opera, Wagner's method of internalising action and making music convey psychological states is carried to the extreme. Wagner's *Ring* is discussed in chapter 8 ('The Elusiveness of Pessimism'). Chapter 9 – 'Wagner and the Transcendence of Politics' –

is concerned with the problem of how we should think about Wagner. Williams holds that we should care about Wagner's works. The problem, however, is that Wagner held some very repellent socio-political views, such as anti-Semitism. In the case of Wagner, can we appeal to the distinction between the man and the work? Are his repellent views present in some form in his operas? These and other problems are discussed.

Remarks on Debussy's *Pelléas et Mélisande* form chapter 10 ('L'Envers des destinées'), Puccini is examined in chapter 11 ('Manifest Artifice'), and a review of an opera book features as chapter 12 ('Comments on *Opera and Ideas: From Mozart to Strauss* by Paul Robinson'). Chapter 13 – 'The Marriage and the Flute: Tippett and Mozart' – is another piece I feel could have been better placed within the book (i.e. placed next to the other Mozart pieces). This piece compares Tippett's *The Midsummer Marriage* with Mozart's *The Magic Flute*. Tippett's *The Midsummer Marriage* was modelled after *The Magic Flute*. There are similarities between the two stories: both stories have two couples, one more spiritual, the other more earthly; and humans in both stories encounter the supernatural, which is expressed in non-orthodox symbols. However, Williams observes that the resemblances between them might not be deep. Whereas *The Magic Flute* manifests a belief in human powers, benevolence and sense, *The Midsummer Marriage* is more drawn towards the supernatural, manifesting Jungian conceptions of opposites and their eventual unification that transcend human understanding.

Chapter 14 is entitled 'Janáček's Modernism'. Janáček is an exemplar of the modernist demand to do more with less. Williams notes that formalism and technicality are not sufficient to satisfy the modernist demand to do more with less. Indeed, they can result in doing less with more. For Williams, Janáček serves as a reminder of what should be done: to use technical complexity and self-awareness to address real concerns, and to articulate and alter emotions. The penultimate chapter – 'Authenticity and Re-creation' – explores the relations between the study, production, and performance of music. The final chapter – 'Naïve and Sentimental Opera Lovers' – explores the distinction

between the 'naïve' and the 'sentimental' as applied by Isaiah Berlin to artists.

It should be clear that this book isn't an introduction to opera or aesthetics. Readers with no knowledge or experience of opera will find the book difficult to digest. Knowledge of operas is at times assumed – which is understandable, given that some of the chapters are from opera programmes – so a familiarity of opera, or willingness to do further reading on the subject, is necessary in order to appreciate the philosophical reflections in this book. This book is not an argument for a particular theory about opera but, rather, is a series of diverse philosophical reflections about opera across sixteen chapters. What unifies the collected pieces is a serious engagement with both the message and medium of opera, of the experiences it produces, and the capacities and resources as its disposal.

Hegel

Frederick Beiser, *Routledge*

&

Schopenhauer

Julian Young, *Routledge*

Andrew Stephenson

Cardiff University

stephenson@bups.org

The two books to be reviewed here are part of a major new series by Routledge, intended to be what one might call 'serious' introductions to some of the great thinkers of the Western philosophical tradition. The series as a whole is also the subject of this review. Already, other volumes in this series include, among others, *Hobbes*, by A. P. Martinich, *Leibniz*, by Nicholas Jolley, *Locke*, by E. J. Lowe, and *Rousseau*, by Nicholas Dent. Soon to be added are books on Spinoza, Kant, Husserl, Rawls, and several more. There is nothing original about the titles of this series – note that the series itself is called, descriptively

but not imaginatively, *Routledge Philosophers* – but what about the content of the books themselves and the concept of the series as a whole? Given the truly bewildering wealth of information readily and popularly available on all these thinkers, is this series justified? Well, it appears to me that against all odds, as regards both content and concept, there is plenty of room in the market for this series. The original contributions that the books in this series seem set to make to the *range* of literature available will prove invaluable to many an undergraduate. So, I need to explain why I think this. Why on earth do we need so many more introductions? First I will deal with mechanics and ideology of the concept of this series. Then I will go on to discuss the specific content of two of the books in this series, *Hegel* and *Schopenhauer*, the reading of which spurred me on to do this review.

