STOA 3.1
STOA: International Undergraduate Journal of Philosophy
published by:
The Center for Philosophical Education
Department of Philosophy
Santa Barbara City College
721 Cliff Drive, Santa Barbara, California 93109-2394
tele: 805.965.0581 ext. 2476  fax: 805.963.7222
e-mail: white@sbcc.net
website: http://www.CPEsbcc.net

Editors
James Chesher
Department of Philosophy, SBCC
Abrol Fairweather
University of San Francisco
Aleksandar Jokic
Portland State University
Joseph White
Department of Philosophy, SBCC
Nicholas Pitney
Student Editor/Department of Philosophy, UCSB

International Consulting Editors
Milos Arsenijevic
Belgrade University
Timothy Crane
University of London
Barbara Herman
University of California, Los Angeles
Nathan Salmon
University of California, Santa Barbara
Quentin Smith
Western Michigan University
Tim Williamson
Oxford University
Hans-Peter Schutt
Universitat Karlsruhe

Subscriptions to STOA should be sent to the Center for Philosophical Education in care of the Executive Director. STOA is printed semi-annually. Free bulk subscriptions are available to institutions. Individual rates are: US/Canada $25.00, International $30.00.

STOA is made possible by a generous grant from the Arthur N. Rupe Foundation
The Center for Philosophical Education

Contents:

A Word from the Editors
CPE: Previous STOAs & Conferences

Articles

Nick Tosh, “Science and Consciousness” 6
Mentor: Prof. Marina Frasca-Spada, Cambridge University, England

Abe Sprinsock, “Introspection and Knowledge of Mental Content” 19
Mentor: Prof. Alan Wallace, University of California, Santa Barbara

Monica Dus, “The Power of Intellect and Freedom” 25
Mentor: Prof. Tammy Nyden-Bullock, University of the Redlands

Richard Warfield, “Must We Save Only the Phenomena?” 37
Mentor: Prof. Daniel Rothbart, George Mason University

Thomas Cochran, “The Problem of Space” 45
Mentor: Prof. Tim Crane, University College London

~ Pacific University Undergraduate Philosophy Conferences ~

“Best Paper: Conference 2000”
Richard P. Brown, “James Sennet’s Epistemic Circularity and Soundness in Alvin Plantinga’s Ontological Argument” 62
George Fox University

“Honorable Mention: Conference 2000”
Ian Hesketh, “‘A Fish in Water’: Michel Foucault and Historical Change” 75
Okanagan University College, British Columbia, Canada

“Best Paper: Conference 2001”
Anthony Reeves, “Consciousness, Oppression and Resistance: A Hegelian Critique of Marx’s Assumption of a Proletarian Revolution” 87
University of Portland

“Honorable Mention: Conference 2001”
University of Washington

“Honorable Mention: Conference 2001”
Wendy Lynn Clark, “A Socratic Unity of the Virtues” 116
Lewis & Clark College
A Word from the Editors

Welcome to Volume 3 Issue 1 of STOA, the ever-growing international, undergraduate journal of philosophy. For those of you who have been subscribers to, and supporters of, STOA, we extend our deepest gratitude for your patience in waiting for this issue. We have been, quite fortunately, the victims of success. STOA now appears on five continents involving eleven countries. Please note, however, that STOA’s vitality remains dependent upon your support. If you know of colleagues and/or departments who are actively involved in supporting undergraduate education in philosophy, please let us know so that we may introduce them to STOA.

In this issue of STOA, you will find our first student papers from China and Turkey. Additionally, in this issue we have a new format – The Symposium. When Professor Keith Lehrer participated in our Virtue & Epistemology Conference, he suggested an intriguing possibility for STOA. What if he submitted a number of outstanding undergraduate papers from one of his classes along with his response to each of those papers? We smiled broadly. This marks an ideal arrangement for STOA as such a format would not only further reinforce the mentor relationship, which STOA is dedicated to nurturing, but it would substantially enhance the value of STOA as a learning tool for undergraduate education. So arose our first Symposium, The Lehrer Symposium. We are excited by the prospect of other faculty utilizing the Symposium format in STOA.

Finally, and at the heart of STOA’s sustenance, is our deep appreciation to the Arthur N. Rupe Foundation. Through the generosity and sincere love and respect for learning on the part of Mr. Arthur N. Rupe, STOA is able to continue thriving. While most libraries and departments remain strapped for funds in trying to maintain meager collections of professional journals, something as unique as an undergraduate journal, while still in its infancy, gets little attention, let alone a share of limited budgets. While institutional and department subscriptions to STOA have been financially difficult for many, for others, as we have been told, in their part of the world, they lack adequate funding for textbooks. Journal subscriptions are simply not possible for them. So, given the overwhelming critical praise and support for STOA, we are now able to provide institutions with free subscriptions. We are deeply appreciative to those departments who have paid for subscriptions and we would hope that those capable of providing financial support for our efforts would do so. We are presently attempting to build an endowment to secure the future of STOA so that this growing international journal will not be held hostage to the traditionally frugal finances of academia.
Again, thank all of you for your support. Please contact us with your suggestions and continue to send us your outstanding student papers. All previous STOA papers are on view at our website - CPEsbcc.net. Conference information, publication of conference proceedings, along with the table of contents of previous STOAs can be found at the back of this issue.

A word on style and formatting. The papers in STOA are not refereed in the traditional research journal sense. Rather, mentored papers are submitted by students and faculty members with the primary role of the STOA editors being that of seeing the papers and journal through the arduous proof stages. We genuinely hesitate to rework grammar and respect the many styles of formatting which are sent to us. Thus, a lack of consistency in format styles within STOA reflects the genuine diversity of styles found presently in the different cultures on planet earth. STOA is an international, undergraduate journal.
Science and Consciousness

Nick Tosh
Cambridge University

1. Introduction

The twentieth century has been characterised by an exhilarating pace of scientific progress, and we have, as a species, become enormously more confident in our knowledge of the world. I think few philosophers would dispute that the effect on their subject has been broadly positive. Certainly it has bred epistemic optimism. While John Locke believed the nature of solidity (say) would forever remain a mystery, bets against science are now no longer a popular intellectual investment. Nevertheless, I want to sound a note of pessimism, and I believe I can do so without extending a Lockean neck to physicists.

It is the claim of this essay that the most interesting philosophical problems clustering around the issue of consciousness will never be illuminated by science. The impediment is not a dearth of creativity, or limited computing power, or shrinking research budgets: as I view it, the impediment is a demonstrable epistemic impossibility.

If the issue is as logically clear-cut as that, why is it that modern philosophical orthodoxy generally takes the opposite view? I regard that as a sociological question, because I don’t believe the dominant influence here is philosophical argument. I think it is science’s cultural status as our paradigm of rational intellectual endeavour. While that paradigm has been highly energising for much of philosophy, in the case of consciousness it has imposed a blindspot.

I will develop and defend these claims in three stages. I will begin by presenting a general philosophical argument. This will lead me to consider some (relatively) modern work on consciousness. I have selected papers by Thomas Nagel and Colin McGinn which I think best demonstrate the problems with the contemporary approach. Finally, I will rewind the clock a couple of centuries and consider the idealism of George Berkeley. Formulated before the present era of scientific hegemony, I believe it provides the basis for a philosophical framework within which both science and consciousness can comfortably coexist.
2. The Problem of Other Minds

If we want to discuss ‘consciousness’ without cloaking it in paradox and mystery, we have to recognise two fundamentally different meanings of the word:

1. The subjective sense experienced by the self (‘I am conscious’);
2. The objective sense applied to others, on the basis of certain qualifying behavioural traits (‘He is conscious’).

Meaning 1 is the consciousness of Descartes’ *cogito*. That I am myself conscious in this sense is perhaps the most certain knowledge I possess.¹ Now the *cogito* obviously will not work in the third person, but I would still like to claim that my brother is conscious, and that my desk is not, without being drawn into a metaphysical dispute. To do so, it seems that I must restrict myself to 2. A thorough analysis of the phenomenological content of type-2 assertions is not relevant to the aims of this essay. One (trivial) fact will suffice: when we state ‘X is conscious and Y is not’, then so long as neither X nor Y refers to the self, we make the distinction on the basis of objectively observable properties alone.

Of course it is often the case that while the predicate ‘consciousness’ is assigned to others on purely empirical grounds, what is meant by the word includes a substantial dose of 1. The ‘problem of other minds’ is to determine how far such a claim can be justified. My position is that it cannot be justified at all. I do not think this should be regarded as a ‘problem’, but merely a fact. After all, an intuitively irresistible belief that defies all attempts at rational justification is hardly a novel philosophical object (consider, say, our trust in the validity of scientific induction).

If I were a 1930s logical positivist, I would defend this position as follows: “Type-1 consciousness is, even in principle, unobservable in objects other than the self. Therefore propositions about its presence or absence in such objects are literally nonsensical. QED.”² A vicious, blunt, bludgeon of an argument, but one for which I must confess substantial sympathy, even in that delinquent form. Nevertheless, positivism has been abandoned, and with good reason. Physicists make frequent and uncontroversial use of unobservable objects in their explanations (quarks, virtual particles, space-time...), and philosophy simply does not have the credibility to pull intellectual rank on science and declare whole swathes of empirically successful theory meaningless. (In fact, strict logical positivism fares little better with everyday objects – who could
offer a rigorous phenomenological reduction of the word ‘chair’?).

The crucial point to note about science is that when unobservable entities do crop up, it is they that do the explaining. If we can compress our empirical knowledge into a tighter, more manageable pattern by inventing an unobservable, then good for us, and good for science. What you will not find in physics is an unobservable explanandum. Such a creature would look ridiculous in any physical theory – scientists do not invent new objects purely for the explanatory challenge their existence would pose (though notice how easy this game would be to play: ‘an invisible spaceship that can go faster than light’ etc). Either there is direct empirical evidence for an object, or the supposition of that object’s existence has sufficient explanatory power to compensate for its own ontological weight.³

Now that we have relaxed the positivist constraints, does type-1 consciousness in others escape the brutal treatment it received above? Unfortunately not, for the concept can add ‘non-explanatory’ to its list of epistemological sins, and a non-explanatory unobservable cannot be employed in a scientifically meaningful context.

On first reading, that claim may seem quite wrong. ‘Surely,’ it might be argued, ‘we find the assumption of 1 in our fellow human beings to be an extremely powerful, prediction-supporting explanation for their behaviour. If I want to anticipate what somebody will say or do, I imagine experiencing the world from their subjective point of view. I try to think their private thoughts, feel their private feelings, and finally guess at their public actions on that basis. And I’m often right.’ What is described here could be called the ‘empathic method’ of behaviour-prediction. It is of course of vital social importance, but in no way does it assume the existence of 1 in the individuals to which it is applied. All we assume when we use the method is that it will work, and this we establish by plain scientific induction. Whether or not the empathic method is effective in some particular case is an empirical fact, and no doubt one of the key criteria for testing the presence or absence of type-2 consciousness, but it need not carry any metaphysical baggage.

Any scientific theory of consciousness will be limited to considerations of 2. It must be non-informative about 1, because, quite simply, 1 has no content with which science can engage. This is true as a matter of principle and so is independent of the power of the scientific theory, even in the (probably impossible) limit of perfect predictive success. Were we to develop a model that reliably predicted the actions and utterances of any individual, down to the slightest gesture, tiniest vocal tone, and most fleeting neuron firing pattern,
then science would have as complete an understanding of the brain as any of its practitioners could possibly hope for. The research field would be closed. Yet even this super-theory does no more than provide a mapping from objectively observable data to objectively verifiable predictions; logically, therefore, it must be equally compatible with 1 and not-1. Nevertheless, 1 is absolutely certain for the self. Science can never capture that Self/Other asymmetry, and since the mind-body problem can only be meaningfully raised for the self, science is immediately disqualified from attempting it. This implies no restriction on the number and range of fascinating ‘body-body’ questions science can answer, but philosophically interesting questions about consciousness are immune to scientific assault. If we want to give answers at all, we will have to be satisfied with constructions that are judged on emotional criteria.

Such pessimism with regards to other minds perhaps merits a brief moral disclaimer. Insofar as we do behave ethically towards each other, we do so on the basis of empirical relationships; certainly we have no privileged access to the putative ‘inner worlds’ of others. My position has no empirical content, indeed, its essence lies in recognising that very absence. It therefore has no implications about what we should expect from others, or how we should act towards them. This point is vividly exemplified by my frequent use of the first person plural. I expect my reader to appreciate and assert the truth of the cogito even while I explain why I can never have rational justification for attributing type-1 consciousness to him. The cogito has the unique property that its truth, even its meaningfulness, is genuinely subjective. Finally, at no point have I concluded that other humans are not 1-conscious; my claim is that the issue can never be settled, one way or the other, in a rational argument. We might conclude that the question is meaningless. Alternatively, we might say that the truth value is a matter of choice, and choose ‘true’ because we prefer it that way. But if we do set a truth value, we must be careful to avoid vesting it with the status of a scientific fact.

3. Late Twentieth Century Philosophy of Mind

There is a tendency amongst modern philosophers of mind to produce speculative anticipations of natural science, rather than philosophy per se. Philosophy-as-conceptual-analysis has to some extent given way to philosophy-as-testable-hypothesis-formation. My intention in pointing this out is emphatically not to register a protest. On the contrary, my personal opinion is that guiding and stimulating science’s research program is one of the most laudable activities a philosopher can engage in. For example, Daniel Dennett’s fascinating conjectural work on the influences genetic and cultural evolution
have had on the modern adult brain may well end up catalysing ‘hard’ laboratory science (Dennett, 1991). That said, I do feel the provocative title of Dennett’s book *Consciousness Explained* is vulnerable to the arguments I outlined in section 2: I want to insert the prefix ‘2-’. The mind-body problem is immune to Dennett’s brand of philosophy for the same reason that it is immune to empirical science.

Now, the existence of philosophy-as-scientific-vanguard is an interesting cultural development, but I do not intend to explore its content any further. More relevant to the aims of this essay is philosophy of mind in the traditional sense. In recent years, this field has witnessed something of a backlash against the ‘materialism’ and ‘reductionism’ of the pro-science lobby. Philosophy had urgent need of such heresy. However, it seems to me that philosophy’s love affair with science debilitated even the self-confessed sceptics.

3.1 “What is it Like to Be a Bat?”, Thomas Nagel, 1974

Thomas Nagel’s famous paper begins with a deservedly harsh rejection of naïve mental reductionism:

Consciousness is what makes the mind-body problem really intractable. Perhaps that is why current discussions of the problem give it little attention or get it obviously wrong. The recent wave of reductionist euphoria has produced several analyses of mental phenomena and mental concepts designed to explain the possibility of some variety of materialism, psycho-physical identification, or reduction. But the problems dealt with are those common to this type of reduction and other types, and what makes the mind-body problem unique, and unlike the water-H₂O problem or the Turing machine-IBM machine problem...is ignored. (Nagel, 1974).

Nagel seems to be drawing the same distinction between the philosophical and scientific problems that I emphasised above. One might therefore presume that he uses the word ‘consciousness’ in sense 1 – the concept which lies at the core of the mind-body problem. The “reductionist euphoria” must forever be confined to 2. But then he continues:

“Conscious experience is a widespread phenomenon. It occurs at many levels of animal life, though we cannot be sure of its presence in the simpler organisms, and it is very difficult to say in general what provides evidence of it.”

On my analysis of consciousness, Nagel’s position is now problematic. The above quote must use ‘consciousness’ in sense 2 (else the claims would be
unsubstantiable assumptions), but on that definition the ‘mind’-body problem
is not unique, and Nagel’s opening salvo no longer applies. If I am correct that
1 and 2 are logically disjoint concepts, then Nagel’s decision to label both with
the word ‘consciousness’ carries a substantial risk: linguistically appealing argu-
ments may turn out to be fallacious.

In fact, Nagel soon tightens up his vocabulary: “Fundamentally an organ-
ism has conscious mental states if and only if there is something it is like to be
that organism – something it is like for the organism.”

The concept defined here has empirical content only if the organism in
question is the self. So Nagel is plumped for 1.

Offering an unambiguous definition of consciousness is an act of com-
mendable philosophical courage, but in so doing Nagel has greatly limited his
room for maneuver. In particular, his assertion that “conscious experience is a
widespread phenomenon” can no longer be defended – or rather, it can no
longer be defended on a truth-functional basis. One might wish to say that the
statement has a ‘choice-value’, in which case Nagel seems to be choosing ‘true’.

The problem is that this choice spawns a mind-body problem for just about
every creature in the biosphere. From Nagel’s point of view, science’s failure
to engage several billion instances of the same general problem – what it is
like to be this bat, that bat, this sheep, that person… – is a shocking indictment
of the collaborating philosophers he pillories in his opening paragraph.

But in the approach I am advocating, statements with ‘choice-value’ free-
dom are best regarded as meaningless. In that case, a meaningful mind-body
problem arises only for the self, and (if we persist in regarding it as a problem)
will resist any rational approach. A fortiori, Nagel’s claim that the current re-
ductionist paradigm is not going to solve it must be true – but it is hardly a fair
complaint against the paradigm.

Although Nagel would probably object to accusations of meaninglessness,
he is well aware that ‘facts’ about what it is like to be an X are epistemically
problematic. But the defence he offers, while emotionally compelling, ulti-
mately amounts to little more than a declaration of political correctness with
respect to other minds. Revealing sentences include:

• “We can be compelled to recognise the existence of such facts without being
able to state or comprehend them.
• “(If we can neither state nor comprehend them, what exactly is it about them
that compels us to recognise their existence? Is it a moral compulsion?).”
And:
• “To deny the reality or logical significance of what we can never describe or
understand is the crudest form of cognitive dissonance.”

When we complete the list of epistemological barriers begun by Nagel, we face the prospect of being permanently unable to describe, understand, observe, or derive explanatory power from what it is “like to be a bat.” Denying the logical significance of such a concept seems more like cognitive tidiness than cognitive dissonance.

I suggested above that it would be unfair of Nagel to attack the scientific-reductionist approach to mentality. Perhaps that assertion needs qualification. Nagel could score points against philosophers who speculate that the mind-body problem might be soluble by empirical science. There is no inconsistency in maintaining that while empirical science is the best tool we’ll ever have for understanding the brain, the mind-body problem is, as a matter of logic, not amenable to it. (That, in fact, is my position). Nagel hints at this argument in the opening of his paper, but later distances himself from it. That latter maneuver is vital to the structure of Nagel’s paper, because he ends up recasting the original philosophical problem in a way which seems to me quite radical:

I am not advertiting here to the alleged privacy of experience to its possessor. The point of view in question is not one accessible only to a single individual. Rather it is a type. It is often possible to take up a point of view other than one’s own… One person can know or say of another what the quality of the other’s experience is…[but] the more different from oneself the other experiencer is, the less success one can expect with this enterprise.

Nagel’s earlier definition of consciousness was unambiguously type-1. This account of the content of conscious experience is unambiguously type-2: he is describing what I have termed the empathic method of behaviour prediction. We have already ruled out its relevance to metaphysics – it can never support a type-1 judgement about another mind. Questions about the subjective character of a bat’s experience are no more (and no less) impenetrable than similar questions about human beings other than the self.

Nagel is of course free to consider both type-1 and type-2 questions about consciousness, so long as he keeps the logical strands separate. What worries me about the above quote is that the first sentence seems retrospectively to redefine all previous references to mentality as type-2 assertions. Much of Nagel’s argument collapses if the 1/2 distinction is no longer available to him
including his rejection of “reductionist euphoria” in the philosophy of mind.

So how should we approach the mind-body problem? Nagel’s parting words on the subject suggest that scientists and philosophers adopt a revamped introspectionism, and attempt to develop a phenomenological vocabulary “that could be used to explain to a person blind from birth what it was like to see.” In Nagel’s opinion, “a phenomenology that is in this sense objective may permit questions about the physical basis of experience to assume a more intelligible form... Otherwise we cannot even pose the mind-body problem without side-stepping it.”

Reports of experience (‘person X said P at time t’) are empirical facts and so qualify as acceptable feedstock for reductionist science, for the same reason that data about neuron-firing qualify. They are just as useless for closing the 1/2 gap, and just as irrelevant to the mind-body problem.

Nagel opened his piece by complaining that fellow philosophers were too enamoured with science to engage the mind-body problem. But Nagel is himself unwilling to deny the relevance of empirical science: in fact, he concludes the paper by endorsing his favourite flavour.

3.2 “Can We Solve the Mind-Body Problem?”, Colin McGinn, 1989

McGinn’s position is the following:

(i) There is no philosophical mind-body problem. A true scientific theory exists which gives a complete physical explanation of mind.

(ii) However, humans are “cognitively closed” to that theory. We will never discover it, nor even understand how its existence might be possible.

McGinn’s basis for (ii) is strong. He correctly notes that (type-1) consciousness is inaccessible to empirical science:

You can stare into a living conscious brain, your own or someone else’s, and see there a wide variety of instantiated properties – its shape, colour, texture etc. – but you will not thereby see what the subject is experiencing, the conscious state itself. Conscious states are simply not potential objects of perception. (McGinn, 1989).

Or, using my terminology, no amount of type-2 data warrants type-1 claims. A cameo appearance by Nagel’s celebrated bat stiffens the argument: “The true psychophysical theory would seem to provide a route to a grasp of the subjective character of the bat’s experiences.”
Since we can’t grasp that subjective character, McGinn argues, we can’t discover the theory.

As I see it (see section 2 above), these arguments demonstrate that a scientific solution of the mind-body problem is an incoherent concept. That we can never discover such a theory follows as a logical consequence, so I naturally must accept (ii). But Colin McGinn starts from (ii), and by shifting the weight of that proposition off the hypothetical theory and onto us, he can just about maintain (i) without outright logical contradiction. This is a very strange philosophical maneuver, and we might be forgiven for expecting something pretty special – certainly more than mere logical possibility – to back it up. Here is McGinn’s defence:

Resolutely shunning the supernatural, I think it is undeniable that it must be in virtue of some natural property of the brain that organisms are conscious. There just has to be some explanation for how brains subserve minds. If we are not to be eliminativists about consciousness, then some theory must exist which accounts for the psychophysical correlations we observe.

That is not an argument: it is a declaration of allegiance to the scientific-reductionist paradigm. Yet McGinn has just offered convincing arguments against science’s applicability to problems of consciousness. And I am genuinely puzzled how McGinn’s reference to “psychophysical correlations we observe” squares with his previous statement: “Conscious states are simply not potential objects of perception.”

I think McGinn is using the concept of “cognitive closure” as a metaphysical license. By definition, if we are cognitively closed to a type of theory, then an author writing about the possible content of that theory type has no epistemic leash to restrain him. He can go where he likes. McGinn wants to use that freedom to defend the explanatory completeness of science – despite the fact that his more logically sound arguments pull in the opposite direction.