Concept

The size of the tomb is, unsurprisingly, the first thing that strikes one upon confronting one of these books. I'm being a bit sensationalist – the contributions in this series certainly do not qualify for such ominous nomenclature. It's just that they are notably bigger than most if not all of the market contenders. *Hegel* runs to just over three-hundred-and-fifty pages, while *Schopenhauer* and *Leibniz* (the latter of which I am yet to read), both run to a relatively modest two-hundred-and-sixty or so. The other available volumes are closer in size to this, which suggests that those to come will be similar – it seems not even the concise and practiced exposition of Frederick Beiser can make *Hegel* easy to wield. The fact that even the slightest of these books is much larger in size than common introductions to lone philosophers could perhaps be inferred from the series' titles' lack of addenda. I think of such gems as, 'The Essential Joe Philosopher', 'Thinking Jane: a Very Short Introduction', or 'Fred Blogs in Ninety Minutes'.¹ The point is that

¹ The *Very Short Introduction* series, by the OUP, is really quite respectable and certainly not superficial, even if they are a little selective and brief. Titled topics range well outside the typical philosophy student's remit, which does seem to suggest that some compromise may have been made between popularity and scholarship. Moreover, many of the books are repackaged from the *Past Masters* series. This said, weighty-named contributors include A. J. Ayer, Michael Tanner, and Tom Sorell. The ... *in Ninety Minutes* series are early versions of *The Essential* ... series, authored entirely by Paul

while the books in this series give all the undergraduate-friendly biographical information, chronologies, and chapter summaries that one expects from an introduction, they go beyond the traditional introduction in both breadth and depth. The poor diligent student is not restricted to simply using them as preliminary reading for a course – you can reference them in essays! And yet in allowing this they stop short of the detailed critical analysis present in the text-based *Routledge Philosophy Guidebook to ...* series. Indeed, perhaps the immense popularity of this lasting sister series encouraged the commission of the series under discussion. Whatever the reason for the series' conception, the books in it fill the market gap between the slim volumes of many undergraduate-aimed introductions and the especially focused and sustained argument of the aforementioned *Routledge Guidebook to ...* series. And as such, an undergraduate looking to go that extra step should have plenty of use for them.

It should be clear, then, that the books are relatively uniform in length. But it cannot simply be assumed that, as a consequence, they are uniform in depth and breadth. There is simply (textually speaking at least) more philosophical material to cover when talking about Hegel than there is when talking about Hobbes. In general, however, you can expect a very good compromise between the three dimensions of essay writing: length, breadth, and depth. In this – the scope – and of course in the titles and aims of the books in the series, there has been successful guidance to ensure coherence. The same coherence cannot be claimed regarding structure, but this is not at all to the detriment of the series as a whole or the individual volumes contained therein. Julian Young's *Schopenhauer*, with a few notable exceptions, works in six chapters through the four books of the crazy-haired German's magnum opus, *The World as Will and Representation*. In contrast, Beiser's *Hegel* separates into five parts based on classic areas of philosophy such as metaphysics and epistemology, each themselves separated by Hegelian topics such as dialectics and spirit. The fundamental difference in

Strathern. Given that the series includes more than twelve books (and of course also given that each can indeed be read in only ninety minutes), it is not surprising that treatment of the philosophers discussed therein is brief and superficial, with precious space being given over to anecdote and hyperbole. These are books for the interested layperson, and most certainly are not reference material.

structure is clearly necessary to the success of the books themselves, and in no way detracts from the uniformity of the series. The editor has rightly trusted to the expertise of the individual author regarding this issue. However, the one notable difference that was brought to my attention was in an area of the books that could be, and on this occasion certainly should have been, dictated by the editor: the index. The index in both *Hegel* and *Schopenhauer* is virtually useless, each running to less than four pages. The contents pages alone of both works runs to four pages as well. Surely there is stuff under the subtitles, stuff that should be indexed, thereby necessitating a larger index than contents page. On the contrary, the index in *Leibniz* runs to a full twenty pages, a much better ratio to the size of the book and its three pages of contents. An index is crucial to study, and a miniscule one is a very real and very practical problem for the undergraduate who wishes to use these books. Given that I have characterised the success of this series largely by its fulfilment of the silent demand of the undergraduate, this is a serious issue, and one which I sincerely hope will be sorted out in the second printings.

Finally, regarding less the mechanics of the concept of the series and more the concept itself, onto the general series editor, Brian Leiter. I primarily know Leiter as a philosopher who works on naturalist interpretations of the big names in the 'Continental' tradition, most notably Nietzsche, but also others including Heidegger and Marx. These naturalist interpretations, given their close affinity with modern science, could also be characterised as an analytic philosopher's interpretations of Continental philosophy. In which case, Leiter can be categorized under the wave of contemporary American philosophers who, hailing from the firmly Anglo-American analytic tradition, have judged it thoroughly worthwhile to tackle those oldies from the Continent – the most famous (and most notorious) of these being without doubt Richard Rorty. The analytic penchants of the editor are reflected in Young, while the naturalist penchants of the editor are reflected in Beiser. In the 'influence and legacy' section at the end of Young's exposition (a section common to all the books in the series in some form or other) early Wittgenstein receives as much attention as Nietzsche. In Beiser, Hegel's organicism is interpreted as central to all aspects of his philosophy, from absolute idealism to theory of the state.

I am not at all suggesting that the slants of these interpretations are biased or wrong. Indeed, I am very keen to promote thoughtful and informed dialogue between the traditions. I genuinely believe (and it is one of the few beliefs I manage to maintain throughout my philosophical progress) each tradition, both Continental and analytic, can learn hugely from the other. Nevertheless, the context of the interpretations should be noted, and with it the pitfalls and benefits of such a context. (It must also be noted that this cannot be taken to characterise every book in the series, the most notable exception being Samuel Freeman's *Rawls* – it would be nothing remarkable to view Rawls from an analytic standpoint).

Content: Hegel

And now, onto the content. It is sometimes difficult to engage critically with such texts as these on a general scale. This is because they often try to be objective, as far as is reasonable. As a consequence, they often say little that is particularly contestable. And this is as it should be. It is not the aim of an introductory text to present as fact or consensus a highly dubitable claim. Having said this it remains the case that certainly, when one reads any book, there will be particular interpretations and specific arguments the soundness of which one does not admit. And Beiser's book is no different in this respect. But this is not the forum to catalogue my pencilled marginal notes (indeed I am sure such a forum does not and should not exist). Instead I will draw out the main philosophic themes of his interpretation, many of which I am in complete agreement with, but with one very notable exception. (My job, by the way, is made all the more difficult by the inadequate index).

After some very informative social and cultural situation, Beiser gives a convincing account of Hegel's early ideals, the influences on his thought, and his first conception of the absolute. Contrary to Kaufmann and Lukács' readings, which it is suggested are rather idealised, Beiser admits finding in Hegel's early work influence from friends and contemporaries in the Romantic movement. The very themes – religion, love, and spirit – that dominate the early work betray the correctness of this appraisal. Beiser's account here is perhaps most persuasive because of its sheer commonsense plausibility – Schelling

and Holderlin where personal acquaintances of Hegel's. Yet Beiser also maintains that, while we cannot understand the development of Hegel's mature works without understanding their origins in a fractured and stressed relationship with Romanticism, nor are his mature works reducible to these origins. And here we move onto the most enduring and central aspect of Hegel's thought – his participation in German idealism and his metaphysics of absolute idealism. Hegel was, like Schelling and Fichte, akin to a grail knight on an eternal quest bequeathed unto them by a past master. He wished to gain knowledge of *das ding an sich* – Kant's greatest legacy, 'the thing-in-itself'. Of this absolute idealism Beiser gives a clear account which is surprisingly logical (rather than mystical). He also draws interesting parallels with Spinoza and the complete modal collapse Spinoza's philosophy engendered. Beiser argues that Hegel's early notion of the absolute led him into the dark night of panlogicism, where all truths are necessary and contingencies cannot be accounted for.