4. Historical Contrast: The Philosophy of George Berkeley (1685-1753)

Eighteenth century ‘science’ was a humbler tradition than its mighty modern successor. ‘Natural philosophy’ and ‘moral philosophy’ were close in name, intellectual status, and aim. It was natural that those seeking to advance human reason should span the divide, and George Berkeley was no exception. His pioneering work on the psychology of vision (Berkeley, 1709) was founded on thorough experimental investigation, and in his own way, Berkeley was as passionate an advocate of science as Dennett or McGinn. Yet many today would
regard his idealist moral philosophy as a direct challenge to the intellectual authority of science. At a superficial level, it is not hard to see why. Berkeley’s target was nothing less than our concept of the external material world.

The ‘common sense’ metaphysics that Berkeley’s work attacks takes roughly the following form:

(i) There exists a world of matter which is independent of one’s perception of it.
(ii) Events occurring in the material world are observed if they correlate, via some kind of causal mechanism, with perceptual events in the mind.
(iii) Perceptual events in the mind are the only direct objects of experience.

Being an empiricist of firm conviction, Berkeley considers what sort of experiential evidence one might have for (i) and (ii). His conclusion is that none is possible, even in principle. To establish a correlation empirically, one is obliged to observe both sides of that correlation; only then can one justifiably claim that event type ‘A’ is consistently accompanied by event type ‘B’. But it is a logical implication of (ii) and (iii) that all one’s observations lie on the same side of the divide. Certainly one experiences all observations as perceptual events in the mind. One might be tempted to claim ‘Yes, but those are experiences of objects that exist independently of my mind,’ but of course that just begs the question. And it is plainly contradictory to speak of observing a ‘naked’ physical event that takes place without an associated perceptual event, because the act of observation is itself a perceptual event.

With all hope of establishing correlations such as (ii) eliminated, (i) is still available as a logical possibility, but it is one for which no evidence exists or indeed can exist. Furthermore, even if (i) is true, there is no way of legitimately inferring the properties of the external world to which it refers. The concept is thus rendered impotent – it is incapable of achieving any of the explanatory work typically expected of it.

Berkeley maintains that an absolute material world in the sense of (i) does not exist. Now strictly, the negation of a metaphysical judgement is itself a metaphysical judgement, and is thus just as unwarranted. It is safer to point out that the whole concept is vacuous, and must therefore be irrelevant to our actual explanatory behaviour. For all practical purposes that conclusion is almost as strong as Berkeley’s.

Notice that particular beliefs about the ‘material world’ are generally un-
objectionable under idealism. My belief that my bicycle is chained to railings in Trinity Street is meaningful because it has (contingent) experiential implications. Idealism is compatible with all possible empirical data, because the claims it rejects as meaningless are precisely those claims that are by definition unverifiable and devoid of explanatory power (e.g. ‘My bicycle has an existence that transcends all my bicycle-related experiences’). A switch to idealism may make certain difficulties in the philosophy of science seem particularly pressing – e.g. how should an individual justify his reliance on the remote testimony of others when forming his scientific beliefs? How does the concept of an objective fact become established through such testimony? – but in fact such questions are equally tricky from a materialistic stance. (And in neither case are they insurmountable.) Certainly idealism imposes no particular threat to the integrity of science.

Let us now assess Berkeley’s position in terms of the 1/2 (and more generally, the subjective/objective) divide that has featured so prominently in this essay. The sole element of ‘common sense’ metaphysics that Berkeley accepts is (iii). That means he is committed to constructing all knowledge from the subjective experiences of 1-consciousness. Science cannot be used to kick-start epistemology: rather, it must emerge as a construction of it. The same may be said of ‘objective’ knowledge in general (including knowledge of 2-consciousness).

Where does that leave the mind-body problem? In my opinion, it leaves it dissolved. Under idealism, 1-consciousness is not a feature of reality, it is reality. Explanation makes sense as a relationship between elements of experience. It is not coherent (and it is certainly circular) to ‘explain’ 1-consciousness – the fact that experience exists – in terms of subsets of that same experience. Every attempt to break the problem of consciousness by scientific assault is an attempt at an impossible epistemological summersault. Berkeley himself convincingly diagnosed the logical impediment in 1713, more than two centuries before modern neuroscience got started:

[A]ll that we know or conceive are our own ideas. When therefore you say, all ideas are occasioned by impressions in the brain, do you conceive this brain or no? If you do, then you talk of ideas imprinted in an idea, causing that same idea, which is absurd. If you do not conceive it, you talk unintelligibly, instead of forming a reasonable hypothesis. (Berkeley, 1713).

We showed earlier that the mind-body problem does not arise for other
brains. It now turns out that even for the self, the ‘problem’ of consciousness amounts to no more than the existence of an epistemological ground floor. Unless knowledge is expected to imply itself into existence *ex nihilo*, something of the sort must be accepted.

5. Conclusion

Philosophically interesting questions about consciousness do not survive translation into scientific vocabulary. The problem was at least glimpsed by Berkeley: science is not, in general, a suitable fundamental currency for philosophical discourse. It is a sophisticated, communally held, high-level epistemological construct, composed of vast bodies of contingent knowledge, and held together by multiple levels of induction, remote testimony, etc. Even the *concept* of an objective fact is non-trivial. (Presumably, elements of experience that are reflected in the linguistic utterances of others are regarded, in some sense, as ‘shared’ experiences and labelled ‘objective’). Precisely how we choose to analyse science is not relevant to the current argument. What counts is that analysis – in fact, reductive analysis – is in principle possible. If we chose not to take that step, if we always demand that ‘the buck stops’ with science, then we should *expect* to encounter insoluble philosophical mysteries when we face up to questions belonging to a deeper epistemological level. But our “cognitive closure” in such cases may be no more than a manifestation of our own obstinacy.

Perhaps our final words on the twentieth century should be taken from a member of the eighteenth:

“I am inclined to think that the far greater part, if not all, of those difficulties which have hitherto amused philosophers, and blocked up the way to knowledge, are entirely owing to ourselves – that we have first raised a dust and then complain we cannot see. (Berkeley, 1710).”

Nominated by: Prof. Marina Frasca-Spada
Cambridge University, England

End Notes

1 Strictly, the Cartesian *cogito* can only establish an instantaneous flash of consciousness – the ‘now.’ To establish any kind of temporal continuity, we must rely on memory for one direction and on induction for the other – both methods open to sceptical doubt. These *particular* sceptical doubts will not be pursued in this essay.

2 It is interesting that A. J. Ayer – who really *was* a 1930s positivist – did not accept this argument: he attempted to *refute* solipsism by defining consciousness
behaviourally. (Ayer, 1971, pp136-143) Having defined the issue out of existence, refutation was of course impossible, and Ayer himself expressed dissatisfaction with his account of other minds. (Ayer, 1971, pp189-190)

3 The more strictly one interprets ‘direct empirical evidence’, the greater is the proportion of scientific knowledge relying on the latter condition. In the limit of pure phenomenology, it might be argued that all scientific knowledge is explanatory.

Bibliography


Introspection and Knowledge of Mental Content

Abe Sprinsock
Religious Studies, UCSB/SBCC Alumnus

Introduction

William James declares that for the study of the mind, “Introspective Observation is what we have to rely on first and foremost and always” (James 185). Patricia Churchland, on the other hand, argues that introspection should play little or no role in the study of the mind, for such observations are invariably deceptive and unreliable. As a result of the investigation I will present in this paper, I reach the conclusion that James’ position is more plausible than that of Churchland. After reaching that conclusion, I will present a counter-argument (the second one of this paper), my rebuttal to which will form the basis for my own position, which differs from that of either James or Churchland.

Clarification of the Issue

The issue here is whether or not introspection is the most reliable method for studying the mind. I will address this issue within the context of the larger question “what is the most reliable method(s) for studying the mind?” In order to answer both of the above questions, we must first establish a criterion for reliability. The reliability or unreliability of any method of investigation can only be evaluated relative to the purpose to which it is applied. Some methods may be very unreliable for one purpose but very reliable for a different purpose. For example, if I wanted to ascertain the topics for the next paper in this class, and one of my classmates were to suggest that I keep a careful record of all the gossip and speculation among the students about what our next paper topics might be, I would consider such a method to be very unreliable. But if my goal were to find out what the gossip and speculation among the students was regarding our next paper topics, then the same method of investigation would be extremely reliable. In fact, it would be the most reliable method possible, for the most reliable source of information about a thing (e.g. gossip among the students) is the thing itself (e.g. gossip among the students). There-
fore, being able to correctly identify the thing itself becomes the first order of business before we can correctly determine what the most reliable method(s) for studying that thing are. And so it is with the mind.

**The Opposing Sides of the Issue**

Thus it is not surprising to discover that the disagreement between James and Churchland as to the most reliable method for studying the mind stems from a fundamental disagreement as to what the mind is. This disagreement can be expressed in terms of the above analogy: Churchland identifies the mind as being nothing other than the brain: “mind is, in fact, the brain” (Houshmand, et. al. 18). From this it follows that the most reliable source of information about the mind is the brain itself, and all the various subjective phenomena accessed through introspection are unreliable and misleading “gos- sip.” But James would claim that Churchland has misidentified the object of study. He would say that the subjective inner “gossip” accessed through introspection is exactly what the mind is: “All people unhesitatingly... distinguish the mental state as an inward activity or passion... I regard this belief as the most fundamental of all the postulates of Psychology” (James 185). In that case, that very same inner “gossip” becomes the most reliable source of information about the mind.

Thus we can see that in order to resolve our issue, we need to determine which of the above two ways of identifying our object of study (i.e. the mind) is more accurate. I will begin this task with a description of, and response to, Churchland’s way of identifying the mind.

**Churchland’s Way of Identifying the Mind**

Churchland considers the mind to be identical to the brain: “Indeed, the relation between psychological states and brain states is probably one of identity” (Houshmand, et. al. 25). This position is known as the mind/brain identity theory, a.k.a. “physicalism,” which is widely held in scientific and philosophical circles. If this theory were true, then every mental state without exception would be able to be explained in terms of the functioning of the brain. Thus, it is incumbent upon physicalists to provide us with a hypothetical model of brain function that could produce every known mental state. The model of brain function provided by Churchland (in agreement with most physicalists), is that of a computer: “I would argue that the brain is a kind of computer” (op. cit. 31).

**Objection to Churchland**

But there are quite a number of mental states that cannot be explained in terms of computational processes. These include emotion, the sensational quali-
ties known as qualia – the “what it’s like” to experience taste, color, etc., understanding, and awareness itself. Nor can anyone explain how computational processes could ever give rise to a first-person perspective, to a point of view, or to a will.

If it is not possible to explain all mental states in terms of any existing models of brain function, then it is also not possible to identify the brain as being the sum total of everything there is to mind. Therefore, it is not possible to identify the brain as the most reliable object of study for studying the mind.

**Counter-Argument #1**

It could be argued that the objections I have just raised are based on a limited understanding of the capacities of computation. Churchland would contend that we cannot limit our concepts of computation to the ones we currently have: “What kind of computer it [the brain] is, we don’t know yet…It is clearly not like the kind one has on a desk. It doesn’t have that sort of architecture, and it doesn’t use those principles” (Houshmand, et. al. 31). Churchland is optimistic about the attempts of science to understand the principles of computation upon which she claims the brain operates: “Currently, simple models of neural networks can perform interestingly complex computations. They seem to be on the right track as models of brain function” (op. cit. 31).

Thus, Churchland holds the view (common among physicalists) that once all the computational functions of the brain have been explained scientifically, all the subjective aspects of the mind (including all those mentioned above) will also have been explained in terms of one or a combination of those functions. At that point, it will have been proven conclusively that the study of the brain is the most reliable method for studying the mind, and it will have become obvious how unreliable the deceptive and misleading inner “gossip” accessed through introspection actually is.

**Rebuttal to Counter-Argument #1**

Churchland’s optimism for the prospect of evidence being assembled in the future to support her present position does not constitute proof of that position. Therefore, there is no basis for accepting her way of identifying the mind as being correct. Of course, she could counter that there is also no way to rule out the possibility that neuroscience will, at some point in the future, succeed in explaining all known mental states in terms of brain function.

But I claim that even if such explanations were to be provided, it still would not undermine the primacy of introspection as the most reliable method for studying the mind, and I offer the following thought experiment to support that claim: Suppose that at some point in the future, neuroscience succeeds in
explaining and mapping completely all known mental states in terms of brain function, and that on this basis it is able to develop instrumentation capable of monitoring and reporting the specific nature, content, duration, and intensity of all thoughts and feelings.

Now the question is this: By what means would it be possible to establish the viability of such monitoring procedures and test them for accuracy? The obvious answer is: Introspection. That is the necessary basis upon which such procedures could ever be developed, accepted, and tested. Why? Because introspection is the only means of gaining direct access to the subjective inner mental experience, which is the only possible standard against which such procedures, as well as any models or explanations of mind, can be judged. From this we can see that the subjective inner mental experience is the primary datum, the thing itself—the object of study we are looking for when we go about studying the mind. Now that we have clearly identified the object of study, we can confidently conclude that introspection, being the only means of gaining direct access to that object of study, is the most reliable method for studying the mind—the method that “we have to rely on first and foremost and always.” Therefore, the position of William James is the more plausible of the two positions we have herein considered.

Counter-Argument #2

The above argument is based on the proposition that the most reliable source of information about a thing is the thing itself. But this proposition can be shown to be untrue. We can find many cases in which the most reliable information about a thing can only be found outside the thing itself, and in which the thing itself is deceptive and unreliable. For example, if we wanted to know the nature of the image we see on a movie screen, we could not accomplish that goal by observing the image itself (to make this analogy more appropriate, let us assume that for this observation we are confined to the seat from which the average moviegoer watches a movie). Such an observation could mislead us toward a mistaken conclusion. To avoid being misled, we would have to become aware of the fact that there is a projection booth from which the image is being projected. At that point we would know that the image is actually a projection of light on a screen. For more complete knowledge of the nature of the image, we would have to go into the projection booth and study the mechanisms and processes therein, after which we would know that the image is actually a projection of light that takes on specific shapes, colors, and transformations by being passed through film consisting of individual frames that are driven (by a motor) through the light so rapidly as to convey the im-
pression of fluid motion.

In the above example, the image on the movie screen is an “epiphenomenon” — a byproduct of certain mechanisms and processes, a disruption which would alter the byproduct or cause it to vanish altogether. It is widely held that the mind is just such an epiphenomenon, whose nature is not in the least self-generating, self-dependent, or self-evident. Therefore, the use of introspection to observe the content of that phenomenon itself, just as observing the contents of the image on a movie screen, is deceptive and unreliable.

**Rebuttal to Counter Argument #2 and My Own Position**

The gist of counter-argument #2 is that epiphenomena are counter-examples to the proposition that the most reliable source of information about a thing is the thing itself, and that since the mind is an epiphenomenon, it is not the most reliable source of information about itself. I could challenge the validity of that argument, but instead I will take a different direction and challenge one of its premises – the proposition that the mind is an epiphenomenon. The considerations that arise in the process of this challenge will form the basis of my own position on our central issue, which differs from that of either James or Churchland.

An epiphenomenon is by necessity a passive entity – a mere appearance, which, by its very nature, can have no causal efficacy of its own. Indeed, one of the basic tenets of epiphenomenalism is that “consciousness does not affect the body but exists as a powerless mental state” (Angeles 88). Therefore, if it can be shown that the mind has any causal efficacy whatsoever, it will have been demonstrated that the mind cannot merely be an epiphenomenon.

It appears to have been known for a long time that the mind has causal efficacy upon the body. The myriad cases of psychosomatic illness, attitudinal healing, placebo effects, etc. have provided ample support for this view. In addition, meditators of various traditions have demonstrated (under laboratory conditions) the ability to significantly alter bodily phenomena such as temperature, heart rate, blood pressure, etc., at will. This by itself appears to disprove epiphenomenalism. And if recent studies purporting to verify the existence of psychokinesis (the alleged ability of the mind to move or affect external objects) were to be conclusively corroborated, epiphenomenalism would soon go the way of the dinosaur.

The above evidence points to the conclusion that the mind is more than just a powerless byproduct of brain function. Nor are the powers and potentialities it appears to possess evident through introspection (at least not the kind of introspection available to ordinary people). That is why I claim that neither
James’ nor Churchland’s position is correct. What exactly the mind is and how best to study what it is has not yet been established to my satisfaction. But if I were required to hazard a guess, I would guess that the mind is an energetic entity with causal efficacy of its own, not only within the body but perhaps outside of it as well, and perhaps connected in some way to macrocosmic causal principles. Therefore, although empirical and philosophical investigation into all its facets and capacities is doubtless important, the most reliable and authoritative methods for studying the mind may turn out to be the meditational techniques used by those who have demonstrated the ability to access and utilize its broadest range of capacities and potentialities. That is my tentative position on the issue.

Nominated by: Prof. Alan Wallace
University of California, Santa Barbara

Bibliography


The Power of Intellect and Freedom

Monica Dus
University of Redlands

I

He has his entire power: here he can fly! Why should he again go down into those muddy waters where he has to swim and wade his wings!-No! There it is too hard for us to live! We cannot help it that we are born for the atmosphere, the pure atmosphere, we rivals of the ray of light; and we should like best to ride like it on the atoms of ether, not away from the sun, but towards the sun!

Nietzsche, The Gay Science

Baruch Spinoza’s concept of freedom does not involve freedom from affections or desires. Neither does it consist of liberty of action, choice, or will. His freedom is something deeper, a concept deeply inherent in human nature, a concept inevitably connected with human happiness. Freedom, for Spinoza, is the inherent power to understand our essence and to self-determine our existence.

Many scholars have argued that there is no place for freedom in Spinoza’s philosophy; yet the philosopher himself proposes ways to achieve liberty. How to compare two opposite statements then? The key to answer this question lies in the word freedom itself: we ought to clearly define not only what is meant by it, but also its relations with men. In our society freedom is usually understood as the magnificent faculty of the mind to attribute judgments, as the powerful tool of the will to suspend, affirm, or deny ideas. This conception, deeply rooted in the Judeo-Christian tradition, has often been incarnated in the sacred notion of Free Will, the liberty to choose our destiny as opposed to predestination. Spinoza’s freedom does not involve the notion of free will. More than this: it cannot.

In Ethics, book II Spinoza affirms: “In the mind there is no absolute, or
The mind is determined to this or that volition by a cause, which is likewise determined by another cause, and this again by another, and so ad infinitum” (E II, 48).

This proposition does more than assert the non-existence of free will. It does more than equate the will and the intellect. First of all it must be pointed out that with will Spinoza intended the ability to affirm and deny, and not to desire. What he is negating here is not the faculty of men to assign judgments, but rather the existence of a separate entity outside or within the mind giving these judgments. For Spinoza, affirmation and denial are not separate from the idea itself. This sharply contrasts with Descartes’ philosophy, where intellect and will are two separate faculties of the mind. For Descartes the intellect is finite; it represents the faculty of knowledge and thus it only allows men to perceive ideas, without affirming or denying them. The infinite will, on the other hand, represents the faculty of choice, hence affirming or denying, avoiding or pursuing the ideas presented by the intellect. Thus for Descartes, ideas perceived by the intellect are subjected to the judgments of the will:

“This is because the will simply consists in our ability to do or not to do something (that is, to affirm or deny, to pursue or avoid); or rather, it consists simply in the fact that when the intellect puts something forward, we are moved to affirm or deny or to pursue or avoid it in such a way that we [do not feel ourselves to be determined by any external force.]” 2 (Med. IV, 57)

Spinoza does not distinguish between the two; will and intellect are not two distinct faculties of the mind, but the same. Hence, because the will, or “freedom of choice,” is not independent from the intellect, and the mind is “determined by another cause,” (see E II, 48, above) we derive that our mind is not free, but affected by external forces. But what does Spinoza really intend when he speaks of determination?

Unlike the conventional notion of determination as unfathomable predestination, Spinoza’s concept involves the notion of causality. The concept of causality among the modes of the universe is central to this philosophy. Spinoza believes that only substance, which can be understood within itself, 3 is not subjected to external causes (or it would not be substance). All the modes in Nature 4 are affected by other modes in an infinite nexus of things; all modes both act and are acted upon. Thus, when Spinoza says ‘determined’ he simply means having been caused, he merely attests to the affection of external causes rather than to something that necessarily had to occur. Although this might
seem trivial, affirming the connection and causal relation of an object with others is essential to his philosophy. We must not forget that this causality necessarily goes both ways: all modes are inherently passive and active at the same time, although they have different degrees of passivity and activity. Second, understanding the causes and determinations of an object means comprehending its essence, for modes are inevitably understood through their causes.\(^5\) More, it means attesting its existence, since only Nature’s existence follows from Her essence.\(^6\) In other words, Spinoza thinks that Nature is the unique, free cause of all things. Because all modes as modes must be conceived through Nature, they are inevitably dependent on her. In so far as Nature is considered absolutely (as substance) they necessarily follow from her, and are determined to exist (they cannot be conceived through themselves alone). Furthermore, modes are also determined to act in a particular way in so far as they are explained through Nature’s attributes. It then follows that nothing contained in Nature is contingent. All modes are affected by external causes (other modes) to exist and act in particular ways.\(^7\) Unlike in Descartes’ philosophy, modes cannot choose to be free or undetermined because perceived ideas (for the mind) are affirmed; they are affections. More than this; it is this very determination that allows them to exist. Without determination, there is no essence and therefore no existence. We now see how Spinoza cannot base freedom on the notion of Free Will. What is freedom then?

In Appendix to *Ethics* book I, Spinoza observes: “All men are born ignorant of the causes of things, and they all have the desire to seek their own advantage, a desire of which they are conscious.”

Because, as it was said before, understanding the essence of a mode inevitably means comprehending the causes of its existence, being unaware of the causes and connections of all things, ourselves included, also signifies being unaware of our own essence. More so, we are conscious of our desires but ignorant of the causes why we seek them. For Spinoza, there is no difference between desire and essence. Nonetheless, this desire is of a special kind: to preserve our own being. Spinoza says: “The conatus with which each thing endeavors to persist in its own being is nothing more but the actual essence of the thing itself” (E III, 7).

Now, every mode in Nature possesses a certain conatus of which the mind is conscious, and is necessarily driven to persevere in its own being. Because every mode that has an essence has a conatus, interactions and affections among modes will inevitably involve interactions among different conati. In addition, because every mode perseveres in its own being, it is possible that sometimes
these conati are opposite and of different degrees. This means that if a mode interacts with other modes possessing stronger and opposite conati, its own essence will be diminished. Because conatus is the power of a mode to maintain its essence, it is clear that in this case the checked mode is not active, but passive. On the other side, if the conatus increases, the mode will be inevitably active.