It is here that Beiser unveils his belief in the centrality of organicism to Hegel's mature thought, and this is something I am convinced by. That is not to say that I find the outmoded metaphor of the world as an organism compelling, but rather that there is value in Beiser's subtextual claim that the quaintness of the idea can be discarded to leave us with the key Hegelian notion of synthesis. The key to understanding Hegel's enormously complex system – including his metaphysics, epistemology, political and ethical philosophy – is the notion of synthesis. At this point Beiser makes several rather ill advised asides. These come in the form of some interesting criticisms of Kant's restriction of knowledge to *transcendental* rather than transcendent, an odd section on spirit and love, and a section on the religious dimension that waxes and wanes throughout Hegel's development. Only then does an account of the conceptual framework within which the key notion of synthesis functions. Hegel's dialectic is the foundation of his metaphysics, and since (Beiser argues, contrary to Foucault) his metaphysics is the foundation of his epistemology and politics, the dialectic is the foundation of Hegel's entire grandiose system. And, of course, the notion of synthesis is the key concept of the dialectic. Since I am writing a review of an introduction rather than the introduction itself, I can excuse myself from explicating these difficult ideas, to which

many trained analytic philosophers may with good reason feel a resistance. I will simply note that Beiser quickly dismisses a myth spawned by the now common English translations of the three key moments in Hegel's dialectic: thesis, antithesis, synthesis. The logic of the progression that leads to the temporarily final moment of *aufheben*, quite literally 'heaving-up', is far more complex and difficult to apply than those popular terms suggest.

I have tried to indicate that Beiser gives us a unified account of Hegel's thought that brings together not only the *Phenomenology of Spirit* and the *Science of Logic*, but also the *Philosophy of Right*, the *Encyclopaedias*, and the lectures on religion and aesthetics. Needless to say there are points of extreme stress in this encompassing interpretation, most notable I think is his account of Hegel's conception of God as an entirely immanent conception. It seems that this reading is positively required by two of Beiser's central principles: unity, and the centrality of organicism. The problem I have with such a reading is that it seems entirely at odds with Hegel's language of transcendence when speaking of God, and indeed the absolute. I am tempted to credit it to the compulsion toward systemization which Beiser's sustained argument sometimes betrays. But, really, this is only a benefit from the point of view of the undergraduate – Beiser's text offers us its jugular and dares us to bite. Therefore I cannot recommend this book more highly to the undergraduate about to study Hegel, whether that student expects to be converted or affronted. It is not perhaps so suitable for the interested bystander, to whom this book may well prove a death blow. But having said that, the same can be said of the analytic philosopher, to whom I would nevertheless mischievously recommend it – mischievous because it might kill them, and mischievous because it might convert them, or at least sow the seeds of corruption.

Content: Schopenhauer

To Young I am not so generous, and this is for two reasons: the author, and the author's subject, Schopenhauer. And yet from saying I am not so generous, it does not follow that I do not see great value in this book, only that I see more faults. In the interest of this review, it is on those I will briefly dwell. The way Young approaches his reading of

Schopenhauer's dualistic metaphysics of will and representation is to argue for a reversal in the later Schopenhauer's thinking, and in this I acquiesce. Just like Hegel with the absolute, Schopenhauer claimed to be able to account for knowledge of the thing-in-itself with his concept of *Wille* – quite simply 'will'. By the time he came to amend and publish the second edition of his masterpiece, *The World as Will and Representation*, Schopenhauer had a lot of new things to say. He added to each of the four books a supplement, together amounting to more words than the original book. Young argues that this so-called second part of *The World as Will and Representation* gives us a complete reversal. No longer does the will correspond to the Kantian noumenal object, rather both will and representation reside in the realm of the phenomenal.