When applied to humans, this concept of conatus becomes even more important; it becomes virtue, and virtue seems to be the road to freedom. Ethics IV says:

“By virtue and power I mean the same thing; that is, virtue, in so far as it is related to man, is man’s very essence, or nature, in so far as he has power to bring about that which can be understood solely through the laws of his own nature.” (E IV, def 8)

In terms of activity and passivity, this identity between virtue and conatus means that, as long as a man seeks his own advantage, he will also increase his virtue, his power, so that when faced with external affections he will then be able to overcome them.

II

“True Freedom is to remain bound by the lovely chains of the love of God.”
Spinoza, Short Treatise on God, Man, and His Well-Being

Of Virtue and Reason

Spinoza thinks that human minds experience transient states, and are affected in many different ways by diverse modes. As we have seen, these affections can either diminish or increase the body’s power of activity, thus determining either pain or pleasure. Because the power of activity is nothing more than the strength of our conatus, modes checking our conatus will inevitably decrease our power. Moreover, according to Spinoza there cannot be anything within a mode causing the mode’s destruction. Because every mode affirms its own preservation, the mind will attempt to think only about “things that affirm its [power] of activity” (E III, 54), and will try to eliminate thoughts of affections checking its conatus. Although the mind will always endeavor to persist in its own being, however, the power connected with these endeavors will vary depending upon the ideas associated with them. Spinoza conceives
two types of ideas: adequate and inadequate. Adequate ideas are those that can be understood through themselves, while inadequate are confused and fragmentary ideas from which effects cannot be explicated through their causes alone. When the body is affected by external modes, the mind will either develop adequate or inadequate ideas on the causes of these affections. It is when the mind imagines inadequate ideas, when it cannot understand the true causes of those affections, that it is passive. On the other side, if the mind develops adequate ideas and preserves its own being according to the dictates of reason, it will be active. But what does this difference in activity signify? How is it correlated to freedom?

Spinoza says that freedom is our power of activity, and that bondage is nothing more than man’s lack of power to control emotions. Now, there are two different kinds of emotions, those deriving from desire, pleasure, and pain, and those deriving only from desire and pleasure. Spinoza affirms that while the former are related to the mind insofar as it is passive, the latter are related to the mind insofar as it is active. Because for the mind to be active means having adequate ideas, it follows that pleasure arising from this activity is different than pleasure arising from inadequate ideas. This pleasure is derived from pure understanding of the real causes of affections, and not from passive reactions to them. It is then clear that bondage and pain are not determined by all emotions, but only by those derived from inadequate ideas.

Although it might not seem intuitive, the process that leads men to greater activity is a purely “epistemological” one. Inadequate ideas are transformed into adequate, complete ideas by understanding the determinations of passions. Passions are not destroyed; they are integrated with a more extensive understanding of their causes, and hence modified into active emotions to propel men’s pleasure. Thus, according to Spinoza, the process leading to freedom, or to an increase in the power of activity, necessarily involves a transformation of inadequate ideas into adequate ones through the understanding of the causes of our desires.

In turn, this means that because activity involves understanding, and freedom is nothing more than activity, then reason is the real road to liberty. In the previous section it is pointed out that Spinoza identifies virtue with power and conatus, and that he asserts that virtue is nothing more than man’s power defined by man’s existence. But if reason is what augments human power, what increases human conatus, and if virtue is identified with conatus, then a question arises on the relationship between the two. Are reason and virtue related? Spinoza says that:
“To act in absolute conformity with virtue is nothing else in us but to act, to live, and to preserve one’s own being (these three mean the same) under the guidance of reason, on the basis of seeking one’s own advantage” (E IV, 24).

Thereby virtue and reason are integral concepts, and it can be deduced that acting virtuously means acting under the guidance of reason to reach freedom. Spinoza asserted that men are not aware of the causes of their desires, and have inadequate ideas of these affections. Furthermore, he maintained that the passage from a state of passivity to a state of activity involves the formation of adequate ideas, which is the integration of inadequate ideas with understanding. It then follows that, because to act accordingly to virtue is to act out of the necessities (determinations) of one’s own nature, and to know one’s nature means to be aware of one’s own determinations, man acts with virtue only insofar as he is active, that is, only insofar as he understands.

For Spinoza the “essence of reason is nothing more than our mind in so far as it clearly and distinctly understands,” (E IV, 25 proof) and we are therefore left with a direct relationship between virtue and reason. Reason is necessary to self-determine and increase the power of activity, and acting under the guidance of reason means acting out of virtue. Thus virtue, a concept commonly associated with ethical connotations, is here related with intellect and reason. In reality, for Spinoza, reason is the very key to ethics. This morality, however, is not characterized by artificial and arbitrary moral values which are set to compromise between individual liberty and social efficiency. It instead finds its very basis on freedom and cooperation among men.

If men were free and not affected by passive emotions they would not associate imaginary causes to their actions and those of their fellowmen. They would actively understand the adequate determinations of events as necessities following from their essence. In fact, as Spinoza points out:

“Man’s judgment is often governed solely by emotion, and things which he believes to make for pleasure or pain and which he therefore endeavors to promote or remove are only merely imaginary” (E III, 51 xc).

Insofar as subjected to passions, men will experience pain and diminish their conatus. As Spinoza asserts though, passions are derived only from the imagination of those lacking an adequate understanding, so that men are “responsible for what they feel” (E III, 51). However, because emotions either
increase or diminish our power, it then follows that men are necessarily and directly responsible for their strength and weakness; in cultivating reason, men become stronger because they transform passions into active understanding.

Insofar as men act under the guidance of reason, they will “judge nothing else to be to their advantage except what conduces to understanding” (E IV, 26). This means nothing more than they will seek things augmenting their power, because the power of the intellect is understanding. At the same time, active men are free men that act with virtue, and this is where ethics comes into play. It is clear that if all men followed reason they would then necessarily share a common ground for judgment, as the previous proposition states. All men would seek to self-determine their own power, and thus have a common goal. According to Spinoza, they would agree with each other; more so, they would have an agreement in their nature. Throughout Ethics, Spinoza clearly states that only when things have something in common they can affect each other. So, if men share common foundation, and if besides being understood through the same attribute, they also agree in nature, it then would follow that they could affect each other. Also, because they share a common purpose, understanding Nature and their nature, and this nature is in agreement, then they will necessarily be advantageous to each other. This is exactly what Spinoza says:

• “There is no individual thing in the universe more advantageous to man than a man who lives by the guidance of reason” (E IV, 35 cor1).
• “It is when every man is not devoted to seeking his own advantage that men are of most advantage to one another” (E IV, 35 cor2).
• “Men will still discover from experience that they can much more easily meet their needs by mutual help and can ward off ever-threatening perils only by joining forces” (E IV, 35 sc).

It is now evident how reason, freedom, and virtue are related; we can understand why Spinoza titled this book Ethics. If men overcome their passions by gaining adequate knowledge of their causes, they become more active, thus increasing their power of activity and freedom. Free men are part of Nature; they act out of necessity following the universal order of Nature, understood only through their nature alone. Because this power is virtue comprehended through men’s essence, acting with reason means acting with virtue, and by acting virtuously men are of great advantage to each other. Spinoza’s freedom is a self-determination that, by starting with one’s awareness of his own power, radiates to all humankind to implode into the sublime understanding of the
nature of God (Nature).

Spinoza says then, that men first must become aware of the causes and determinations of their own essence, and only then join forces with other men and embark in the same quest leading to the highest form of freedom: the knowledge of Nature. Spinoza does more than this; he gives his definition of True Freedom, and how it is intrinsically related to virtue and reason:

“Firm existence, which our intellect acquires through immediate union with God so that it can produce ideas in itself, and outside itself effects agreeing well with its nature, without its effects being subjected, however, to any external causes by which they can be” *(Short Treatise on God, Man, and His Well-Being, 149).*

This passage reveals through its sibylline fascination the way to True Freedom. But the same sibylline fascination also obscures the meaning. What is this immediate union with Nature, and how is it possible not to be subjected by any external cause?, especially considering that Spinoza himself affirms that men cannot exclude themselves from the causality nexus of the universe. We must not forget that men are modes explained either under the attribute of extension or thought, and not through themselves. This signifies that men as modes in Nature can reach True Freedom only if they lucidly and simultaneously understand themselves and their emotions, and the essence of things. True freedom in fact does not merely involve the second type of knowledge, which is the passage from the imagination of things into adequate ideas of the essence of the attributes of Nature. The highest form of human freedom consists in the passage from the formal knowledge of these attributes to the knowledge of the real essence of things. In other words, true freedom is the absolute knowledge of Nature as *natura naturans* and not *natura naturata*. Spinoza says that:

- “From this third type of knowledge there arises the highest possible contentment of the mind” *(E V, 27).*
- “The highest virtue of the mind is to know God, that is, to understand things better by the third kind of knowledge and this virtue is all the greater the more the mind knows things by the third kind of knowledge. So he who knows things by this third kind of knowledge passes to the highest state of human perfection, and consequently is affected by the highest pleasure” *(E V, 27 proof).*
True freedom is therefore nothing more than the power of the intellect through which one can not only embrace, but lightly merge with Nature and conceive things as eternal and necessary, reaching the supreme contentment of the mind.

It has been argued that such a mental state is impossible to achieve, and that in Book V Spinoza’s thoughts resemble the ramblings of a “God-intoxicated man.” Although Spinoza himself admits the difficulty of reaching the knowledge of Nature, he also affirms the possibility of understanding Nature. Trying to elucidate Spinoza’s arguments for *scientia intuitiva* (intuitive knowledge) and the Intellectual Love of Nature is far beyond the purpose of this paper, but it is necessary to point out that this higher knowledge is not a loftier, nobler, and esoteric state of the mind negating reason. One ought to know the attributes of things before understanding their formal essence; without the cognitive process leading to the formation of adequate ideas it is impossible to reach the intuitive, immediate, all-encompassing understanding of the essence of Nature. Augmenting our power of activity and freeing ourselves from the bondage of passion is absolutely necessary to experience the ‘highest contentment of the mind” (E V, 27).

The third type of knowledge is not separate from the second type of knowledge (guiding humans to virtue); but it is only the final step to perfection. It is the evolution of that very freedom, understood in terms of activity, virtue, and reason, that leads to *scientia intuitiva*. The difference between the two knowledges is not solely quantitative, but it also involves a change in the ontological state of humans; a change in the conception of their own being. In the second type of knowledge men are modes understood under the attributes of Thought or Extension, so to this respect they will never be “free” as Spinoza intends it: “That thing is said to be free which exists solely from the necessity of its own nature, and is determined to action by itself alone” (E I, *def* 7).

At the third stage of knowledge the mind is eternal and conceives Nature as the immanent cause; it perceives everything as existing, with its existence derived from its essence. For Spinoza only Nature’s existence follows from His essence, so saying that everything is perceived as existing means that everything is necessarily perceived as Nature, and not merely in Nature.

Spinoza calls this state True Freedom because all affections are understood as Nature, not only as following from the nature of things. For instance, he affirms:

“It may be objected that in understanding God to be the cause of all things we
thereby consider God to be the cause of pain. To this I reply that in so far as we understand the causes of pain, it ceases to be a passive emotion; that is, to that extent it ceases to be pain. So in so far as we understand God to be the cause of pain, to that extent we feel pleasure” (E V, 19 sc).

This is more than modifying inadequate ideas into adequate ones. It involves not only the understanding of the causes related to oneself as a mode, but the wisdom of perceiving the original, immanent cause. In this wisdom of perception, as Spinoza affirms in the last part of the proposition, no pain can be experienced. All pain becomes pleasure; everything becomes pure power of activity, pure Freedom. But again, this freedom does not originate from some separate entity as will, it is the power of the intellect, it is reason, it is understanding.

Thus True Freedom and Freedom (for humans as modes) are not only inevitably linked to one another, but they also emerge through the same process: reason. The answer to the question posed earlier, whether human freedom exists in Spinoza’s philosophy, is necessarily affirmative. Freedom does not only have different degrees, but it is related to the intellectual maturation of men. Above all, it can be achieved even without the final liberation, the ultimate proclamation of the eternity of the mind. This freedom is human freedom reached through the second type of knowledge; it is the moral freedom of acting under the guidance of reason, with virtue and understanding. It is the liberty of behaving according to an intrinsic ethos, common to all of those sharing the same nature.

“The good which every man who pursues virtue aims at for himself he will also desire for the rest of mankind, and all the more as he acquires a greater knowledge of God” (E IV, 37).

Spinoza unambiguously explained the steps to freedom, consciousness, and human cooperation. This self-determination culminates through the union of the mind with the whole of Nature. It is an abandonment of one’s self that does not require arbitrary faith in decadent values, or blind beliefs in castigation, sacrifices, and pining in order to obtain liberation. This freedom is here; it is man’s very essence, very air, and his power to soar into the depths of Being.

Nominated by: Prof. Tammy Nyden-Bullock
University of Redlands
End Notes


4 Spinoza uses the words Nature (Natura) and God (Deus) interchangeably: Deus, sive Natura, (God, or Nature) His conception of God is not an anthropomorphic, theological entity, but an infinite, immanent, all-encompassing substance. “By God I mean an absolutely infinite being; that is; substance consisting of infinite attributes, each of which expresses eternal and infinite essence.” (E I, def 6). Throughout the paper the word Nature as meaning God has been capitalized, while ‘nature’ corresponds to the ‘essence’ of a thing or mode.

5 E I def 5; E I, 23

6 E I, 24, 25

7 See E I, p 29.


9 E III, 4, 6

10 E III, 1, 3

11 The word ‘imagine’ is crucial to understanding this notion of passivity. We have pointed out how inadequate ideas, which make the mind passive, are confused and fragmentary ideas. It is important to comprehend that this confusion is not inherent to ideas a priori, but rather involves incompleteness. Ideas are not ‘wrong’ or ‘right’; they involve the degree of reality of bodies. In Nature, all ideas are adequate; they follow from its nature and are necessarily contained in it. However, in men’s minds these ideas can become inadequate when affection of the human body are thought to be caused by other external bodies that are not in reality present (See E II 17 sc).

12 E III, 58

13 It must be noted that such a transformation is possible only because of the absence of any metaphysical principle outside or inherent into substance to follow. In fact, unlike the Judeo-Christian tradition, Spinoza does not set an absolute good and evil in a metaphysical sense. His conception of good and evil in fact is based only on what increases or decreases the state of perfection of a particular mode, and does not involve arbitrary judgments or “values.” Nature is neither good nor evil, but only an eternal, purely active and perfect substance. Moreover, Nature does not act with ‘freedom of will’ (E I, 32 cor); there isn’t any external or finite cause why he is determined to act (E I, 17 cor1, 2, proof, sc).
14 E IV, 35
15 E I, axiom 5; E I, 3

**Bibliography**


**Must We Save Only the Phenomena?**

*Richard Warfield*

*George Mason University*

In the beginning of his article “To Save the Phenomena,” Bas van Fraassen states the terms of scientific realism: “[T]he aim of science is to give us a literally true story of what the world is like; and the proper form of acceptance of a theory is to believe that it is true.”¹ We might want to continue this definition to say the guiding view of realism is that our descriptions of the world accurately and truthfully reflect it. Van Fraassen’s position, counter to the realist’s, holds that the acceptance of a scientific theory in no way requires the belief that the theory is true. Rather, for van Fraassen, the acceptance of a theory requires only that it be the foremost explanatory account of all relevant empirical phenomena on the basis of its pragmatic consequences—i.e., its predictability and success in advising other scientific endeavors. According to van Fraassen, when we have a theory that employs non-empirical causal properties or highly theoretical entities we can accept it if it meets both of the above criteria: pragmatic success and empirical adequacy. However, we cannot assent to its truthfulness because of the non-empirical and undetectable elements it entails; this is to say, if we cannot detect particular constituents of a theory, despite the theory’s supposed success, we cannot say it gives a literally true story of that aspect of the world to which it pertains. The best we can do, according to van Fraassen, is to accede to a knowledge of the observables of a given theory and acknowledge the law-like characteristics but ontological uncertainty of the causal properties and highly theoretical items associated with that theory. My paper is designed to show how the empirical anti-realism of van Fraassen falls prey to realist objections of a certain kind, as well as to demonstrate the noteworthiness of these realist arguments. Let us turn now and see how van Fraassen illustrates his view with a discussion of Newtonian theory.

Van Fraassen begins by noting the distinction between the ‘real’ phenomena and the theoretical model devised by Newton to replicate these phenomen-
ena. He calls the empirical, measurable, and interrelated workings of the actual phenomena *appearances*, and the presumed corresponding elements grounded in a mathematical “Absolute Space” inhering in Newton’s theory are referred to as *motions*. Van Fraassen believes that for Newton to claim his conceptual model accounts for all known phenomena or is “empirically adequate” means “all actual appearances are identifiable with [isomorphic to] motions in that model.” However, van Fraassen maintains that Newton’s theory (TN) extends beyond its correspondent role with the use of “Absolute Space” and the attendant notion that the center of gravity of the solar system in Absolute Space has zero velocity. Because of certain mathematical considerations, Newton concedes that the appearances would be unaffected if there were to be any change at all in the velocity of the center of gravity: “But, as he himself [Newton] noted, the appearances would be no different if that center were in any other state of constant absolute motion.” Van Fraassen calls the introduction of any changed velocity back into TN, TN\(_v\). Whereas TN has velocity 0, only TN\(_v\) can have any numerical value for velocity \(v\) provided that it is a constant. When factoring in the aforementioned synonymy of appearances in light of a change in velocity in TN, TN and TN\(_v\) are both empirically adequate and therefore empirically equivalent. Having established empirical adequacy between Newton’s theory and possible derivations of it, van Fraassen addresses the issue of determining the truth-value of our proposed theories TN and TN\(_v\).

Van Fraassen suggests we suppose there is someone \(L\) who takes the Newtonian idea of “Absolute Space” to be empirically insupportable. \(L\) does not believe in the appearance of Absolute Space and cannot say that the motion of Absolute Space is correct in TN. This does not preclude \(L\)'s acceptance of TN as empirically adequate, though. In other words, TN sufficiently accounts for the appropriate phenomena but is not isomorphic, or true, because it erroneously postulates motions that are in point of fact empirically questionable. This incongruity lies in the employment of an idealized grounding that renders the velocity of the center of both TN and TN\(_v\) equivalent. There is only one set of appearances but more than one set of motions to correspond to those appearances. And since each set of motions is empirically adequate we have no way of adjudicating between different views. It would seem that no view is true, or at least since we cannot know whether a particular view is true, we are, says van Fraassen, “still free to believe that each is false.”

Theories as a whole, including non-empirical components, are not true or do not literally say what the world is like. However, the portion of the theories that are isomorphic to ‘appearances’ are true. Van Fraassen does not deny that
a theory may offer knowledge of observables, but he would be reluctant to say that theories reflect knowledge of the highly theoretical; he assents to the reality of observables but remains noncommittal as concerns the associated theoretical parts. He does not, therefore, state outright that one exists and the other does not; only we are unable to account for the theoretical parts because of “our limitations, the limits of observation, which are not incapacitating, but also not negligible.”

But if we cannot accept any one theory as true, since it is possible to posit a number of empirically equivalent theories, where does that leave science and what are we to make of theories? Van Fraassen does not advocate abandoning scientific theories simply because they utilize empirically dubious causal laws and the like. In van Fraassen’s picture, theories do not give us absolute knowledge about theoretical entities, but instead provide us with an efficacious guide if we utilize the theories. By applying them we can determine which ones display law-like features and which ones do not; then we can accept the theories that display these features and use them in guiding further scientific work. The choosing of one theory over another, says van Fraassen, “is a pragmatic superiority and cannot suggest that theories, empirically equivalent in the sense explained, can nevertheless have different empirical import.” Let’s turn to some critical comments on van Fraassen.

The following critique aims at revealing how van Fraassen’s espousal of empiricism and empirical adequacy is too ontologically restrictive of scientific theories claiming realism and truthfulness: his argument is too strong. At the same time, however, van Fraassen’s argument is too weak inasmuch as his other criterion—that of pragmatic utility—equivocates on the status of ontological matters when there is good reason to believe many or even most of our scientific theories are at least approximately true. In order to argue this last point, it is instructive to enlist a basic understanding of Donald Davidson’s articulation of realism: “[B]eliefs are by nature generally true…[one] has only to reflect on what a belief is to appreciate that most of his basic beliefs are true.” To explicate further, we are right to believe most of our beliefs are true, especially the ones which, if changed, would issue axiomatic and extensive adjustments throughout our belief-system because of the implausibility of a ‘radical translation’—the Quinian thesis that asserts, essentially, if any set of noises or marks cannot be translated to some degree into our conceptual schema, there is no justifiable reason to call them parts of a language or reasoning-system at all. Analogously, we may state that we are right to believe most of our scientific theories are true because we could not understand what it would
mean for them to be generally false. We can move from the generality of beliefs to the specificity of scientific propositions because our beliefs constitute our understanding of the world and our basic scientific beliefs are part and parcel of this belief-system; our basic scientific beliefs are gotten from and entwined with the reasoning-system of our basic overall beliefs. A decisive split between the two would simply evoke the problematic of a radical translation. Of course it is also the case that the more theoretical our scientific beliefs, the more questionable their truth-claims become—a kind of supervenient relation between the basic and more theoretical (though this is unlike van Fraassen’s claims because his doubts concerning the truthfulness of the highly theoretical are based on empiricism). In effect, Davidson is purporting a coherence theory of truth that surmounts questions about the piecemeal dilemma of the verification of the truth of any prior theories used to substantiate new theories by principally reversing the premise and conclusion such that theories are true not by virtue of the fact they cohere, but they are in principle true, therefore they cohere. A longstanding difficulty with most standard coherence theories of truth has been the difficulty of offering—which Davidson appears to do—a non-question-begging definition of coherence.

Van Fraassen desires to make a distinction between elements in a conceptual theory that account for empirical phenomena and those parts of the theory that do not appear to correspond to any empirical phenomena. This distinction between the empirically real elements and unreal theoretical postulates allows van Fraassen to largely reject truth-claims of theories about nature, supposing that only what is empirically observable or detectable is real; this separation of theoretical and empirical permits the notion of empirical adequacy since the empirical must remain uncontaminated by anything non-empirical to be the ultimate arbiter of scientific propositions. But does van Fraassen’s argument hold under closer scrutiny? If we can muddy the margin between the empirical and theoretical components of our scientific picture, van Fraassen’s criterion of empirical adequacy, and hence his anti-realism will be compromised. In order to achieve this we will have to invoke a Davidsonian understanding of coherence since a standard coherence theory of truth entails methodological flaws of its own and is therefore suspect in probing van Fraassen’s argument.