In general this is fine, and before I go on to present a case study in disagreement, I should note that the two chapters on Schopenhauer's philosophy art in Young's book are both excellent. They really do justice to the wealth of material in Schopenhauer's aesthetics, although they do perhaps neglect the massive debt he owes to Kant. Now, the praise over with, there is a section early in Young's book where I think he is particularly dismissive of what I find to be an intuitive observation made by Schopenhauer and contradicted by Wittgenstein. (The reason I choose to focus on this particular point of my disagreement with Young is twofold: it highlights the limitations of an underlying concept of the series – the worth of the dialogue between the analytic and Continental traditions; and it restricts the valid extrapolations one is allowed to make from the common fact that Wittgenstein, strangely, read and was influenced by Schopenhauer throughout his life). In the small part of the *Tractatus* where Wittgenstein deals with solipsism, 5.6-5.641, he offers a diagram to show how perception is not. I speak of the famous circle within and to the edge of an oval. I will try to represent this diagram symbolically in order to liken it to a diagram Young also offers us to reject. Wittgenstein's and Young's positions are similar and contrary to. I think Schopenhauer is right, and I certainly don't think Young justifies his immediate adoption of the Wittgensteinian stance. Again my point is twofold: to show that Young's dismissal is unjustified, and to show that his close alignment of Schopenhauer with Wittgenstein is flawed.

Let S represent the subject of perception, a person like you or me. Let O represent the object or objects of perception, a tree or a cat. Let square brackets define the domain of experience for S, while soft brackets simply help with clarity.

For Wittgenstein and Young, it is both conceivable and correct for the perception to be like so:

a. S [O]

The domain of S's experience is only that outside of S, namely O. Experience is not had of internal episodes, and nor is awareness of S (self-awareness, as it were) necessary to awareness of O. Clearly much of Wittgenstein's time was taken up with arguing for this position, with the private language argument and his critique of the role of internal episodes for example. Young offers no such support, only a picture of a stick man imagining a tree in a thought bubble.

As Young shows us with a different picture, this time of a stick man imagining a man imagining a tree in a thought bubble, Schopenhauer disagrees:

b. [S (O)]

This is not to disagree specifically with Wittgenstein's observation regarding his diagram that 'really you do *not* see the eye' (5.633), for this it is difficult to deny. But still there remains logical room to make the Kantianesque claim (in keeping with Schopenhauer's immense debt to Kant) that the subject of perception somehow necessarily informs the objects of perception; that is, the subject is always present in perception, so that (a.) is not possible. Kant and Schopenhauer both argue this position from the intuition that it is impossible to conceive of a tree, say, without being aware of a subject who is conceiving of the tree. It is not within the remit of this review to settle the issue or even argue for either side. But whatever you decide in this debate, it seems certain to me that there is at least a debate to be had, and this it seems is what Young fails to acknowledge.

It seems, then, that even sizable introductions like these must make assumptions which will displease some if not many. Nevertheless, the particular depth of Young's book is ideal for the student of Schopenhauer, as much of the available literature is either too brief – the tendency being to see Schopenhauer as merely an egotistical, extravagant Kant – or too detailed and scholarly, at least for the newcomer. I hope it is clear that, as a resource for undergraduates, I have a lot of time for this series, and specifically for the two books I have focused on. Indeed, my enthusiasm is probably manifest in the length this strange review has grown to. For that reason I will now stop and take no more of your time – you're really going to need it if you plan to tackle the whole of this series.

Upcoming BUPS events

Philosophy is of course much, much better if you're with people who are passionate about the subject and know what they're talking about. BUPS exists to bring together undergrads who love philosophy. Our events offer opportunities to give or discuss really great papers, to meet and mix with other undergrads who think worrying about ethics or the fundamental structure of mind and world is kinda cool. To build an understanding of how philosophy is done across the country. To meet other students who like this stuff as much as you do, have done their reading and want to talk. We also organise the UK's only big, annual national undergraduate philosophy conference.

Interested?

Good. You should be at the events listed over the page then! You can see a typical programme or download a past conference report at our website – **www.bups.org**. If you're not already on the BUPS-L mailing list for announcements, you can also subscribe through the site. Don't worry – BUPS membership is free and our conferences are all tailored to fit a student budget. Submit a paper or come along when you can – we'd love to meet you!

Latest details of all our activities, profiles of the committee and an updated list of upcoming events are always available at: www.bups.org

Any enquiries can be addressed to: info@bups.org

**BUPS
Philosophy
Skills
Conference 2007**

*St. John's College, University of Durham
30th March-1st April*

Keynote by Professor Robert Hale (Sheffield)

Excellent philosophy papers by undergraduates

Skills development sessions

Outstanding papers to be offered publication in the BJUP

Friendly, in-depth discussion and great social events

Attendance with Bed & Breakfast - £58

Attendance without B&B - £18

Book now to avoid disappointment.

www.bups.org

Subscribing and submitting papers to the BJUP

BJUP subscriptions

The BJUP is the English-speaking world's only national undergraduate philosophy journal. We publish the best papers from BUPS' conferences, but also accept high-quality essays by direct submission.

Our non-profit status keeps the cost of subscription to our print version down, and all BUPS members receive the electronic version of the journal for free. New issues go out quarterly. We offer three levels of subscription:

BUPS member subscription (electronic)

Becoming a member of BUPS is really, really easy – all you need to do is join the BUPS-L mailing list. The electronic version of the journal is distributed to all BUPS members. We hope you enjoy it!

Individual subscription (print)

An annual subscription to the print version of the journal costs £40 in the UK.

Institutional subscription (print + electronic)

Institutions (libraries, schools, universities) wishing to subscribe to the journal receive both a print copy and a personalised electronic copy licensed for unlimited distribution to, and printing by, current students of the institution. This package costs £60 per year for UK delivery, slightly more for overseas postage.

Subscriptions run for a single academic year, a current subscription covering the print version of issues 1(1)–1(4). Full details of how to subscribe, and methods of payment we accept are available at the journal's webpage:

www.bups.org/BJUP

Submitting a paper to the BJUP

Most papers we publish will be 2,000 – 2,500 words in length. However we will consider papers of any length. We would suggest that you limit your submission to a maximum of 5,000 words, though, since papers longer than this are often better dealt with as a series of shorter, tighter, more focused essays.

What we're looking for in papers that we publish is actually quite simple. We like work that is:

- carefully structured
- argumentative rather than merely descriptive
- clearly written
- knowledgeable about a given subject area
- offering a new argument or point of view
- not just written for area specialists

As a general tip, don't write with 'This is for a journal, I must be technical, formal and use lots of jargon to show I know my subject...' running through your mind. Explanation to others who may not have read the same authors as you, clear laying out of thoughts and a good, well-worked-out and -offered argument that says something a bit different and interesting: these are the key characteristics of the best papers we've received. Don't be afraid to tackle difficult or technical subjects – we're all keen philosophers here – but do so as carefully and clearly as possible and you have a much better chance of being published.

Most of our papers are analytic, but we are delighted to accept and publish good papers in both the analytic and continental traditions.

We accept papers electronically as Microsoft Word .DOC or Adobe Acrobat .PDF files. If you have problems sending in these formats, please contact us and we will try to find another mutually acceptable file format.

Papers should be submitted via email to **bjup@bups.org** and should be prepared for blind review with a separate cover sheet giving name, affiliation, contact details and paper title.

Don't worry about following the journal's house style before submission. The only requirement we have in advance is that you follow English spelling conventions. Any other requirements will be made clear if your paper is accepted for publication.

Please do not submit papers for a BUPS conference and the journal at the same time. We'll make suggestions for rewriting or restructuring papers we think could be publishable with a bit of work. Please do not re-submit a particular paper if it has been rejected for a BUPS conference or the BJUP.

Reviewing papers fairly is a difficult and time-consuming job – please give us a couple of weeks and do not submit your paper elsewhere in the meantime.

We run the journal on the minimum copyright requirements possible. By submitting work you license BUPS and the BJUP to publish your work in the print and electronic versions of our journal, and agree to credit the journal as the original point of publication if the paper is later published as part of a collection or book. That's all – you are not giving us copyright over your work, or granting a licence to reprint your work in the future. We're philosophers not lawyers, so we hope that's pretty clear and fair...

(Anyone seen Rab recently?)