Let us say we have any number of competing theories that predict certain outcomes—though, it is inconsequential whether a detection of the phenomena under investigation precedes any explanatory model or whether certain calculations hypothesize the existence of as yet undetected phenomena (the latter might have seemed to bolster a standard coherence theory of truth) be-
cause confirmation still amounts to empirical confirmation. Therefore, of course, any claim to empirical adequacy is subject to testing. But when testing the outcomes of each theory to see if they are indeed empirically adequate, we are compelled to draw upon our theoretical backdrop—that we may say is at least approximately true—through the use of ancillary theories in our testing procedures which are apt to be vitally bound to our overall scientific beliefs and practices (recall the theory of radical translation); we could not forward or test a theory except by reference to previously employed theory-terms. Theory bears substantially on the empirical in our scientific procedures. Empirical adequacy as an exclusive test for a theory’s truthfulness follows only in reference to individuated phenomena that could not in fact be individuated empirically. The idea of there being a certain theory equivalent to all ‘relevant’ phenomena without having a causal connection to other phenomena on the basis of an indispensable appeal to some scale of theoretical backdrop (scientific, linguistic, etc.) is untenable unless the phenomena are divided a priori and in contention with van Fraassen’s empirical efforts; were this the case we could affirm the correctness of our theories; the bind is that, of course, this information would be ineffable and beyond our apprehension. Van Fraassen assumes a critical demarcation of phenomena when he employs words such as “relevant” and “appropriate” to describe the affinity between appearances and the particular corresponding elements in motions because even though may be ‘appropriate’ to , how are we to segregate the specific under investigation without reference to some set, or sets, of conceptual models not equivalent to —this is to say, to other, frequently used theories. Terms like ‘relevant’ require a causal background (or backdrop more generally) we apply to decide (pragmatically) what counts as relevant or not; although we can say our theories are approximately true, the need to differentiate one from another is a pragmatic one, but one that is necessitated by our use of conceptual models. Relevancy and the like is, accordingly, a theoretically necessary precondition for deciding empirical adequacy. Van Fraassen’s distinction between the empirical and theoretical forces an intra-empirical differentiation of phenomena, and this ultimately commits him to the ontological reality of theoretical postulates. If he says that only the empirical is real then he must say that pertinent causal/theoretical elements are real as well, because empirical phenomena would not be without them. And if a concordance of empirical/non-empirical occurs at this level (appearance) it must also occur in conceptual models (motions) that hope to replicate these phenomena given that empirical phenomena require a theoretical background to delimit them and make them utilizable for what van
Fraassen commends as furthering scientific endeavors; the only way phenomena are communicable for affecting (and generating) scientific practice is through a cohering theoretical arrangement. This draws us to the question of coherency, and this is also where the advantage of Davidson’s argument is seen.

Our use of ancillary theories—which we see are necessary—would help to confirm or disconfirm one of our theories eventually, with ample ancillary theories and therefore with reference enough to associated phenomena to warrant good reason to be convinced of confirmation or not, as there is no solely empirical way to separate certain phenomena from others in order to argue the cogency of empirical adequacy. We could distinguish among them because one would begin to cohere with other background theories that are in principle true, the ones that cohere the best are most apt to be true. But the truthfulness of our constituting and dialectically evolving scientific beliefs is not founded upon the ambiguity of determining if ‘cohere’ means something like ‘entails’ or ‘is in keeping with,’ or the dilemma of how we could gain a sense of ‘coherence’ without the concurrence of a linguistic set-up of which it would have to be a part. Rather, our definition of ‘cohere’ is instituted if we say most of our theories must in principle be at least approximately true. By deflating all possible descriptions, substitutions, or ways of defining ‘cohere’ to what scientific methods and beliefs are, we can state that coherence is precisely the status of these methods and beliefs; hence we skirt any problems of ambiguity or foundation. Our scientific account is (approximately) accurate and truthful without having to be an entirely independent and detached description. This is similar to how an understanding of truth in Davidson’s account avoids circular descriptions of what truth is by deriving it from the state of our general beliefs; we simply say, ‘truth is what most of our beliefs are since it would be literally incomprehensible for them to be otherwise.’

I would like to mention a last argument for realism of a non-Davidsonian type—an argument from instrumental reliance—that would hypothetically undermine empirical adequacy and pragmatic superiority. I do so because I believe it conjures up the same weak-points as a standard coherence theory of truth, and so emphasizes the favorability of adopting a Davidsonian argument for realism against the anti-realist empiricism of van Fraassen. Take, for example, the following: If causal properties which, in van Fraassen’s view are non-empirical, were not correct (or close to it), then any instruments that we use to observe scientific phenomena would not work as planned. A productive outcome for a scientific experiment assumes that causal properties associated
with the proper functioning of the experiment will remain unchanged. Those causal properties inhering in the table (molecules of wood and the like) that my hypothetical experimental device is placed upon follows the theoretical laws that we have discovered about it—in other words, the table stays put. If the experiment functions properly then it would be safe to say that the conditions/theories helped to explain or inquire into empirical phenomena. But this argument is circular just as our standard coherence theory of truth is likely to be. The theoretical structure we use to test whether our experiment turned out as planned is the same one used in the predictive process, so it is not a reliable judge of its own truthfulness but simply a sign our theory-structure is ‘consistent’, and if any determination is made based on consistency, it runs the same risk of ambiguity as ‘cohere’. However, if we accord an approximate truthfulness to our theories, as Davidson’s notions allow us to do, then the purpose of successful experiments appears to be the refinement of our ontological beliefs, to what extent they are true, and not the fulcrum over whether they are at least approximately true or not. And in attempting to solve (and continue to solve) the question of a degree of accuracy, realism of the kind expressed here extends beyond advising other scientific endeavors (by way of successful experiments) on the basis of what van Fraassen believes is nothing more than a “pragmatic superiority.”

Whereas van Fraassen would like to claim that causal properties have no correspondence in the empirical realm, it can be shown that causal properties are requisite for the detection of empirical phenomena. Causal properties are far from empirically irrelevant. If van Fraassen claims that certain elements are right or true because they are isomorphic to empirical phenomena, he has to admit the correctness of the causal properties that are linked to these elements because the empirical phenomena cannot be accounted for without causal properties. Van Fraassen’s empiricism excludes too much. And his pragmatism admits too little insofar as it admits only a law-like acknowledgment of theoretical workings and entities, when, analogizing from Davidson’s coherence theory of truth, we know it would simply make no sense if the causal properties attributed to the empirical were not at least approximately true.

Nominated by: Prof. Daniel Rothbart
George Mason University

End Notes
1 p. 187. van Fraassen, Bas. “To Save the Phenomena.” *The Philosophy of Sci-
  2 ibid p. 188.
  3 ibid p. 188.
  4 ibid p. 189.
  5 ibid p. 191.
  6 ibid. p. 190.
The Problem of Space

Thomas Cochran
University College London

Introduction
What is space? One answer would be to say that space is the medium in which all physical substance exists. Yet when we try to isolate space from the stuff contained within it what are we left with? Is space something? Does it have features which are distinctively features of space and not matter or energy, or when we divorce ideas of substance from space are we left with nothing whatsoever? It is one of the most fundamental concerns of physics and metaphysics that any analysis of matter requires an understanding of the assumed medium in which it exists and yet something, which seems so obvious when first looked at, can have many surprising and controversial implications.

Probably the most famous debate over space, the Leibniz-Clarke correspondence, was instrumental in forming a split between science and philosophy and helped physics emerge as a distinct discipline. The correspondence was carried out during the period of 1715 to 1716 but the source of the debate is much more far reaching. Both positions will be looked at in light of their historical context, and also how they should be viewed given contemporary theories of space.

The debate centers on whether space should be regarded as real, a concrete feature of reality with an independent existence from the matter we suppose to be contained within it. Thus, it is argued that if space has an independent reality it is at least possible that there could be absolute space, where there is no distinguishable matter or substance to be found in the universe. If this is the case, it would follow that it is not necessary that any matter exists and hence if we regard space as empty, a pure vacuum, then the question ‘why is there something rather than nothing?’ has genuine meaning. Although the question of whether vacuums exist is implicit in any discussion of space, I shall be concentrating mainly on arguments concerning the possibility of absolute space, where only space exists. If absolute space is a coherent concept then naturally,
I would suppose, so is the concept of a vacuum, though it should be noted that the existence of vacuums, or an intelligible concept of space that doesn’t require notions of matter, by no means necessitates the possibility of absolute space. This is because arguments concerning possible states of everything have radically different implications than those concerning particular elements within it.

The arguments for and against absolute space shall be looked at in various different forms, as a logical, mathematical and imaginative possibility that looks at the problem in epistemological terms. Also as a metaphysical and nomological possibility which requires an analysis of actual and possible laws of physics. I shall argue that whilst within current conceptions of space we should regard it as a definite part of our ontology, it is by no means clear that it is independent of matter. I shall attempt to show that the idea of absolute space is incoherent because the notion of the absolute employed, central to which is the idea that everything is the same, is inevitably fraught with contradictions. Thus something must always exist. This position has various philosophical implications, which I will try to explicate by arguing that this position supports some contemporary theories in physics, which although counter-intuitive remain philosophically satisfying and coherent.

**Leibniz versus Clarke**

The background for Leibniz’s refutation of absolute space is based on his rejection of the kind of atomism proposed by the early Greek Atomists. By espousing elemental particles that were exactly the same, the Atomists violated laws of reason that Leibniz felt he had established. Thus, Leibniz’s rejection of the “pleasing imagination of vacuum and atoms” is based on analytic and a priori reasoning on how the world must fit his principles of logic.

As an alternative explanation of space Leibniz accepted the Cartesian picture of an infinite plenum populated by infinitely divisible bodies. Descartes’ own refutation of the existence of vacuums is based on his claim that the essential property of everything within the physical realm is extension. Matter and physical space are identical and hence the idea of a void or absolute space is incoherent. He argued that if there were nothingness between any two objects, the two objects would be next to each other. One of the best reasons why we should believe there is empty space is because it allows objects to move around freely. So Descartes espoused the existence of vortices, the circular motion of which would communicate movement. Leibniz largely accepted this picture, where motion is carried by means of impulses through the plenum, although his reasons for this were quite different from Descartes’.
Leibniz bases his rejection of absolute space on the grounds of his two fundamental laws of reason, the identity of indiscernables and the principle of sufficient reason. The law of the identity of indiscernables states that if any two objects have all their properties in exact accordance then it is in fact the same object. Some philosophers reject this on the grounds that two objects can be qualitatively identical. For example, two machine-produced chairs can have all their properties in common and yet be numerically distinct. However, for the purposes of refuting the possibility of absolute space we need only take the principle in its weakest form, i.e. that the features of an object include all its relational properties, its position in time and space relative to other objects. This is usually taken to be trivially true, that two qualitatively identical objects cannot share the same position in space and time without being numerically identical, yet we shall see later that accepting this raises a problem for the Newtonian conception of absolute space.

The principle of sufficient reason is a little more debatable, since it focuses on teleological rationale, and is seen as purely heuristic for the purposes of elegant and efficient explanation. Kant, for instance, said that it is not a substantive principle but serves as a philosophical guide. The principle states that for everything that occurs there is a reason why it happens in the way it does and not otherwise. Thus, in his monadology Leibniz uses this principle to explain why God created this particular universe, the sufficient reason being that it is the most maximally self-consistent possibility, that is the greatest variation from the simplest laws.

Leibniz’s strongest argument against the possibility of absolute space runs as follows: In absolute space each point of space is identical to every other because each has the exact same feature, being devoid of any property we could attribute to a substance, except perhaps the most basic qualities of dimension. In all respects, absolute space must be completely uniform, so that nothing except space could be found in it. Also, since absolute space is infinite, each point within it bears the same relational qualities as every other because infinite distance extends in every possible direction no matter where you specify any point within it. It seems then that if the principle of the identity of indiscernables is true, one cannot coherently distinguish one point in space from any other and hence that the idea of an infinite amount of space creates a multiplicity of spatial points which cannot possibly be a genuine multiplicity. The concept also violates the principle of sufficient reason in that there seems to be no reason why space should be orientated in the way it is rather than the other way round, or why for example, the whole thing was not positioned an
inch to the left.

It may be argued however, that in order for Leibniz’s argument to work he
does not require absolute space to be infinite. Certainly his arguments against
the possibility of any void whatsoever, do not require spatial infinity in exten-
sion, although given his theory of matter, I would suppose he would assume
void to be infinitely divisible. If it was finitely divisible we could ask why
could there not be just one atomic point of void.

Leibniz bases his refutation on the identity of spatial points. Whether we
are talking about all of space or just a limited portion of it, any area of space
you care to pick out will be identical to any other geometrically similar area of
space. Hence, one is not able to pick out any area of space at all. But clearly
when we have a finite region of space we can pick out different sections within
it, for example the edges can be distinguished from areas further in. Areas in
the void can be distinguished by virtue of being in different locations relative
to one another. After all, in his Monadology (1714), although monads differ in
respect of their internal properties, all this comes to is that the internal differ-
ences are in terms of relational properties to other monads.

Only if we take the identity of indiscernables to be true in its boldest form,
that there are no two qualitatively identical objects in the universe, can Leibniz’s
refutation of the possibility of the finite void work. Yet there seems to be no
necessity that we should accept the bold version of the law to be true, at least
not necessarily true, as Leibniz would demand. The same could be said for the
principle of sufficient reason. Although it would serve as a refutation of the
possibility of absolute finite space, we are under no obligation to accept that it
is true. Leibniz is only using the principle of sufficient reason in the same
sense as he does when arguing that this is the best of all possible worlds, that is
on the basis that there is always a best reason to have as much variation as
possible. In all, we have as good a reason to accept the obvious appearance of
things, that there is empty space, as we do of accepting some of the logical
laws of Leibniz. We should concentrate then, on the formulation of Leibniz’s
argument as I have first used it, that the weakest version of the identity of
indiscernables refutes the possibility of absolute infinite space.

The reason why Clarke accepted absolute space was probably mostly be-
cause of the empirical evidence in favor of the vacuum. Guerike’s experiments
where air is pumped out of a vessel and Torricelli’s work with mercury barom-
eters which shows that substances have different densities and hence relative
degrees of void within them, seem to provide irrefutable empirical evidence
that vacuums exist. Leibniz did not take these experiments particularly seri-
ously, claiming that there would be things such as beams of light and magnetic
effluvia present that are simply ‘void of heaviness’.

As well as empirical evidence in its favor, the Newtonian picture of space
contained an extremely elegant and satisfying explanation of how the planets
moved, by means of the force of gravity attracting objects from any distance.
Even though the notion of action at a distance is a very unnatural idea (Leibniz
called it an occult quality, in reference to the rejected notions of Aristotelian
and scholastic philosophy), it is what actually appears to be the case. No evi-
dence could back up Descartes’ concept that the planets revolve because they
are contained within a giant vortex. This is why the debate was seen to have
been ‘won’ by the Newtonians who were able to set the course of modern
physics for the next three hundred years.

Ultimately, of course, the Newtonian laws of gravity, the inverse square
law, as well as classical mechanics in general have been superseded by Einstein’s
theory of general relativity and quantum mechanics. Einstein’s conception of
space is one where gravity is determined by the curvature of space-time (thus
negating the need for action at a distance) and how this curvature is relative to
the density of objects such as stars and planets contained within it. However,
whilst it may seem as if Leibniz has been vindicated in his view that space and
time are relative, we need to see whether the reformulated refutation actually
applies. That is, must absolute space be infinite?

Infinity

When formulating concepts about absolute space in the sense that only
space exists, the idea of infinity is often used. From the earliest Greek thought
about space it is often assumed that space must be infinite. However, much
contemporary physics theory is undecided on this issue by claiming that it is
possible for space to be finite and simply unbounded. Since my formulation of
Leibniz’s primary argument against the possibility of absolute space is grounded
on the assumption that space is infinite it is important to question whether in
any sense space must necessarily be infinite. If absolute space must be infinite
then Leibniz’s argument will be a great deal more forceful. If not, then by
elucidating why space is not necessarily infinite we will gain much more in-
sight as to the nature of space. However, the concept of infinity has been the
source of continued disagreement so we need to be clear as to what kind of
infinitude is relevant to absolute space and whether this in turn is applicable to
Leibniz’s argument.

Probably the most challenging problems concerning space and infinity are
to be found in Zeno’s paradoxes of motion. Seemingly the most self-evident
feature of space, smoothness and hence being analogue, that is, where between any two points there is always a third, is apparently shown to be paradoxical. The ‘race-course’ or ‘dichotomy’ paradox (recorded by Aristotle) is probably the clearest and runs as follows: In order for someone to traverse any distance he must first cover half the distance and then half the remaining distance and so on \textit{ad infinitum}. In all he must cross an infinite number of points in between and as it is impossible to complete an infinite series, motion is impossible. Quite obviously however, people can move, so an absurd conclusion has been reached which must be discharged. It seems rather ironic that space is what we believe to allow free movement and yet its smoothness should prevent the possibility of motion.

Zeno himself used his paradoxes to support the thesis that everything is one and hence that things such as motion are illusory as maintained by another of the Eleatic philosophers—Parmenides. It was in order to resolve this paradox that the Atomists Leuccipus, Democritus and later on Epicurus denied the possibility of infinite divisibility, and hence postulate the existence of indivisible atoms. However, Aristotle was to provide a number of important distinctions within the concept of infinity that allow us to make much better sense of how infinity can be applied.

For Aristotle the notion of the infinite is the same as saying that it is untraversable. Since traversing the infinite would require infinite time, ideas of the actual infinite were practically a contradiction for Aristotle. This is because the actual infinite must be presented all at once whereas the only way you could traverse the infinite was through continuation of time. By contrast the mathematics of the infinite, the concept of constant addition, is close to a tautology because continuation of number is akin to continuation of time (this analogy is also explicit in Kant’s idea of the synthetic \textit{a priori}). It is through this notion of constant continuation that Aristotle developed the concept of the potentially infinite and hence, although anything is potentially infinite simply because it can be conceptually extended or divided indefinitely, this is never actually presented in reality.

What does this tell us then about the nature of space? It tells us that any notion of space must be intimately connected with the notion of time. This idea of course is intrinsic to current conceptions of space-time as Einstein’s theory of relativity demands. It also tells us infinity isn’t so much a property that something can have. Rather, infinity can be attributed when there is no sense that something can be limited in any way, as something that cannot be traversed or measured or complete.
Leibniz’s refutation is grounded on the paradox between infinity and one-ness. Since the view of space derived from Einstein’s theory allows curved space, and hence the possibility of looping space, it seems possible for space to be finite and yet still be unbounded, which would stop us from ever reaching a boundary or any point beyond which we could not traverse. Here the Aristotelian concept that infinity is the same as untraversability is clearly rejected since the most important feature of Aristotle’s account of infinity, that of constant continuation, is compatible with finite space.

The most common argument for this view of space is based on the analogy of the sphere that is finite in three-dimensional space and yet one can walk around its surface in two-dimensional space without ever needing to stop. Thus, in the same way space could be finite and unbounded as a totality because the overall curvature of space time allows us to continue forever within it, although it makes sense to say there is some region which is distinctly outside it, even if we could never be in it. Does this really mean that a state of absolute space could be finite? Absolute space is when everything is uniformly space, yet a finite and unbounded region of space seems only finite and unbounded in respect of the nothingness surrounding it. Thus space can only be finite if it is not everything; i.e. it is distinguishable from nothingness. The problem with the sphere analogy is that it cannot be applied to everything. The n-dimensional finitude is only possible, at least in the only way I can conceive it, given an n+1 dimensional surrounding area. If we are not to create an infinite regress of finite spatial containers we must arrive at a surrounding area with infinite dimension. An area with infinite dimension is just one without any dimension, that is, without any sense to which it can be coherently measured. The only possible form of this is infinite nothingness. Hence, if absolute space is considered as something it cannot be both finite and everything. The concept of absolute space in so far as it is the possibility of nothing rather than something requires that everything is space.

Mathematics

It is worth noting that Aristotle made a crucial distinction between the metaphysically infinite and the mathematically infinite. He decided that while the study of mathematics required a notion of the infinite he attacked the idea that it actually existed either in the form of infinite extension or divisibility. It is important then that when we question the nature of space we must be careful not to confound the mathematical perspective with the alleged difference in how things actually are.

Yet the reliance upon mathematical methods and concepts is central to the
rationalism of the philosophers who were later to inquire upon the nature of space, most notably Descartes and Leibniz. The debate between physics and philosophy that was epitomized in the Leibniz-Clarke correspondence may well have turned on how they differed in their views of mathematics. Leibniz and Newton independently formulated one of the most important discoveries in mathematics and logic, the calculus. However, while I think it would be fair to say that Newton and physicists since then have regarded mathematics as simply a tool, often just a language to communicate their theories to other physicists, many philosophers, especially the Rationalists, have seen mathematics as the paradigm field of knowledge and try to apply its methods to their metaphysics.

The Rationalist philosophers seem to have a very realist view of mathematics. This is not just the long-standing philosophical issue of whether in our ontology we should include entities like numbers, but rather whether we should take mathematical concepts to be akin to discoveries about reality and whether we should accept theories because they are mathematically elegant and coherent. Although these ideas are prevalent throughout the discussion of space it is especially notable that the current vogue in physics is to accept theories based on their mathematically aesthetic qualities. This is because sciences such as cosmology, particle physics and attempts to find a grand unified theory of everything have reached such an abstract level that they are practically impossible to test. Perhaps physicists would do well to apply their ideas to philosophical scrutiny. Instead of accepting theories purely on the basis that they work in the case of quantum mechanics, they should endeavor to finally establish just what view of the world they are implying. It is important not just that the theory works but that we know the reasons why it works, and when it comes to astronomical theory, it is important to know that the mathematics used is justified.

Kant claimed that it is necessary for the mind to perceive the world as being contained within space and that we could know by means of synthetic *a priori* knowledge geometric truths about space. Certainly the nature of space seems to be fundamentally the domain of geometry. However, since Kant’s day, the idea that mathematics is necessarily true, or forms one totally coherent system, has been shaken. Einstein showed that gravity could be explained based on principles of geometry that question some of the most seemingly self-evident assumptions.

What none of the philosophers of the 17th and 18th centuries knew was that Euclidean geometry is not the only coherent geometrical system. The possibil-
ity of non-Euclidean geometry has shown us that geometry no longer gives us a priori truths about the world, just truths whose meaning is given by the axioms of whichever system we are working in. A distinction has to be made between geometry as a pure mathematical system and the geometry of the space we live in, where for instance, a straight line is not just the shortest distance between two points but the line traced out by a beam of light in space.

Thus, if we look at the mathematics of absolute space, we should recognize that it can only bear an analogous relation to its actual possibility. In mathematics the formulation of pure void is paradoxical. It seems self-evident that in this sense of absolute space there is nothing to number and hence it could be represented as zero. Yet because pure void is infinite in extent this zero has to be multiplied by infinity, which mathematicians will tell us equals one. There is a contradiction then between one and zero, but this result tells us simply that we cannot properly formulate absolute space, in the sense of pure void, in mathematical terms. Thus, although the concept of zero is necessary for working in arithmetic, it is not something we can apply very easily to the existence of a pure void.

Conceiving of Absolute Space

We have two possible ideas of absolute space. Firstly, the concept of space as pure nothingness or void which is the alternative to something in the question of why there is something rather than nothing. This is the strictest concept of absolute space. The second is what our contemporary notions of space would look like if there were no distinguishable matter within it to speak of. Although this conception is not absolute in the sense of space as everything, it should be looked at in terms of the possibility of there being no matter in the universe.

Our first notion of absolute space is in line with the Ancient Greek and Newtonian concepts and is characterized by infinite extension and hence infinite divisibility since it is utterly smooth, uniform and thus without any boundaries or distinguishable points. Arguments against its possibility can be made on purely conceptual and a priori grounds since its features are only relevant to abstract mathematical and logical inspection. In order to differentiate this concept from contemporary ideas of space it shall be referred to as pure void.

It is in contrast to the idea of pure void that the contemporary ideas of space gain most of their appeal. I have already described some of the ideas behind it above but of most importance is the willingness to question features such as infinite extension and to have a more positive conception of it than as simply nothing. However, as will be seen in the arguments against the possibility of a pure void, any claim that this kind of space could be absolute or total
is undermined because whether everything is taken as purely one thing or absolutely nothing the notion is entirely incoherent.

Probably the biggest problem when talking about the concept of pure void is that it is impossible to form a positive idea of it. If we take the view that all thoughts are intentional upon an object then how can we form an intentional idea of nothing? It seems natural to suppose that to think of nothing is just not to think. The concept of pure void is deeply puzzling because we are supposed to have a viewpoint or intentional content of a concept which by its very essence excludes all content and viewpoint whatsoever. It is like trying to achieve what Ernest Nagel calls “the view from nowhere” because pure void is completely relative in the sense that everywhere is the same with not one feature to distinguish one point from another.

The objections made by Leibniz described above focus on this problem, yet it seems to me that his arguments can only be directed against the possibility of conceiving of a pure void. Leibniz as one of the paradigm Rationalists may have taken the inconceivability of pure void to be enough evidence that it could not exist. However, we should not discharge the possibility of pure void on these grounds since we might be able to form a relative idea of pure void that is, in terms of not being something or other.

In many scientific and mathematical contexts whether or not something is imaginable can be a good guide as to whether it could exist. As I have mentioned, the conceivability of non-Euclidean geometry led to a view of space that Einstein’s theory of relativity demanded. However, anyone who wishes to defend the possibility of pure void will be unwilling to let the limits of our imagination refute the idea.

Yet when we try to formulate the concept of pure void, in terms of mathematics or logic, again we run into problems. In logic there is simply no way of expressing the sentence, for instance, that ‘there is a state such that nothing exists’ without being self-contradictory. You simply cannot blankly negate everything but rather only something or other. We can deny that any matter exists, but cannot form any fuller conception of pure void than that.

The problem is that the languages of mathematics and logic are intended to examine the natures and compatibility of things that exist. The medium in which this is supposed to be going on is rarely questioned. Thus, it is doubtful that logic or mathematics is in any way independent of what it analyses in the same way as measurement only makes sense in terms of the things or events it is measuring. If pure void is essentially a concept, the coherence of which we are trying to determine, it is highly probable that as a concept at least, it can
only be relative to concepts of substances from which it is distinguished.

**Can space ever be absolute?**

So we are left with a relative concept of pure void, but is it possible that in reality substance could have been or ever will be entirely absent? Here we are dealing with the possibility of absolute space in the metaphysical sense. Probably the strongest argument against this idea is that nothing can come from nothing. It is incoherent that anything that is truly nothingness could possess a force by which substance is created. Also, there could be nothing in the future since even if we were to infinitely disintegrate all the matter that exists such that it is mathematically zero, that is never going to be a state which is actual since the state is only approached given an endless span of time.

Here again, the problem of space is explicitly a problem about infinity and whether it can be regarded as a totality or something absolute (central to the claim of absolute is the idea of oneness, a single independent state). It just isn’t coherent to say that everything is infinite in the same way, as it is incoherent to say that there is a greatest possible number. Yet it is a commonly held claim that in infinite time all things happen. This claim confounds a notion of a single state where all possibilities are achieved with our notion of time as a continuing series which, if infinite, never achieves a final state. Why couldn’t there be just one event that repeatedly occurs again and again forever? This is a possibility that would stop all possibilities from happening. Hence, I cannot believe that actual infinite totalities are possible because we must always invoke a notion of time in any discussion of the infinite. The best way of grasping the idea of the infinite for me is to imagine an event that can be repeated indefinitely.

Kant regards the notion of a totality as the backbone of his concept of the antinomies. The antinomies are to be seen in disputes such as whether there is a first cause or an infinite past or, particularly relevant to this essay, whether the universe does or does not have a limit. The argument against infinite space, (the thesis) is already familiar to us. It states that because infinity is a constant succession of extension it can never actually be at any point that space is infinite. The argument against space as finite, (the antithesis) states that if space were finite there would be a boundary beyond which there was empty space. Since however, space is nothing but the relation of objects this cannot be. Similar arguments are put forward about the finitude or infinitude of time and at the end of this paper I will try to show how a third option is possible. However, at this point it is enough to say that Kant is supporting here in both arguments the impossibility of absolute space. Also, Kant’s arguments do not pose a problem
for the possibility of finite and unbounded space-time as in the theory of relativity. Relativity does not claim that there is any sense to what is beyond space-time, in the same way as there is no sense to saying what is south of the South Pole. In addition, empirical, not a priori, reasoning supports relativity.

Anyway, since both positions in an antimony can be proved, Kant concludes that it is a problem with our concepts of the totality which confusedly take arguments from the phenomenal world to be relevant to states beyond any possible experience of it. The world as presented to us in experience is as a community of the co-existence of inter-dependent objects. It is this kind of view that I am espousing when denying the possibility of the absolute. The only legitimate view of reality is one of an inter-dependency between something and nothingness. There can be no infinite plenum if we are to account for how bodies appear to move, but neither can there be infinite nothingness.

Until Einstein’s theory of relativity, space seems only to have been viewed as a purely empty void, which as we have seen, is rather difficult to isolate as a concept. The contemporary conception, however, makes a fundamental shift in the way we view space. Rather than being a mere empty void through which certain forces act, space can be said to be a definite something with concrete topological features. If only we could see radical and tangible curvature in space such as would be experienced when being sucked into a black hole, we would be very wary of denying that space is an entity in its own right. However, this does not convince me that space is genuinely independent of matter and hence that there could be just space. The features of spatial curvature seem to rely on the density of actual objects within it and so although space is intelligible when considered in itself, without the presence of matter we would not have anything to make sense of.

At this point we must draw a distinction between local space, i.e. the space in between galaxies and so forth and cosmological space, the vast regions of space which remain uninfluenced by any matter at all. It is the overall cosmological space which contemporary physics regards as expanding at an ever-increasing rate. Yet even here it is difficult to separate notions of space from notions of matter. It is currently conjectured that space is generated at the big bang along with matter and its expansion can be described as space becoming less and less dense.

Not only that but there is no such thing as a true vacuum, a purely empty void. All space has a definite energy state that has come about due to the annihilation of matter and anti-matter into energy. Thus space must be distinguished from nothingness and even when we regard space as something its features are
inextricably linked with the processes of energy and matter.

Here space is given as a kind of something, which has varying degrees of curvature and density according to the overall expansion and contraction of the universe. When talking about the expansion of the universe it is important to note that we only have genuine evidence that things such as galaxies are moving farther and farther apart, not that the universe as a whole is getting bigger. However, I think it is probably correct to say that our universe is finite and unbounded since the notion of space being infinite in extent has been shown to be deeply problematic.

As I have mentioned, we cannot say that our current conceptions of space can be absolute in the sense that it is the only thing that is, or that it is everything because there are other things, namely what is outside this bounded region of space-time. There are various theories in physics today that claim that there is a multiverse, consisting of infinite possible universes all existing parallel to each other, or the mega-universe theory which states that there are regions of isolated space-time which continue to branch out of each other.

What we are left with then is a question of whether one particular region of space-time could ever approach a state of complete absence of distinguishable matter within it. I have already mentioned the fact that there is no such thing as a true vacuum since all space has a definite energy state contributed by the annihilation of matter and anti-matter. However, is it possible that all matter and anti-matter that exists could mutually annihilate leaving us a kind of porridge space that is a uniform region of energy? One problem is that there doesn’t appear to be an equality of matter and anti-matter for this to be able to occur. Yet it is just about theoretically possible that there could be an equality and that it could all simultaneously collide (in exactly the right proportions and without separating back into matter and anti-matter again) to leave us with what could be said to be just space. However, the kinds of physical laws and symmetry that could create this possibility are not to be found in our universe. Perhaps in another region of space-time this could occur, but it is not for me to speculate about this since we cannot rule out the possibility that the regions of space-time could affect one another.

**Self-Causation**

Thus the possibility of absolute space must be denied at all levels whether as a fleshed out conception or as a metaphysical possibility. Although it is intelligible to say that there could be no matter, and that a finite vacuum can exist, it is only given that space is something and even then any vacuum cannot be isolated from the various forces within it.
Something must always exist, but what should we make of this claim? The most obvious answer would be to say that the universe has an infinite past and will have an infinite future. However, we have already seen some of the problems that infinity can present us with and should recognize Kant's points about understanding the universe as a totality. It just does not seem philosophically satisfying that there is a continual chain of one event to the next stretching back and forwards forever. How is it that any of this should be there? We need a positive idea of why there is something rather than nothing rather than blankly claiming that there is something because there cannot be nothing. Without resorting to cosmological proofs of the existence of God, I suggest that this positive reason is self-causation.

The physicists Richard Gott and Li-Xin-Li have suggested that “the universe emerged from something rather than nothing—and that something was itself” (New Scientist 24/1/98). Obviously we are not dealing with the ordinary notion of causation here, for something to cause itself would practically be a contradiction in terms for our everyday understanding of the concept. Yet within the theory of relativity, there is a provision for loops in space-time. Just as travelling through space, we would eventually arrive at the same point in space we started at, so it is possible to arrive back at the same time. When applied to everything we have not only the possibility of a finite and unbounded universe in space but finite and unbounded in time. In fact the notions of space and time are not really separable; we should take them to be describing just two aspects of the same thing.

What makes loops in space-time self-causation is that it is exactly the same time that is being repeated, there is no external time to measure the repetitions against. Actually, to use the word ‘repetition’ is to make it sound as if there are numerically distinct although qualitatively identical universes going on here. When the overall space-time is looped, this looping is infinite. We could start to imagine this as an infinitely large circle of identical universes but given that they hold all their relational properties in common, they all lie in the middle of an infinite chain, we can reject this confused multiplicity, and by application of the identity of indiscernables, postulate just one self-existing universe. The infinities cancel one another out.

This theory does not violate any laws of physics. Instead it gives us an explanation of why there is anything at all. If it seems that we cannot resist thinking of it as holding oneself up by one's own bootstraps, this is because we have a confused idea of a perspective outside the universe. In fact no such perspective is possible or even coherent. One cannot get outside of everything
and look at it—which outside perspective would have to be included in everything as well.

It just doesn’t make sense to have a view of everything except as entirely self-contained. Questions concerning what is outside everything as is implied by the question ‘into what is the universe expanding?’ or ‘what happened before the big bang?’ are meaningless. The idea of self-causation isn’t a conception of the universe as a totality but rather can only be made sense of given a plurality of events that loop back on themselves. Hence I believe we can avoid the problems of Kant’s antinomies, where any absolute conception of a totality is contradictory, and claim that the universe is completely relative to itself.

To end this paper, I will conjecture how self-causation is nomologically possible. The most obvious way in which self-causation could occur would be by means of an oscillating universe. An oscillating universe is one where expansion occurs, eventually slows down and stops because of the pull of gravity, and then contracts in a ‘big crush’. If the same initial conditions were to obtain at the point where it has contracted to the extent that another big bang could occur, an exact repetition will emerge. I should point out here that this conjecture assumes determinism to be true. Some have claimed that the necessary limit of our predictive powers in the field of quantum mechanics refutes determinism. This is not the case. Although we can never know certain facts about sub-atomic events such as both the position and momentum of an electron at the same time and thus can only calculate in terms of different possible quantum states, there are coherent deterministic interpretations of quantum mechanics. By appealing to the idea of a hidden variable we can say that whilst there may be more than one epistemic possibility there need not be more than one metaphysical possibility. The compelling arguments in favor of determinism then will persuade us not to rule it out in virtue of just one of the various possible interpretations of quantum mechanics.

However, the oscillating universe theory is currently ignored because there does not seem to be enough matter in the universe to slow expansion, and the energy within space, known as lambda, seems to provide a repulsive force which would mean the universe will continue to expand forever. However, it is by no means certain that the force of lambda is not one that could change. It is possible that the matter/anti-matter annihilation in the early universe that gave space its energy could separate back out into matter and anti-matter again, providing us with enough of a gravitational force to halt expansion and cause contraction.

Also, even if this universe, this particular finite region of space-time, were
to expand until all that is left is Hawking radiation, it does not necessarily mean that oscillation is impossible. Another region of space-time, that is another branch of the supposed mega-universe could contract and cause a chain reaction so to speak, such that the eventual singularity reached by one universe could draw in the other branches of the mega-universe. This is possible because the espoused mega-universe does not so much postulate a disconnected universe, but rather that during the early stages of expansion, known as the inflationary period, the expansion is so fast that it exceeds the speed of light. This, so far as we are concerned, creates a horizon that prevents us from ever communicating with the other branches of space-time.

Thus, although it is rather far fetched, I believe that in order to preserve what I take to be the inescapable conclusion of a self-causing universe, an oscillating universe is required. I have shown one way in which such a thing could occur that agrees with the current laws of physics and I think it provides a good example of the kind of philosophical implications that scientific theory can have. Hence we have seen how far, what appeared to be a very simple idea of space can take us and how important it is to clarify its nature if we are to understand what kind of universe we live in.

Nominated by: Prof. Tim Crane
University College London

Bibliography

Dr. Linda Smith, Cosmologist, UCL. Interview conducted, 25/11/98.


New Scientist, 11/4/98 “To Infinity and Beyond” (referring to the science of space and the possible future of the universe).

New Scientist, 24/1/98 “In the Beginning” (referring to self-causation).

Alexander, H.G. The Leibniz-Clarke Correspondence, Manchester University Press, 1956.


What is the relationship between faith and reason? This central epistemological question has come through the influence of analytic philosophy to entail a project of defining both terms. What is rationality? What is faith? At the outset, it is easiest to propose a definition of reason: a set of related skills that provide for us believing, knowing and understanding. Faith is more difficult and may perhaps be the source of all controversy and ambiguity in the debate. Is faith devoid of reason? Is it intuition? A substitute until reason prevails? A leap to belief after rationality has carried us some distance? Philosophers are divided into three common positions: that faith and reason are opposed, containing no coherence; that faith and reason are convergent (right reasoning would coincide with good faith); and that faith and reason are complementary.

The Ontological Argument has been proposed as perhaps the strongest argument for the convergence of faith and rationality. Anselm sought to establish a proof of God’s existence in the eleventh century and it has since been blessed with an illustrious history. While becoming a hallmark of Catholic theology, many influential philosophers such as Kant and Hume, among others, attempted to refute it. While many theologians and philosophers have considered the argument as an archaic failure, Alvin Plantinga, a contemporary theologian, has returned to the argument stating that it is difficult to show in its history of opposition when and where it was actually refuted. After following tradition in dismissing the original form of Anselm’s argument, Plantinga details his own version, rendered by James Sennet as follows:

(1) There is a possible world in which unsurpassable greatness is
exemplified.

(2) The proposition *a thing has unsurpassable greatness if and only if it has maximal excellence in every possible world* is necessarily true.

(3) The proposition *whatever has maximal excellence is omnipotent, omniscient, and morally perfect* is necessarily true.

Therefore,

(4) There actually exists a being who is omnipotent, omniscient, and morally perfect, and has these properties in every world. (Sennet 19)

In the context of Plantinga’s philosophy, it is strange to find an argument for God’s existence. Belief in God, according to *Reason and Belief in God*, is supposed to be properly basic. With this tension aside, the argument must still be examined for its validity and soundness. The conclusion, the existence of God, is an extremely ambitious goal. Does Plantinga’s ontological argument provide, as he claims that it does, “not the truth of theism, but its rational acceptability?” Anselm originally intended the argument as a reductio absurdum argument for non-Christians. He drafted premises that he thought were equally agreed upon by believers and unbelievers, intending that the conclusion be seen as rational by all. Plantinga’s Ontological Argument is considerably different; everyone does not equally endorse the first premise. He admits that it suffers from a similar fate as this argument:

(5) Either God exists or $7 + 5 = 14$
(6) It is false that $7 + 5 = 14$
(7) Therefore God exists. (Plantinga 71)

Plantinga admits at the end of the essay that probably no one would endorse the only premise of the argument, that it is possible that there is an instance of maximal greatness, unless they already agreed with its conclusion—the belief that God exists in the actual world. The first premise in the illustration above (and also in Plantinga’s Ontological Argument) is meant to be a premise that theists might endorse and non-theists find themselves handicapped to label “irrational.”

James Sennet, the author of *Modality, Probability, and Rationality: A Critical Examination of Alvin Plantinga’s Philosophy*, identifies that Plantinga’s argument really has only one objectionable premise. All the others follow necessarily because of their entailment by definition. However, Sennet mostly
avoids examining the original premise, claiming that argument employs “epistemic circularity.”

Is the first premise a good premise? Is the argument circular? Does the Ontological Argument, in Plantinga’s new form, provide for the convergence of reason and faith, in so much as it claims that theism is “rationally acceptable?” While an analysis of Sennet’s response reveals an untreatable flaw in his criticism, in the end Plantinga’s argument still fails to establish any additional rational acceptance for theism.

On Epistemic Circularity

Sennet provides two general analyses concerning Plantinga’s ontological argument. After translating it into modal logic symbols, he shows its validity. However, regarding soundness, James Sennet maintains that the ontological argument as proposed by Plantinga is trivial and employs “an epistemic circularity…that is unacceptable” (Sennet 26). Sennet’s accusation of epistemic circularity is found through his development of a notion called obvious entailment:

A proposition P obviously entails a proposition Q just in case it is impossible for a cognizer S to justifiably believe P and fail to be justified in believing Q…If P obviously entails Q, then S’s justifiably believing P and failing to be justified in believing Q is impossible precisely because understanding P is not possible without understanding that it entails Q. (Sennet 2)

The ramifications of this definition of obvious entailment will soon be discussed, but first a more general analysis is appropriate. Sennet has identified two parts of an argument, a proposition P and a mysteriously related proposition Q “obviously entailed” by P. The logical picture of their relationship might be depicted like this, where the symbol stands for “obviously entails”:

P Ê Q

One reason to identify the notion of obvious entailment is that a person might be said not to properly understand a proposition P, whether simple or complex, unless that person understands all of the statements that it obviously entails. Sennet offers the following example:

(8) Martians smoke marijuana.

obviously entails
Sennet suggests that a person that endorses the proposition “Martians smoke marijuana” must necessarily endorse the proposition “There is life on Mars.” It is impossible for a person to believe the first without also believing the second.

But what kinds of statements are obviously entailed, or necessarily assumed, by a proposition P? Sennet does not offer any list of the kinds of obviously entailed statements, merely constructing the rule: a proposition Q is obviously entailed “if it is impossible for a cognizer S to justifiably believe P and fail to be justified in believing Q.” This definition of obvious entailment at least includes all valid deductions. Moreover, a quick examination of Sennet’s Martian example reveals that obvious entailment includes more than valid deductions: obvious entailment is broader than logical entailment. Obviously, (8) is not solely adequate for deductive support of (9). What might these other kinds of statements like (9) be? Although Sennet never describes these statements outside of their unitary property, I believe they can be fairly interpreted as necessary and sufficient conditions. For example,

(10) Peter is watching A Field of Dreams.
(11) Peter can see.
(12) Something is causing the projection of A Field of Dreams.
(13) Peter is watching a movie.

Statements (11) and (12) are necessary conditions in order for (10) to have a true truth-value. In order for Peter to be watching, in the ordinary meaning of the word, he must have the ability to see. Something must also be the cause of the projection, whether projected on the wall or TV. Statement (13) is also an obviously entailed statement, but in a different sense. As it necessarily follows from (10) with the addition that A Field of Dreams is a movie, it is not part of the functionally minimized set of necessary and sufficient conditions, but should be included because obvious entailment includes all necessary statements.

Why, or how, are statements such as (9), (11) and (12) obviously entailed? Judging from Sennet’s Martian example, a cognizer must supply a set of necessary and sufficient statements that are commonly accepted as true. This will provide through deduction an explanation of the reason a certain proposition
Q must be true if proposition P is true also. For example, a cognizer should supply the following propositions in order for (9) to be deductively deducible from (8).

(14) Anything that smokes marijuana must exist.
(15) Martians are beings on Mars.
(16) If there is a being on Mars, then there is life on Mars.

With an understanding of the rule of obvious entailment, a description of the kind of statements obviously entailed, and a program through which the necessity of obviously entailed statements can be understood, it is possible to proceed to Sennet’s criticism of Plantinga. He proposes that it is possible to make a structural distinction between sound and unsound arguments according to the following rule (which just so happens to use the notion of obvious entailment):

It seems obvious to me that, if S justifiably believes P on [ground] G and P obviously entails Q, then it is impossible that (i) S fail to be justified in believing Q on G and (ii) S be justified in believing Q solely on G plus knowledge that P entails Q. If P obviously entails Q, then simple knowledge that P entails Q, added to G, cannot give justification for Q if G alone cannot justify Q. So, under these circumstances, S’s belief of P on G is not justified. If I believe [8] on G, and come to know that [8] entails [9] which G alone does not justify for me, then the simple fact that I understand [8] to entail [9] does not give me justification for believing [9]. (Sennet 2)

Sennet has developed his logical picture a little further, where is the symbol for “gives justifying grounds”:

\((G \rightarrow P) + (P \rightarrow Q)\)

The rule that has been introduced is host to all sorts of question, but before these questions can be asked Sennet provides intriguing explanations behind this analysis of an argument. He makes two zealous claims. He begins by stating that Q is not justified solely on the basis of the following true conjunction: G grounds P and P obviously entails Q. Examining the pictured relationship of the statements, this seems tantamount to destroying the foundations of
logic. This will be discussed shortly. The most important issue at hand (and the one that will be shown to be of most concern to Plantinga) is his second claim: “showing that G cannot justify Q for S…is sufficient to show that S does not justifiably believe P” (Sennet 2).

Despite Sennet’s insistence that “it seems obvious” that this rule is true, it is not obvious to me and I also believe it is false. However, while it is easy to notice the problems it introduces into logic, it is difficult to point out its flaw.

It is easily noted that Sennet’s position seems to introduce a logical absurdity. It was discussed earlier that his broad definition of obviously entailed statements (Q’s for shorthand hereafter) included at least all valid deductions from proposition P. By definition, valid deductions are necessarily true (if the premises are true). Of course, the necessity of deduction fits nicely into the definition of something that that cannot possibly fail to be true—which is true of all Q’s by proposition P. A problem thus easily presents itself:

\[(17) \quad A + B \]
\[(18) \quad (A \lor C) + (B \lor D)\]

From the rules of logic, (18) follows necessarily from (17) after a few implicit steps. As a valid deduction, (18) can be construed as a Q and (17) as a P. Let us also suppose that there are justifying grounds for (17). Substituting these specific propositions for Sennet’s general formula:

\[G \quad (17)] + [(17) \lor (18)]\]

If Sennet’s rule is to be applied, even though it seems (17) is properly grounded by G, this is not really the case if G can not also directly support (18). Therefore, an argument for (18) based on (17) grounded by G, coupled with the knowledge that (17) obviously entails (18), is not sufficient. This obviously seems to be a logical absurdity and that somewhere Sennet has gone astray. How can the obvious entailment from (18) affect the grounding of (17) by G?

It might be helpful to approach Sennet’s argument analysis from another angle. In ordinary language, he says that:

\[(19) \quad \text{To constitute good reason for its conclusion, an argument must be such that one can justifiably believe all the premises without understanding}\]
them to entail the conclusion. (Sennet 24)

The symbolic representation holds true for this non-technical version of Sennet’s rule of reason. However, this is even a more blatant rejection of fundamental logical principles. Nowhere does Sennet discuss “justifying grounds” (G) of the premises, but he speaks directly of the inference: in order for there to be good reasons for a conclusion, the premises (P) cannot obviously entail the conclusion (Q). If the picture of obvious entailment is right (PÊQ) then Sennet believes this inference is invalid even if the premises are true.

Despite the startling implications of Sennet’s rule, what would happen if one supposed that Sennet were right and attempted to take it seriously? Is there any sense that can be made out of it?

Before these questions are approached, much has been said of the form of Sennet’s criticism without explaining its application to Plantinga’s particular arguments. In Sennet’s words, his:

criticism of Plantinga will center on the fact that theistic beliefs crucial to Plantinga’s arguments are such that the grounds they have cannot justify the belief that God exists. Hence, such arguments are not good arguments for the rationality of theistic belief. (Sennet 3)

The symbolic representation of the relationship between theistic beliefs and theism is:

Theistic beliefs Ë The belief that God exists.

Sennet’s rule allows him to claim that no theistic belief is justified because the belief that God exists is unjustified. His accusation is one of epistemic circularity, and there does seem to be a hint of it in the ontological argument. He maintains that the ontological argument as formed by Plantinga, in which the belief that God exists in the actual world follows necessarily from the first premise concerning possible worlds, is constructed so as not to meet the requirements to be a “good reason for its conclusion” according to (19). In order to illustrate the problem in conclusions obviously entailed from premises, Sennet offers the following example:

(20) God exists and 2 + 2 = 4.
(21) Therefore, God exists. (Sennet 26)

Necessarily, the justification in believing (21) must be antecedent and prior to proper justification of (20). This, then, might be the missing clue in understanding how Sennet’s seemingly erroneous rule might actually be a brilliantly concocted blessing. Maybe the picture of the relationship between P and Q has been misrepresented from the outset. What kind of a statement is antecedently and independently related to a proposition P? Sennet, in his chapter on the ontological argument, states that:

The idea behind [19] is that an argument serves as good reason for its conclusion for S only if that argument can take S from a previous acceptance of the premises to a belief of the conclusion that is rational, but would theretofore have been unjustified. But if the conclusion expresses a proposition that S must understand to be entailed by one of the premises in order for S to accept that premise, then that argument cannot function as a good reason for the conclusion for S. There is an epistemic circularity here that is unacceptable. (Sennet 25)

This paragraph further cements that the symbolic depiction of Sennet has been woefully misconstrued (due to Sennet’s obscure explanation). A proposition P is justified by a conjunction of true premises. What might these antecedently and independently justified statements be? What might be fairly extracted from this excerpt is that at least some obviously entailed statements Q do not belong subsequent to proposition P. But if they do not come subsequent to P, where do they belong? What does this do to the relationship between P and Q? It has been shown that obviously entailed statements include at least the necessary and sufficient conditions for P. What if Sennet means by his whole analysis that these Q’s function to give truth value to P? The relationship might be symbolized in this manner, where Ú symbolizes “is one of the necessary and sufficient conditions of P”:

$Q \Ú P$

This picture image allows us inside Sennet’s world. If he is reread from the beginning and understood to mean that Q’s do not follow from P but precede them (antecedently and independently), then his system becomes entirely more alluring. An illustration and a more formal example can provide case
and point.

Take, for instance, a farmer crossing through his field one evening. He is out for a night stroll to enjoy the multiplying pumpkins in his patch and, as with all badly written stories, the cool breeze. He suddenly finds himself perceiving, through sight, a galloping horse and upon it a horseman—only headless. It is no coincidence that he holds this experience to be called grounds G. He naturally finds himself inclined to believe a proposition labeled P that the Headless Horseman is pouncing through his pumpkin patch. However, pursuant to James Sennet, he recalls that one of the necessary and sufficient propositions of such an event (an obvious entailment Q) is that the Headless Horseman exists. Of course, the farmer is much too familiar with the storybook legend and knows through his high-priced education that the proposition “the Headless Horseman exists” is false. This story provides an understanding of how grounds G might attempt to give justification for P but fail as it could not simultaneously support Q, a belief independently examined but not overturned by such a sordid sight.

A more formal demonstration is appropriate – akin to a proposition of this sort, where A and B are necessary and sufficient conditions for P:

\[(A + B) \supset P\]

Proposition P will have a true truth-value if A and B have been determined to have true truth-value as well. Sennet’s goal with his notion of obvious entailment is thus to note that a proposition P is a function of its necessary and sufficient conditions that must be established antecedently and independently.

How does this apply to Plantinga’s ontological argument? Admittedly, by both Plantinga and Sennet, the argument really has only one controversial premise: proposition (1) (as proposition P) claims there is a possible world in which unsurpassable greatness is exemplified. By design, the conclusion of the argument necessarily follows from this premise. However, upon examination, the conclusion is found to part of the set of necessary and sufficient conditions for (1) to be true, namely that maximum greatness is exemplified in every possible world, including the actual world. As exemplification of maximum greatness in the actual world is equivalent to God exists in the actual world, this proposition Q, according to Sennet, must properly be antecedently and independently established.

Concerning Sennet’s main criticism of Plantinga (that the belief of God
exists is unjustified and therefore theistic beliefs crucial to his arguments are also unjustified) this reforms the relationship between theism and theistic beliefs as:

\[
\text{The belief that God exists } + X \leq \text{ theistic beliefs}
\]

What is to be made of Sennet’s analysis? At first glance it seemed the equivalent of logical heresy but now it seems as though his thinking indeed is quite clear. I stated earlier that I thought Sennet’s rule of reason was false, but difficult to show where it falters. It is best to begin by noticing that the criticism of Sennet comes about when the relationship between P and Q is P \(\leq\) Q, and his genius when Q \(\geq\) P. I have also argued that I think necessary and sufficient conditions are an accurate interpretation of what Sennet means by obvious entailment. A quick journey into logic will reveal where his error lies. But first, another story.

This plot does not concern a farmer, but me. Right now I am sitting in my living room enjoying the aesthetics of the Christmas season around me, in particular the rainbow of lights on my Christmas tree. According to Sennet I can establish the proposition P, “The lights on my Christmas tree are on and beautiful” by appealing to its necessary and sufficient conditions Q’s: lights are on the tree, electricity is running through the bulbs, there are no burnt out bulbs in the string, the colors of the lights on the string are beautiful, and so on. These must be antecedently and independently shown to be true to validate the proposition, “The lights on my Christmas tree are on and beautiful.” But why? Do my grounds for such a proposition have to be limited to an examination of its necessary and sufficient conditions? Is this even what is ordinarily done?

The answer to the latter two questions is no. My grounds for P might be, “I see the lights on my tree and I like them.” What has gone wrong in Sennet’s rule of reason? A journey into logic reveals that the proper formulation for necessary and sufficient conditions Q’s and their entailed proposition P is:

\[
A + B \Rightarrow P
\]

In other words, the truth-value of P is not merely a function of the truth-value of A and B but also the reverse, the truth-value of A and B is a function of the truth value of P. If it can be shown that there can be grounds G for believing P independent from A and B, then P can be grounded and through deduc-
tion the truth of A and B can be shown also. The moral of the Christmas light story is that proper grounds for P do not have to include grounding the justification of every necessary and sufficient condition.

Thus, the structure of Plantinga’s ontological argument, in so much as the conclusion is obviously entailed by the first premise, is not open to Sennet’s criticism of epistemic circularity. If, in fact, there can be shown to be grounds G for the first premise then the argument can be sound.

**On Soundness**

Now free from the constraints of Sennet’s structural criticism and provided with a directive for establishing its soundness (searching for grounds for the first premise P), is it possible to appeal to any commonly held beliefs to justify belief in (1)? The project immediately becomes intimidating in the light of Plantinga’s admission of the problem detailed by statements (5)-(7): that a person will not be persuaded by the argument because, like the argument in (5)-(7), the ontological argument “obviously…isn’t a proof; no one who didn’t already accept the conclusion…would accept the first premise” (Plantinga 71). Unfortunately, Plantinga’s example is misleading. It seems as though he is accepting the criticism of James Sennet, that the conclusion is such that it must be antecedently believed before one will accept (1). Has Plantinga admitted that belief of God’s existence in the actual world is a necessary precursor to justification of (1)? That the truth-value of (1) is merely a function of what is already believed about what it obviously entails? I do not believe that this is a completely accurate reading of Plantinga. He does not go so far as to admit that belief in God is grounds for believing in the possibility of an instance of maximal greatness.

Perhaps he means that the two propositions, (1) and (4), tend to come as a pair. A person with an outlook that believes God exists is a person that is often simultaneously influenced in believing that it is possible that there be an instance of maximal greatness. What might this outlook be that gives grounds for (1) and through Plantinga’s argument gives some type of support for (4)? This is a difficult question, yet it remains true that if some grounds can be furnished for (1), the soundness of Plantinga’s argument is still possible.

Has Plantinga’s argument established the “rational acceptability” of theism? The rational acceptability of theism, because of the necessity of the deduction, is a function of the rational acceptability of the first premise. Plantinga’s claim is that the first premise is of such a nature that “there is nothing contrary to reason or irrational in accepting this premise” (Plantinga 71). How does
one establish this claim? This can only be done by examining the statement and all of its entailments to see if there is any justified reason to hold that any of these is contrary to reason or irrational, providing prima facie reason to doubt (1). If a teleological outlook does not justify (1), then what are the grounds or justifying evidence for (1)? The absence of observing grounds for (1) should not be construed as providing a prima facie reason to doubt (1). Any such construction would be an argument from ignorance. There also needs to be an examination of the statements obviously entailed by (1). Is there any justified reason to hold that one of them is contrary to reason or irrational? All the propositions obviously entailed by (1) are of the form: a being has maximum greatness in this possible world. The only one of these entailments available to observation is the one concerning the actual world. Is there any justified reason to hold that belief in God in this world is contrary to reason or irrational about belief in God in this world? However, now we have reached an unfortunate digression. The rational acceptability of the first premise will be the degree to which belief in God is rationally acceptable in the actual world. The irony I believe this presents, is that Plantinga’s argument will not add any additional rational acceptance towards belief in God then already directly exists without it, unless justifying grounds for (1) can be provided.

A last minor issue concerning Plantinga’s argument. Anyone who understands modal logic will understand the statement is either necessarily true or impossible. There is no middle ground for contingency. However, the first premise is written as if it were contingent. This is not a formal or informal fallacy but an underhanded maneuver by Plantinga. Any person unfamiliar with the rules of modal logic may not understand that the first premise, while true in a possible world, cannot be merely possible.

Plantinga’s ontological argument, by examining James Sennet, is not epistemically circular. However, Plantinga’s claim that the argument provides for the rational acceptability of theistic belief is true only if grounds can be offered to support the first premise. Otherwise, the claim for the rational acceptability of the first premise is dependent on the rational acceptability of belief in God established externally to the ontological argument, meaning that the argument proposed by Plantinga offers no additional rational acceptance.

If Plantinga is correct in discovering the fatal flaw in Anselm’s argument – and his own valid argument proves to be uneventful – what can be said of the ontological argument? Are there any sound versions of the ontological argument? Plantinga’s argument and the other versions discussed in his essay may be open to the criticism that they reflect the fading significance of the onto-
logical argument. As a contributor to discussing the relationship between faith and reason, the argument fails to be a prototype example of their convergence.

End Notes
1 Because I am unfamiliar with modal logic symbols, I assume Sennet’s analysis to be correct.
2 Sennet here uses the classical existential assumption. It is not true in all cases that claiming a theorized object has properties necessitates its existence.
3 X is all other necessary and sufficient conditions for that particular theistic belief.

Bibliography
Campbell, James. From Belief to Understanding. Australia: The Australia National University, 1976.
“A Fish in Water”:
Michel Foucault and Historical Change

Ian Hesketh
Okanagan University College, British Columbia, Canada

I would like to write the history of this prison, with all the political investments of the body that it gathers together in its closed architecture. Why? Simply because I am interested in the past? No, if one means that writing a history of the past in terms of the present. Yes, if one means writing the history of the present.1

Like Marx, Hegel, and many others, Michel Foucault saw definite shifts appearing in history. However, unlike Marx and Hegel, he was not concerned with production or the dominant political system of the time, he was concerned with episteme shifts. The term episteme is meant to portray the underlying construction of knowledge “that allow[s] different objects and different themes to be spoken at one time but not at another.”2 These epistemic shifts constitute Foucault’s account of historical change. The Order of Things will provide the focus for this analysis of Foucault’s view of historical change because it is in this study that a concept of ‘Universal History’ can be derived from Foucault’s thought. The argument of this paper is simple: Foucault’s analysis of the underlying, irrational structures of history is extremely useful because it challenges historians to look at the past on its own terms. His ‘failure’ to determine what causes epistemic shifts is also viewed by the present author as a positive attribute, despite obvious claims otherwise. However, in his attempt to point out the irrationality of history, and destroy any intellectual links with the Enlightenment, Foucault has essentially created a theory not unlike those he attempted to distance himself from.

Intellectual Background and Methodology

Michel Foucault was born in Poitiers, France, in 1926. Poitiers was also the birthplace of Jean-Paul Sartre in 1905. Both Sartre and Foucault studied Friedrich Nietzsche and were greatly influenced by him. They both engaged
in the philosophy of praxis, were outspoken intellectuals of French society and challenged social injustice. At times they did this shoulder-to-shoulder. However, Sartre’s later intellectual influences came, not from Nietzsche, but from Husserl and Heidegger, whereas Foucault remained a Nietzschean also heavily influenced by Batialle. While Sartre went on to be one of the main proponents of an existential philosophy, Foucault oscillated in and out of structuralism and is now considered a poststructuralist. For Foucault, Sartre’s existentialism was a philosophy of humanism and consciousness derived from the intellectual premises of the Enlightenment. Foucault rejected the intellectual foundation of the Enlightenment, believing that promoting the values of reason and humanism has only led to the creation of a form of disciplinary power that assumes the central role of humanity in the world. Foucault, therefore, could not adhere to Sartre’s existentialism because he viewed it as just another form of Western intellectual domination. As Arthur Herman makes note of:

Foucault rejected Sartre because his philosophy still contained the same bankrupt assumption about man as a self-regarding subject. The modern Western image of man, Foucault would assert, “is not a proposition to be defended but a product of social-historical processes,” that is to say, of capitalist civilization.

Furthermore, Foucault carved a space in the philosophical world that was all his own, especially through his critique of reason as a mode of domination. Although he was probably greatly influenced by Sartre, the generation that separated the two could have been a lifetime.

As well, Foucault is not a historian in the traditional sense of the word. His main purpose was to take a theme, like the human sciences, and show how it has changed throughout time – to construct a Nietzschean genealogy. His main purpose, then, was to begin with the present and move backwards until a difference could be located in the past. In so doing, Foucault could reveal how the particular aspect of the present has changed, evolved, and perhaps become something that is far removed from what it once was. As Foucault writes, “[i]t is no longer a question of judging the past in the name of a truth that only we can possess in the present, but of risking the destruction of the subject who seeks knowledge in the endless deployment of the will to knowledge.” Indeed, Foucault attempted to delegitimise the past while putting into question the present order of knowledge.

In the introduction to The Archaeology of Knowledge, Foucault set about
to chastise historians for their practice of ordering the past around their present state of knowledge. Foucault argued that historians have continually written about the past in a way that inevitably leads to the present. Foucault’s arguments supporting this are worth noting:

Continuous history is the indispensable correlative of the founding function of the subject: the guarantee that everything that has eluded him may be restored to him; the certainty that time will disperse nothing without restoring it in a reconstituted unity; the promise that one day the subject – in the form of historical consciousness – will once again be able to appropriate, to bring under his sway, all those things that are kept at a distance by difference and find in they might be called his abode.8

Indeed, Foucault felt that historians have made the past into a comfortable place that they can enter at ease – the past has become the historian’s ‘abode’. It was Foucault’s argument that the past is not a place of rationality, where everything and everyone make sense. Furthermore, it is by exploiting difference that continuous history can be avoided and by that very premise Foucault’s genealogy is an attempt to discover change and explain change. “My aim is to uncover the principles and consequences of an autochthonous transformation that is taking place in the field of historical knowledge.”9

Foucault, then, is an historian of discontinuity.10 As Mark Poster illustrates, “[Foucault] does not narrate the evolution of the past; he does not tell the story of how ‘the seamless web of yesteryear’ leads slowly and inexorably into the present.”11 Furthermore, because Foucault is mainly concerned with searching out discontinuities, it is fairly easy to extract a view of historical change from Foucault’s writings. Although this essay will concentrate on his most comprehensive book – The Order of Things – because his view of historical change is most obvious in it, is possible to see his view of historical change, whether in Discipline and Punish or Madness and Civilization.

The Order of Things

The first page of The Order of Things outlines an old taxonomy of animal classification from a Chinese encyclopedia. The classification system that the Chinese employed when the encyclopedia was written does not conform to our present state of knowledge. Foucault felt that the taxonomy, instead of being chaotic and ridiculous, challenged the limits of modern knowledge and therefore prompted him to analyze how societies think certain things at one time and not at others.12 This allows Foucault to reflect on the question of how societies organize knowledge into an identity – a coherent orderly system –
from a past that appears to have little in common with the present.  

Through Foucault’s archaeological study, he finds that there are two great 
discontinuities “in the episteme of Western culture: the first inaugurates the 
Classical age (roughly halfway through the seventeenth century) and the sec-
ond, at the beginning of the nineteenth century, marks the beginning of the 
modern age.”  

These discontinuities mark periods in which a transformation 
occurring in the way people not only thought, but understood themselves, their 
culture, and their entire mode of being. Foucault defines discontinuity as “the 
fact that within the space of a few years a culture ceases to think as it had been 
thinking up till then and begins to think other things in a new way.”  
The epistemic shift into the Classical age marks the first rupture in Western thought. 

Although the human sciences had not been developed yet, Foucault at-
tempted to deconstruct that process, searching for an origin. Prior to the Clas-
sical age, Foucault argues, thought was based on resemblance. Up to the end 
of the sixteenth century, the interpretation of nature and texts were based on 
resemblances and similarities. Taxonomy was based upon whether or not the 
plant or animal resembled theological symbols and virtues. Foucault also dis-
covered that the grammar of the 16th century was constructed the same way 
nature was ordered. 

Words grouped syllables together, and syllables letters, because there are vir-
tues placed in individual letters that draw them towards each other or keep 
them apart, exactly as the marks found in nature also repeal or attract one 
another. The study of grammar in the sixteenth century is based upon the 
same epistemological arrangement as the sciences of nature or the esoteric 
disciplines. 

The underlying factor in the episteme of the sixteenth century is the fact that 
God was the source of all knowledge and placed man at the center of the uni-
verse. However, this would all change in the first great discontinuity Foucault 
envisaged. 

Foucault claims that at about 1660 a great discontinuity engendered the 
Classical period (the Enlightenment) in which thought was structured by rep-
resentation based on identity and difference instead of resemblances, en-
gendering the practices of general grammar, the analysis of wealth and natural 
history. The creation of the verb in language, the creation of value in the 
analysis of wealth, and the creation of the character in natural history formed 
the basis of the episteme of the Classical age. However, God was still center
of the universe, but His position was beginning to erode as man was beginning to reflect on himself.

The discontinuity that marks the shift into modernity at the beginning of the nineteenth century sees the creation of the natural sciences. Philology, biology, and political economy began to “take up an area where [general grammar, natural history and the analysis of wealth] did not exist... The object of knowledge in the nineteenth century is formed in the very place where the Classical plentitude has fallen silent.” However, it is the creation of History that shows the decisive break with the Classical period into modernity. History took the place of Order in the Classical age. “History gives place to analogical organic structures, just as Order opened the way to successive identities and differences.” Furthermore, God had lost his place as the anchor of human knowledge. For Nietzsche, this is the era of the death of God; for Foucault it is the birth of man. In other words, prior to the nineteenth century, “man did not exist (any more than life, or language, or labor).” Therefore, man is a relatively new creation as the death of God and the creation of the human sciences have left man to become both the subject and object of study. And what underpinned all of the new sciences is that they sought out an origin and an end for the object under study, which became humanity.

The Death of Man

The epistemic shifts that compose Foucault’s view of historical change would not be complete without his vision of the future. However, Foucault appears to distance himself from other philosophers of history because his primary methodological stance was to analyze the past entirely removed from the modern epistemological structure. Foucault argues in Discipline and Punish that he refuses to write “a history of the past in terms of the present,” arguing instead for a “history of the present.” Foucault even argues that Marxism is merely a production of the nineteenth century episteme:

At the deepest level of Western knowledge, Marxism introduced no real discontinuity; it found its place without difficulty, as a full, quiet, comfortable, and goodness knows, satisfying form for a time (its own), within an epistemological arrangement that welcomed it gladly (since it was this arrangement that was in fact making room for it) and that it, in return had no intention of disturbing and, above all, no power to modify, even one jot, since it rested entirely upon it. Marxism exists in nineteenth century thought like a fish in water: that is, it is unable to breathe anywhere else (my emphasis).

Furthermore, it is obvious that Foucault does not feel his own theory, or lack of
theory, fits into this modern epistemological arrangement. However, it is possible to see that Foucault falls into the same trap as modern philosophers in their view of the present as being some point of precipice, or as Foucault claims, a discontinuity, which his idea of the death of man represents. Foucault’s call for the “history of the present” may distance himself from using the “terms of the present” but he seems unable to avoid the same errors of methodology that he claims of the Enlightenment progeny like Marx.

Foucault argues that as the Classical period marks the ordering of society around the inquiry into God and modernity marks the death of God and an inquiry into man, post-modernity (this is not his term), will see a death of man. *The Order of Things* leads the reader through these two great discontinuities in Western epistemology and opens the door to a new rupture that Foucault can see in the near future. Foucault essentially argues that the emergence of history is a sign of a new rupture that will appear and render the historical mode of thinking foreign to modern thought.23 “By revealing the law of time as the external boundary of the human sciences, History shows that everything that has been thought will be thought again by a thought that does not yet exist.”24

One can therefore see that Foucault’s archaeology of the human sciences is based on the rupture he sees occurring before him. He ordered the past around the environment of the present. Foucault, then, does not break out of the modern epistemological framework. The death of man seems as prophetic as Marx’s withering of the state in favor of a global socialist society. He used his archaeological analysis to show that “man is an invention of recent date. And one that is perhaps nearing its end.”25 Furthermore, he placed a structure on the past in order to have it lead up to his conclusion of the present. Perhaps his critique of Marx could be just as easily placed upon himself – he is still looking to the past in order to explain the present as a distinct turning point in history.

In an interview, Foucault claimed that philosophers tend to engage in “harmful habits” when reflecting on the past with the present in mind. Foucault asserted that one “harmful habit” is “the analysis of the present as being precisely, in history, a present of rupture, or of high point, or of a returning dawn, etc. The solemnity with which everyone who engages in philosophical discourse reflects on his own strikes me as a flaw.”26 Foucault would have to be the first to admit that he falls into this trap, especially when he claims that man will be erased “like a face drawn in the sand at the edge of the sea.”27 Perhaps Foucault’s archaeology of the human sciences is really only a description of the present through a facade of historical processes of discontinuity. His view
of historical change fits right in with the changing world he sees around him. He is therefore, criticizing himself in the interview: “I think we should have the modesty to say to ourselves that...the time we live in is not the unique or fundamental or irruptive point in history where everything is completed and begun again.” This does not by any means downplay the significance of Foucault’s contribution to the study of historical change, though. However, understanding that he perhaps falls into many traps he chastises historians for doing in the introduction of The Archaeology of Knowledge, it becomes much more easy to question his archaeological studies.

**Sartre’s Criticism**

Although there are many differences separating Sartre and Foucault, they both adhered to the concept of discontinuity in history. Sartre, who had Marxist leanings, understood and readily agreed with Foucault on the concept of a culture’s abandonment of one way of doing things for another in a very short period. Comparing Marx with Foucault, one can see that they both saw the same time periods as distinct discontinuities, although in a somewhat different light. Marx focussed on a shift in the mode of production during the discontinuities and that it was a conflict between two differing ways of producing that engendered the change. This is obviously where Marxist thought departs with any relation to Foucault. Although Foucault would definitely agree with the Marxists that the mode of production changed a great deal, he would have trouble agreeing with the notion that it engenders the epistemological shift. This is where we can find Sartre’s Marxist critique of Foucault and The Order of Things.

After the publication of The Order of Things, Sartre acknowledged the accomplishments of Foucault but had one major criticism. Sartre argued that “Foucault avoided the question of history, how one episteme is supplanted by another.” Sartre felt that a causal relationship should be described, or else any findings that Foucault maintains are virtually irrelevant. Whether or not causation is “the question of history” is certainly debatable, and indeed, part of the problem, but Sartre’s criticism is quite relevant. In The Order of Things Foucault takes the reader on an exhaustive journey, exploiting the dominant episteme of each time and even gives precise dates as to the ruptures in each, but he avoids purporting any gesture of causation. As Geoff Bennington and Robert Young mention, Foucault, “may, in principle at least, be able to explicate everything within the epistememes, but can say nothing more powerful than ‘it happened’ about the shift from one to the next.”

Foucault definitely does not hide the fact that he does not provide any
causal relationship. He claims, though, that to try to discover what causes one *episteme* to be revoked in favor of a new one would be beyond our present comprehension of our current *episteme*. Foucault argues that the discontinuities “could be appraised and measured only after a quasi-infinite investigation concerned with nothing more or less than the very being of our modernity.”

Furthermore, any attempt to discover the “fundamental event” that leads to the discontinuity would simply be based on the epistemological structure of the time and therefore continue the methodological problem Foucault directs at contemporary historians. Moreover, the reason Foucault did not provide a causal analysis is for the same reason he depreciates the value of Marxism. He wanted to avoid “the search for an original foundation that would make rationality the *telos* of mankind, and link the whole history of thought to the preservation of this rationality, to the maintenance of this teleology, and to the necessary return to this foundation.”

Foucault, then, could easily refute Sartre’s argument because it implies everything Foucault was trying to avoid. However, Foucault does acknowledge in *The Archaeology of Knowledge* that a fuller analysis of the transformation “must be left until a later study, in which [he] can give it [his] full attention.” Moreover, Foucault even realizes that this is an issue that his readers want him to engage, but for some reason that “later study” never comes about. Obviously Foucault’s comments were for the purpose of silencing his critics. He did not want to maintain a causal analysis because that would lead to a reduction of possible inquiries. In other words, Marx claims that the changes in the economy engender all others, but he and his followers are constrained to this explanation for every change that occurs. Foucault does not fall into the trap of reducing all possibilities to a single world-view that could explain anything and everything. That is perhaps Foucault’s greatest strength. Because Foucault merely attempted to present the discontinuities, his analysis cannot be blinded to other aspects that do not fall into the boundaries of a specific theory. Foucault would argue that Marxists really cannot understand all of the changes that take place because of their deterministic views and therefore they fail to recognize aspects that would disprove that world-view.

Therefore, by giving absolute answers, he would have undermined everything that he stood for. Foucault realizes that he is not describing how change occurs and furthermore has no desire to. Foucault wants to leave his book open to interpretation. “This book is not simply the object that one holds in one’s hands; and cannot remain within the little parallelepiped that contains it: its unity is variable and relative.” The fact that Foucault wanted his book to
be interpreted by the individual reading it is his greatest strength. He does not want to provide absolute answers. In doing so he would be repeating all the problems he finds with modern philosophy.

**Foucault’s Contribution**

By avoiding a claim to a causal relationship, Foucault avoids examining the history of the world as a *telos* of humanity. Furthermore, Foucault shows us that history is by no means rational. His strength is in the fact that he distanced himself from others in that he did not analyze history in an attempt to rationalize it – he did not link concepts to concepts and events to events. On the contrary, Foucault searched for difference and sought out the obscure – that which contemporaries would have a difficult time comprehending – choosing to locate the underlying, irrational link between statements, avoiding linear and diachronic ordering. Foucault denied the role of the Enlightenment as providing absolute answers. Reason and rationality do not entirely constitute reality. Indeed, these ‘enlightened’ notions are new and spread out forms of domination, power and knowledge. However, is Foucault really that far removed from his ‘enlightened’ forebears?

Jean-Francois Lyotard claimed that *The Postmodern Condition* is “increduity toward metanarratives”\(^3\) and Foucault would certainly agree. However, Foucault has essentially created a metanarrative of history. He produced a concept of historical change mysteriously similar to those philosophers of history grounded in the intellectual remnants of the Enlightenment. His search for the irrational is rational; his search for a lack of progress is progressive; his denial of the Enlightenment confirms it. Foucault, no less than Marx, is a fish in the Enlightenment’s stream of philosophers of history.

Perhaps this analysis of Foucault is correct, but what disturbs the present author is that I have essentially committed all of the errors that Foucault claims intellectual historians continually make.\(^3\) I have placed a structure on Foucault’s thought that he would argue is simply not there. I have viewed his analysis of history as a continuous stream of thought, devoid of contradictions and heresy. I did not search for statements in his thought, but ideas. I turned his complex historical investigations into a metanarrative that he would argue quite simply does not exist. Perhaps I am the follower of the Enlightenment, not Foucault. Therefore, my findings were predetermined by my methodology, by searching for an overarching link – a *telos* in Foucault’s archaeologies. Keeping that in mind, one must wonder if it is possible to represent someone’s ideas in a coherent way without conforming to the dominant *episteme* of the time. If anything, this is what Foucault has taught us. But is it possible, then, as Foucault
believed, to write the “history of the present” without “writing a history of the past in terms of the present”?

End Notes

5 This point is made in Gilles Deleuze, Foucault trans. Senn Hand (Minneapolis: University of Minnesota Press, 1988), 1-23.
6 Poster, Foucault, Marxism and History, 64.
9 Ibid., 15.
10 Foucault did not necessarily agree with this label. He certainly did not set out to find discontinuities but it is hard not to see that his method of historical analysis—whether genealogical or archeological—would inevitably lead to discontinuity. His argument against this is worth noting: “This business about discontinuity has always rather bewildered me. My problem was not at all to say, ‘Voilà, long live discontinuity, we are in the discontinuous and a good thing too,’ but to pose the question, ‘How is it that at certain moments and in certain orders of knowledge, there are these sudden take-offs, these hastenings of evolution, these transformations which fail to correspond to the calm, continuanist image that is normally accredited?’” Foucault, Truth and Power, in Power/Knowledge: Selected Interviews and Other Writings 1972–1977 ed. Colin Gordon (New York: Pantheon Books, 1980), 111–2.
11 Poster, Foucault, Marxism and History, 75.
14 Foucault, The Order of Things, xxii.
15 Ibid., 50.
16 Ibid., 35.
17 Ibid., 200–204.
18 Ibid., 207.
19 Ibid., 219.
20 Ibid., 344.
21 Foucault, Discipline and Punish, 31.
22 Foucault, The Order of Things, 262.
24 Foucault, The Order of Things, 327.
25 Ibid., 387.
27 Foucault, The Order of Things, 387.
28 Foucault, Politics Philosophy Culture, 36.
29 Poster, Foucault, Marxism and History, 5.
31 Foucault, The Order of Things, 221.
32 Foucault, The Archaeology of Knowledge, 13.
33 Ibid., 65.
34 Ibid., 23.
36 Mark Poster makes a similar comment in his analysis of Foucaultís thought in Cultural History and Postmodernity: Disciplinary Readings and Challenges (New York: Columbia University Press, 1997), 152.

Bibliography


Consciousness, Oppression and Resistance: 
A Hegelian Critique of Marx’s Assumption of Proletarian Revolution

Anthony Reeves
University of Portland

“Marxism teaches that exploitation and degradation somehow produce resistance and revolution. It’s hard to say why. What I’ve learned from women’s experience with sexuality is that exploitation and degradation produce grateful complicity in exchange for survival. They produce self-loathing to the point of extinction of self, and it is respect for self that makes resistance conceivable…”

-Catherine Mackinnon, Feminism Unmodified

As Marxist scholar and feminist Catherine MacKinnon notes above, Marxism does indeed seem to teach that oppression and exploitation are the central factors in bringing about ‘resistance and revolution.’ Marx appears explicit enough in his elucidation of how capitalism and the early industrialism of the 19th century was oppressive to the proletariat. However, he seems much less clear as to how this oppression necessarily brings about a proletarian revolution – he assumes that exploitation itself will spark a resistance among the exploited. In light of what MacKinnon notes above in relation to the exploitation of women and more so in relation to what Marx’s predecessor Hegel had to say about self-consciousness and agency, this assumption on Marx’s part seems ill-held.

To examine this fully, it is first necessary to understand how Hegel believed that individuals come to independent self-conscious agency, that is, acting upon one’s own initiative. This will mainly be examined through Hegel’s account of “Lordship and Bondage” in the Phenomenology of Spirit. Secondly, a look at Marx’s account of alienation and how, from a Hegelian standpoint, this appears to prevent for the industrial laborer a natural emergence of self-
consciousness and independent agency. After this, the implications this has for the possibility of a proletariat revolution arising out of oppression will be laid out as well as what this entails for oppression in general. At the end of this examination I think it will be clear that, at least from the standpoint of Hegel, degradation itself does not imply resistance, but in fact perpetuates the very processes which produce it.

From the standpoint of Hegel, human consciousness does not begin as self-consciousness, but rather self-consciousness emerges out of recognition by others. “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.” Hegel argues this is simply because to be aware of something, it appears that one must be removed from that which they are conscious of. For example, I am aware of this cup of coffee because I see, taste, and feel its warmth and this awareness arises out of my removal from it, my ability to ‘view’ it with my consciousness. My consciousness, on the other hand, is “immediate” to me; that is, I cannot view my consciousness because it is the very thing that is viewing. Thus, it is necessary for another consciousness to recognize my own consciousness, and one which reflects my existence back to me, for me to attain consciousness of myself.

Prior to this self-consciousness, human consciousness is entirely within a state of “sense-certainty,” a form of consciousness where the individual is only aware of its environment and the environment’s ability to fulfill her natural desires. The individual is not aware of herself as a subject experiencing the objective world, but instead is only aware of the environment itself. As Jeffrey Gauthier put it in the example of a newborn baby, “The newborn baby displays awareness of its world as it reaches for an object, struggles to suckle at its mother’s breast, cries in response to gas in its stomach, and so on. In the beginning, however it is not aware of its awareness of those things…it lacks self-consciousness.”

The full implications of this initial lack of self-awareness for human consciousness will be made apparent as Hegel’s account of the emergence of self-conscious agency is further explored. To illustrate how self-identity is born out of this “sense-certainty,” Hegel offers the parable of the master and the slave in the section of the *Phenomenology of Spirit* entitled “Independence and Dependence of Self-Consciousness: Lordship and Bondage.”

In this parable, two ‘consciousnesses’ in some primitive historical period, though Hegel does not intend that this was necessarily and actual occurrence in human history, encounter one another and “recognize themselves as mutually
recognizing one another.” This is uncomfortable for each consciousness because each is being objectified by the other, that is, each consciousness is gaining a definition of itself from the other consciousness. This is discomfiting because neither has control over the way in which the other objectifies him and, since each consciousness wishes to control their own self-identity, this leads to armed conflict between the two.

Furthermore, by showing a disregard for one’s own life through violent conflict, each consciousness supercedes the immediate objectification of the other by demonstrating that they are more than mere living things. As Hegel puts it, “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. And it is only through staking one’s life that freedom is won; only thus is it proved that for self-consciousness, its essential being is not [just] being…”

Thus, not only must I prove myself to the other, but I must show, through risking my life, that I am not merely a physical body that can be objectified by the other.

This encounter could easily lead to the death of one of them, especially since each is willing to risk their lives to escape objectification. However, this does not occur because, as Hegel scholar Charles Taylor summarizes, “one side gives in, recognizes its attachment to life, and becomes subjected to the other. The victor spares the vanquished in order to make him a slave.” The loser quickly realizes that its consciousness is dependent on life and submits, and the winner does not kill the loser, for the loser is the very witness to his own victory and is the very consciousness which can reflect back to the winner his own superiority. Thus, one consciousness becomes the lord, or master, and the other the bondsman, or slave.

The relationship that develops between the master and slave, and the slave’s eventual triumph over the master, is essential to understanding Hegel’s concept of self-conscious agency. As noted above, the slave becomes the witness to the master’s superiority. But, by this very token, the master’s self-concept is dependent upon the slave’s recognition, a recognition that is never fulfilling for the master since it comes from a defeated consciousness. More simply, how can the master possibly trust the recognition of a lesser being? The master understands that the slave will reflect back unto him, the master, whatever he, the master, wishes; thus the veracity of what the slave tells the master about himself is always in question for the master. Thus, the master never attains satisfactory self-definition for himself.

More important is the situation of the slave. For Hegel, the consciousness
of the slave was fundamentally changed by the fear of death that was incurred during the conflict with the master. This fear of death brought about a profound self-recognition on the part of the slave that was affirmed by the overpowering presence of the master. For the slave, consciousness had an “absolute melting-away of everything stable” which is the “essential nature of self-consciousness, absolute negativity…”

Also, the slave must labor in service of the master. In doing this, and transforming the material world, the slave imprints his own ideas upon the world. Thus, by shaping the natural environment to his own will, the slave creates a reflection of himself and since, as noted above, the slave has achieved self-consciousness, he can attribute this reflection to himself. That is, the slave gains an independent identity from his labor because he recognizes the products of his labor as expressions of himself. “Through this rediscovery of himself by himself, the bondsman realizes that it is precisely in his work...that he acquires a mind of his own.”

What the slave achieves in this is free thought, ability to act on that thought, and thus, self-conscious agency. The slave realizes that he can transform the world according to concepts and thus that he is a thinking being. This allows the slave the ability to, as noted by Taylor, “create models, archetypes, and then change things to conform to them.” Essentially, the slave has overcome the master because he, the slave, having gained an independent identity, can execute actions of his own rationality. Furthermore, the once slave now has the ability to continue to define himself independently of others through his labor.

It is important to emphasize here that for Hegel, both self-recognition and self-expression through labor are necessary for this permanent rational self-agency to exist. Without self-recognition, which for the slave was gained ‘socially’ through the conflict with the master, one would be doomed to never recognizing one’s labor as one’s own. Without self-expressive labor, which the slave gained through his labor, one’s limited self-consciousness would fade and an identity independent of others would be impossible.

Thus far, a review of the emergence of self-consciousness as given by Hegel has been done. What will be important to remember as this investigation continues into Marx is that, for Hegel, true and lasting self-identity is only possible through self-expressive labor and that this sort of self-identity is necessary for independent agency – that is, action induced by one’s own rationality. The violent and socially created identity that is found within the parable, the violent clash that initially took place between the master and slave,
Hegel limited to primitive humanity where there were not permanent social institutions to deal with conflicting ‘consciousnesses.’ In modern societies, Hegel would assert, such social conflict would be mediated in different ways.\textsuperscript{18} Thus, it will only be necessary to remember, as far as this analysis goes, that consciousness is socially created and that independent identity and agency are contingent upon individual material creativity.

Karl Marx’s understanding of human identity is in fact very similar to that of Hegels. For Marx, labor is the natural activity of human beings and it is the way in which humans express and identify themselves. It is by shaping the world around her that an individual is able to come to some sort of identity and consciousness of what herself is. This is not limited to physical production, any sort of activity in which the individual is able to affect the world in a personal sort of way, whether intellectual, physical, or otherwise, is labor. In fact, Marx believed that this was the only way in which an individual could identify herself. As Marxist scholar Erich Fromm notes, “For Marx, man is alive only inasmuch as he grasps the world outside of himself in the act of expressing his own specific human powers, and of grasping the world with these powers. In his productive process, man realizes his own essence, he returns to his own essence.”\textsuperscript{19}

There is another aspect to this as well. For Marx, self-expressive labor is what defines the human species and separates it from other forms of life. Marx, in his \textit{Economic and Philosophic Manuscripts of 1844}, notes that “an animal produces only what it immediately needs for itself or for its young…It produces only under the dominion of immediate physical need, whilst man produces even when he is free from physical need and only truly produces in freedom therefrom.”\textsuperscript{20} Thus, for Marx, one is human only insofar as they are freely expressing themselves through their labor.\textsuperscript{21}

Here it is possible to understand what Marx intended by alienation. Alienation in general is the “tearing away of man from the object of his production”\textsuperscript{22} where one is no longer identifying herself with the product of her labor. The worker views what she is and what her products are as fundamentally distinct, the object of her labor being completely alien to her. As Fromm puts it, “Alienation is essentially experiencing the world and oneself passively, receptively, as the subject separated from the object.”\textsuperscript{23} No longer does one seek production as a means of identifying oneself, for labor has lost its validity in providing a reflection of what one is. Thus, for Marx, one is experiencing such alienation is unable to achieve authentic self-identity.

More specifically, alienation occurs on four different levels for the industrial
laborer. The first, estrangement from one’s product, occurs where a worker’s labor is objectified in his product, that is, the worker places his essence through labor into a product, and then this product is taken from worker and turned against him. Thus, the more a worker produces under industrial conditions, the more he experiences a “loss of reality” for that very thing which defines the worker, his labor, is continually being taken away from him and turned against him through the forces of political economy.24

The second type of alienation is alienation from productive activity. Since the worker does not choose her productive activity, it is not self-expressive for her; she cannot find identification through labor, because the labor itself and the type of labor is forced. Thus, labor becomes not a satisfaction of the need of self-expression, but a means to satisfy other needs (food, shelter, etc.) and man is reduced to his animal functions – satisfying these functions become ultimate ends.25

The third is estrangement from humanity’s species being. As noted above, Marx understands man’s essence as a creative laboring species. This is degraded as laboring activity becomes not a self-expressive and defining activity, but a means to more animal-like ends. Also, Marx believed that humans naturally have all of humanity in mind in their actions. Capitalism, on the other hand, forces us to be competitive with our fellow humans and thus is, again, alienating.26

Lastly, in relation to the third level of estrangement, since humans are estranged from their species, competition arises between them. “An immediate consequence of the fact that man is estranged from the product of his labor, from his life-activity, from his species being is the estrangement of man from man.”27 In the capitalist system, this is not only evident in the relations between the bourgeois and proletarian classes, but even between the workers themselves as they compete against one another for the same jobs.

Marx understood alienation as the fundamental exploitative and oppressive feature of capitalism. It is only when another owns or controls the means of production that the product of one’s labor can be taken from them and turned against them and that labor can be controlled in such a way as to stifle the individual creativity of the worker. This sort of systemic model fundamentally degrades the worker, for Marx, as she is unable to do that very thing that essential to her species.

There is a way to free humanity from such a plight, and that is the positive transcendence of private property. Since alienated work is labor that must belong to another, private property is the collection of this alienated labor.
Thus, its positive abolishment is “the real appropriation of the human essence by and for man.”

Here seemingly arises the contradiction in Marx from the standpoint of Hegel. Marx agrees that self-expressive labor is integrally connected with self-identity, yet he assumes that a solution to a degrading system will arise out of a class, the proletariat, which has been denied their self-identity by the economic system. From Hegel’s standpoint, the proletariat, having been denied expressive labor by capitalism, would be unable to engage in the independent agency necessary for a solution because genuine self-identity, which is necessary for such agency, has been denied to them. Resistance to this sort of oppression seems quite unlikely considering that first of all, the individuals of the proletariat would have to recognize themselves as individuals deserving something better, and two, would have to believe that they could transform the world according to their own rationality. These both, Hegel would argue, are contingent upon a strong sense of self, that which Marx himself has shown is nearly impossible for the laborers of an industrial society to possess.

To look at the first point further, if Marx’s elucidation of alienation is to be taken seriously, it appears that by his own account he has denied the working class the possibility of ‘human’ agency. For if a laborer can be denied his species being by destroying labor’s expressive qualities and making it into a means to more animalistic ends (again, food, shelter, etc.), it would seem that the laborer is in fact reduced to an animal. Marx himself even seems to hint at this when he comments that “estranged labour tears from him [the worker] his species life.” How then, if the proletariat is reduced to such a state, are the workers supposed to face a system of exploitation that can only be ideologically overcome rationally? More simply, as long as the degraded working class is provided the basic means of sustenance and its animalistic needs are taken care of, why would they even think to change the system or believe that the system could in fact be changed?

To put this same criticism into a Hegelian model, if society denies the workers their ability to gain self-recognition through their labor, how can one assume that they have left the realm of “sense-certainty?” Since genuine identity for Hegel, as shown above, is gained through expressive labor, then who is to say that the workers have left the state of consciousness where one is aware of one’s environment insofar as it is fulfilling of a particular desire? Thus, if the economic system provides for these desires, it appears that the proletariat has little reason to rebel.

Of course, workers do have contact with each other, which would seem to
provide a basis for self-consciousness for each individual gains social reflection from others of his class. However, the content of this reflection, the information reflected back to the subject, seems limited to the overall state of the proletariat in general, for this would be all that the proletariat would seemingly understand - a state defined by, as shown by Marx, animalistic fulfillment of desire. For example, I could be the son of a steel-worker and see that all my parents do is eat, sleep, watch TV, and have sex. From this, I would essentially learn that is the sort of existence I should expect and it is only conceivable that others of my class would give this same expectation, or ‘reflection,’ to me. There is no reason to think that I would understand that I deserved and should strive for more.

The second criteria for resistance, believing that one can through their rationality shape the world, also seems defunct for the proletariat from Marx’s analysis. Since the workers are denied their fundamental expressive means in relation to the objective world, they are forced, as noted by Erich Fromm above, into a state of passivity and receptivity. It can hardly be assumed, from a Hegelian stance, that individuals inherently understand that they can shape the world to their will. As Hegel notes in his parable, it seems to be the case that individuals learn that they can shape the world according to their ideas by actually doing so; the slave only gains his independence through his actual expressive formation of the world. For an alienated laborer, this concept would appear to be completely foreign, for such a worker would never engage in the objectification of their ideas. Thus, for the worker, engaging in some sort of broad formative activity, such as the creation of a communistic society, would seem impossible.

Thus, Marx’s assumption that the general oppression of the proletariat will necessarily bring about revolution hardly seems justified. It appears that capitalism itself takes from the proletariat the very tools it needs to overcome capitalism. In other words, independent agency, that is having a strong concept of self and believing that one can change the world to fit this idea, seems to be fundamentally denied to the industrial labor in that the modes of production alienate the worker from the very thing which allows self-definition. In fact, the oppression provided for in capitalism seems to be exactly that which allows the same oppression to continue - the alienation of the laborer from their labor. In the capitalist framework, it appears that the very degradation of the worker is what allows the systematic oppression of that class.

To take this to a broader level, this appears to point out something about oppressive practices in general. By the very terms degradation, oppression,
and exploitation, it is understood something essential is being taken from someone. By the very act of exploitation, for example, one gains an advantage over another because one is gaining something that the other no longer has. In fact, the exploiters are always going to have the advantage for they are always going to have a greater means of enforcing their will since the exploited are, well, exploited and the exploiters retain the fruits of that exploitation. This is true not just in terms of material circumstances, i.e. one person or class having more money or power than another, but, as I think has been shown, in terms of consciousness and self-identity as well. It appears that acts of exploitation are by their very nature self-perpetuating by their encouragement of docility on the part of the exploited and their empowerment to the exploiters.

Is it thus hopeless from such a viewpoint? It is honestly hard to say, though I don’t think one could argue conclusively either way with what here has been examined. What I think is most important to understand from this specific analysis is that, apart from what a libertarian might argue, the oppressed cannot be expected to naturally undertake the means necessary to engage their freedom and to achieve their liberation. It requires something other, literally an other for Hegel, to give the oppressed an identity apart from that which they gain in their state of oppression. It is necessary to give to the exploited, the worker in this specific example, that which was taken away from them in the process of exploitation – in this case, the worker’s expressive labor.

Again, it is important to keep in mind that the purpose of this paper is not to specifically attack the assumptions of Karl Marx, for of course Marx’s perspective on the situation of the proletariat is far more complex than what has here been given. What makes the proletariat so unique for Karl Marx is that a piece of the educated class, i.e. a possible other for the proletariat, will eventually find itself in the proletariat and will be able to give to the proletariat the self-identified means to undertake a revolution. What is important to take from this critique is that exploitation itself, in any form, is self-perpetuating and the task of liberation cannot necessarily be left to the exploited. This could have broad implications for feminism (as is clearly pointed about by Catherine MacKinnon), present day worker exploitation, economic globalization, issues of race, and issues of sexuality.

In this paper, I first examined Hegel’s conception of the emergence of independent agency. Secondly, I looked at Marxist alienation and how this related to Hegel’s assertions and from this showed the difficulties of assuming a proletariat revolution due to the inability of worker to gain self-identity. Lastly, I examined the implications of this for understanding oppression in general
and briefly looked at what this means for overcoming such oppression in the future. It seems clear that reform and resistance are only possible through a consciousness of reformation and resistance.

Endnotes

2 Ibid.
4 Ibid., 58-63.
5 Ibid. 111.
6 58-63.
7 J. Gauthier, “Consciousness, the Other, and the Self,” 2.
8 Hegel, *Phenomenology of Spirit*, 112.
9 Ibid., 114.
11 Hegel, 115.
12 Ibid., 116-117.
13 Gauthier, “Consciousness, the Other, and the Self,” 5.
14 Hegel, 117.
15 Hegel, 118-119.
17 Hegel, 119-138.
18 Taylor, 154-158.
21 Ibid., 76-77.
22 Ibid., 76.
23 Fromm, *Marx’s Concept of Man*, 44.
25 Ibid., 73-74.
26 Ibid., 75-77.
27 Ibid., 77.


29 Marx, “Estranged Labour,” 76.


Bibliography


Gauthier, Jeffrey A. “Consciousness, the Other, and the Self.”


I. Introduction

In the pages to follow, we will pursue, into the pitfalls of the problem of the asymmetries of explanation, and back out again, the pragmatic, context-dependent, account of explanation forwarded by Bas van Fraassen, as defined in his text *The Scientific Image*.

In their 1987 paper, ‘*Van Fraassen on Explanation*,” Wesley Salmon and Philip Kitcher criticize van Fraassen’s account of explanation on the grounds that it can not be taken as a rival to long-standing theories of explanation (as, for instance, the Hempelian account) without, at minimum, addressing the (yet standing) problems of the rejections and asymmetries of explanation which undermined those accounts. They question his claim to have solved the problem of asymmetries, criticizing his fable of “The Tower and the Shadow” as an inadequate response (while praising his response to the problem of rejections). Their critique of that example in both the paper, and briefly, in Wesley Salmon’s text *Scientific Explanation and the Causal Structure of the World*, is sound. This criticism prompts their claim that van Fraassen has offered an interesting and very useful account of the pragmatic aspects of explanation, and in his system’s not resolving the traditional problem of asymmetry, that he has not provided, as he set out to do, a pragmatic theory of explanation to be considered alongside other such accounts.

While I am convinced of the success of their attack on the example van Fraassen offers with regard to the problem of asymmetries, I am equally convinced that his system fares far better with regard to that problem than his example portrays. Van Fraassen’s understanding of explanation affords a unique perspective on the problem of asymmetries and their cause. As such, it provides a unique solution to the problem.

This examination will begin with a brief description of the problem of
asymmetries, followed by van Fraassen’s account, his defending example and Salmon and Kitcher’s criticisms. It will conclude with a defense for the system, which I take to be stronger than that provided by van Fraassen’s own example, but which is nevertheless drawn from implicit (and explicit) aspects of his overall account.

We will be forced to conclude that van Fraassen’s account, in providing a working solution to the problem of asymmetries, is in fact an account to rival other accounts of explanation. And so, it is a pragmatic theory of explanation.

II. The Problem of Asymmetries

An account of explanation must include the machinery for distinguishing good explanations from non-explanations. The problem of the asymmetries of explanation arises because a particular kind of non-explanation looks to some account’s machinery exactly like a good explanation. A distinction between the two, the explanation and the non-explanation, should be found in the account, but is not. Here are a couple of traditional examples of this problem, the “flagpole” and “barometer”:

It is possible to explain the length of a flagpole’s shadow, by giving the height of the pole, the elevation of the sun at a given time and some general information or laws about the principles of light and optics. If someone wondered about the length of a flagpole’s shadow, a clever person could pull out those interesting tidbits of fact as an explanation of it. If, on the contrary, someone were to wonder about the height of the flagpole, the reverse could not be said; intuitively one could not hand them the length of the shadow, the principles of light and optics, and the elevation of the sun, and say “that explains the height of the flagpole”; the explanatory relation does not “run” in that direction. Perhaps that is because the flagpole causes the shadow of the flagpole, and the shadow does not cause the flagpole. Unfortunately some accounts of explanation, most notably the Hempelian D-N model of explanation, do not have the machinery to differentiate between the two ‘explanations,’ the one that is and the one that isn’t.

The second example, the “barometer,” is also a traditional example. It is a fact that a barometer’s level falls almost inevitably with the coming of a storm. This is because barometers are made to indicate air pressures identifiable with storms, where their levels drop. A clever person seeing the dropping of the barometer’s level might explain the event by saying, “a storm is coming.” It is not deemed appropriate for a meteorologist to explain the approach of a storm by noting, “oh yes, that happened because the barometer level is dropping.” Once again, an account of explanation whose machinery takes the latter to be
as explanatorily appropriate as the former is insufficient.

There are a number of other examples of this traditional problem. But, these two should serve sufficiently for our discussion.

III. Van Fraassen’s Account of Explanation

Salmon and Kitcher’s Reading

Bas van Fraassen’s account of explanation is at the heart of this examination. The reading of the account found in “van Fraassen on Explanation” differs from my own in its de-emphasis on the role of context. An understanding of the objections Salmon and Kitcher level depends upon a clear picture of their reading of the account. Let us first get their reading of it on the table. Following that we will extend and correct for our own reading, and then in section IV we will more closely examine the example that van Fraassen chooses to stand in for his account against the problem of asymmetries as described.

Salmon and Kitcher sum van Fraassen’s account of explanation as follows:

According to van Fraassen, an explanation is an answer to a question Q of the form: Why Pk? where Pk states the fact to be explained—i.e., the explanandum (phenomenon). Any such question can be identified as an ordered triple <Pk, X, R>, where Pk is called ‘the topic’ of the question, X = \{P1,...,Pk,...\} is its contrast class, and R is its relevance relation. Such a question is posed in a context that includes a body of background knowledge K.

Q also has a presupposition, namely:

(a) Pk is true;
(b) each Pj in X is false if j \neq k;
(c) there is at least one true proposition A that bears relation R to <Pk, X>.

(a) and (b) together constitute the central presupposition of Q. The why-question Q arises in the given context if K entails the central presupposition of Q and does not entail the falsity of (c). That is, it is altogether appropriate to raise Q even if we do not know whether there is a direct answer or not, provided the central presupposition is fulfilled.

If the question does not arise in the context, it should be rejected rather than answered directly. This can be done by offering a corrective answer, i.e., a denial of one or more parts of the presupposition. If the central presupposition is satisfied but (c) is in doubt, a corrective answer to the effect that (c) is false may be suitable.

If the question arises in the given context, it is normally appropriate to provide a direct answer. The canonical form of a direct answer [B] to Q is:

(*) Pk in contrast to the rest of X because A.
The following conditions must be met:
(i) A is true.
(ii) Pk is true.
(iii) No member of X other than Pk is true.
(iv) A bears R to <Pk, X>.

A is the core of the answer, for the answer can be abbreviated ‘Because A.’

. . . [Of] the person Sq who asks the question Q . . . [and] the respondent Sr, . . . in keeping with van Fraassen’s approach . . . understand Sq and Sr . . . [to be] operating in a common context with a common body of background knowledge K determined roughly by the state of science at the time. Thus, K may contain many propositions that neither the questioner nor the respondent knows. Moreover, Sq may have false beliefs that are in conflict with propositions in K. Sr may therefore offer corrective answers to flawed questions by pointing to items in K.

Whether A, the core of the answer to Q, is relevant depends solely on the relevance relation R. If A bears R to <Pk, X> then A is, by definition, the core of a relevant answer to Q (Kitcher and Salmon, p.317-318).

This summary captures the formal content of van Fraassen’s account with the exception of his section on the evaluation of answers (summarized and discussed elsewhere in the Kitcher/Salmon article) and some details regarding “the topic” Pk that are inessential to this discussion.

An Extended Reading

We need to consider the format of the why-questions more closely in order capture clearly the emphasis in van Fraassen’s account:

Why (is it the case that) Pk in contrast to the rest of X?

In this question, the contrast class X, is one of the contextually determined elements of the explanation. As such it enjoys a role of importance. Salmon identifies van Fraassen’s contrast-class with an “explanandum-partition,” and although he praises its usefulness in identifying the question behind a given sentence, he objects strongly to its role in the answers to why-questions. In the question, “Why (is it the case that) there is water on the sidewalk?” the contrast-class could be {there is water on the sidewalk, there is snow on the sidewalk, there is snow and ice on the sidewalk}, or, it could be {there is water on the sidewalk, there isn’t water on the sidewalk}, and so on. The contrast class chosen helps to define the question beneath the surface question, and will be determined by the context of the situation and the interest and perspectives of
Although background knowledge does play a contextual role, the other big player in this regard is the relation of relevance, R. When a given question is asked, the context of the question (questioner’s interests included) determines the relevance relation, what van Fraassen describes as “the respect-in-which a reason is requested.” Van Fraassen’s account places a great deal of emphasis on this relation. It is his contention that without context, and the subjective interests in particular aspects of a question that questioners have, (all of which is largely defined by, or located in, the relevance relation) explanations would not exist at all (more on this below, and see p. 130, The Scientific Image).

The relation of relevance does not admit of formal characterization at present. Although van Fraassen indicates that a project might be made of classifying the various types of relevance relations, his account suggests that there are quite a number of them. One example that he gives of the role of this relation is in the modern-dressed lantern example, from Aristotle’s Posterior Analytics II, 11 in which the relevance relation defines, for one interrogative, two different questions. A father asks his son “Why is the porch light on (and not off)?” To which the son replies, “because the switch is closed and electricity is flowing to the light.” The son is answering either of (at least) two distinct questions. The father may wonder whether or not guests are expected, in which case, van Fraassen says, the lad is being “impudent”; or, the father and son may be involved in fixing the home’s circuitry, in which case, the father may be wondering whether or not he has caused a short circuit, bypassing the light’s switch. This emphasizes the role of the relevance relation quite neatly: the relevance relation is the definition, in a sense, of both the question, and the relation to it of the desired response, which stems from the question’s context.

There are, as we will see, substantial problems with the relevance relation. One of which is my own concern, that van Fraassen, in attempting to provide a theory of explanation, has instead provided an interesting and useful theory around explanation, leaving explanation itself in a black box and calling it the “relevance relation.” This concern will not be addressed here.

Salmon and Kitcher are particularly frustrated by the lack of formal constraints on the relevance relation, describing it as providing an “anything goes” variable into explanation. As I will defend below, I disagree.

IV. ‘The Tower and the Shadow’ Example

Section 3 of “The Pragmatics of Explanation,” in van Fraassen’s The Scientific Image, is entitled “Asymmetries of Explanation: A Short Story.” ‘The Tower and the Shadow’ is the example this section contributes to his
account, an example of ‘an asymmetry of explanation’ being reversed by a change in context.

Van Fraassen’s account takes the asymmetries of explanation to be caused by the contextual relevance relation hidden in requests for explanation. He feels that the ‘reversal’ of an asymmetry by contextual factors, an example of an asymmetry becoming a good explanation, confirms this view and so, his account.

Salmon and Kitcher offer the following summation of the example, as it is taken from The Scientific Image:

In van Fraassen’s story, a character offers the following explanation of the height of the tower:

That tower marks the spot where [the Chevalier] killed the maid with whom he had been in love to the point of madness. And the height of the tower? He vowed that the shadow would cover the terrace where he first proclaimed his love, with every setting sun—that is why the tower had to be so high. (van Fraassen p. 133-4; Kitcher and Salmon p. 316)

The thrust of this example is that it appears (at least at first blush) to turn the flagpole example on its head, explaining the object with its shadow.

V. Assessing the Example As A Response to Asymmetry

Although van Fraassen chose the Tower and the Shadow example to spearhead his account’s attack on the Asymmetries problem, the example is insufficient. Demonstration of this fact is explored in the next four sections where I try to separate concerns Salmon and Kitcher raise.

In sections (I) and (II), I examine the aim of van Fraassen’s example, addressing the question: If the example proved successful, would it be sufficient to defend the account? By (II) I consider an objection to that aim, presented in Scientific Explanation and the Causal Structure of the World. Section (III) presents an alternative goal attributed to van Fraassen by Salmon and Kitcher and its proposed failure. While maintaining the failure of the example in that role, I argue that van Fraassen’s aim in the example differs. In (IV), I display Salmon and Kitcher’s dissatisfaction with the relevance relation in particular.

(I) “The context . . . determines relevance,” (van Fraassen p. 129) van Fraassen repeatedly tells us in The Scientific Image. He continues a page later, “the description of some account as an explanation of a given fact or event, is incomplete. It can only be an explanation with respect to a certain relevance relation and a certain contrast-class. These are contextual factors . . .” (van Fraassen, p. 130, bolded emphasis added). “If that is correct,” he says
in 3.1, and, I think, here is his intention with regard to the Tower and the Shadow example, “if the asymmetries of explanation result from a contextually determined relation of relevance, then it must be the case that these asymmetries can at least sometimes be reversed by a change in context” (van Fraassen p. 130).

The Tower and the Shadow example does, perhaps, show a reversal of an asymmetry on the surface, and it is caused by a change in context, although Salmon and Kitcher’s objection in section (III) will give us reason to wonder whether it is an acceptable reversal. I think however there is a larger question. Why does van Fraassen think that giving us an example of an asymmetry being reversed by a change in context, will show either, (i) that an explanation is not an explanation unless it is context-determined, or (ii) that his context-based account of explanation thereby solves the asymmetries problem? (The latter of these two questions is among our main interests in this paper.)

These passages indicate that he takes the asymmetries of explanation to be attributed to the, previously unnoticed, contextually determined relevance relation. If so (this is assessed below) would an example of the asymmetry being reversed by context solve the problem, by displaying this attribution and showing that his account has the machinery to distinguish explanations from non-explanations that run the wrong way? Perhaps his account does have such machinery, (and perhaps the asymmetries are so attributable,) but, an example of the asymmetries being reversed by a change in context, as given in the Tower and the Shadow example, does not display that feature of his account.

(II) Wesley Salmon, in Chapter 4 of his Scientific Explanation and the Causal Structure of the World, notes that the example points out “that, in a certain context, the length of the shadow can explain the height of the flagpole . . .” he concludes, “none of us would have doubted that claim.” (Salmon p. 95) Indeed, for the problem of asymmetry it is an uninteresting revelation. Although there might arise circumstances, well-displayed by van Fraassen’s method, in which the explanatory relation runs in the opposite direction, that is unfortunately no reason, in itself, to believe that his theory prevents those instances in which it doesn’t (run in the opposite direction) from slipping through i.e. being taken as explanatory when they are not. The example does not demonstrate a mechanism of distinction. As such, the example can not be taken as successfully displaying the account’s solution to the traditional problem of asymmetries. Van Fraassen considers this reversal under context a “crucial test” for his account, but I think that it is not. Even where the reversal is taken
as successful, his account is not thereby displayed as having the capability to deal with asymmetries.

(III) The problems with this particular example may extend further. In “Van Fraassen on Explanation” Salmon and Kitcher interpret the example as “an attempt to show that the claim of explanatory difference is shortsighted.” They maintain that:

...failing to appreciate that arguments are explanations (the basis of explanations) only relative to context, we assess the explanatory merits of the derivations by tacitly supposing contexts that occur in everyday life. With a little imagination, we can see that there are alternative contexts in which the argument we dismiss would count as explanatory” (Kitcher and Salmon, p. 316, emphasis added). 

Their claim is that the example fails at this attempt. We will assess this claim before we again address the question of van Fraassen’s intended aim.

The claim that this example fails to achieve the presumed goal arises because the argument behind the explanation given in the story, is not the argument van Fraassen appears to present as wrongly dismissed as unexplanatory. On this reading, the argument that is incorrectly taken to be unexplanatory is the one that counts the length of the tower’s shadow, the elevation of the sun and the principles of optics as the explanation for the height of the tower. The argument Salmon and Kitcher identify that is actually forwarded in the end of the story as an explanation for the tower’s height is one that involves the Chevalier’s desires or intentions, the principle of rationality and some background “principles about the stability of the height and position of large physical objects” (Kitcher and Salmon, p. 317). If this is correct, van Fraassen will not have given an example of an argument dismissed that is actually explanatory, but rather of an argument dismissed, and another argument that is actually explanatory. In which case, his example will have done little, if any, work for him at all.

This claim looks to be correct. Imagining the goal of van Fraassen’s example as Salmon and Kitcher interpret it to be, the example fails to accomplish it. The argument behind the explanation given in the story is not the one that we would dismiss, the one where shadow length, optics, and sun explain the height of the tower; so, we have not erred in dismissing that explanation as unexplanatory; in fact, a different explanation, one involving intent and desire, is in order.
Given this result, it is time to reconsider van Fraassen’s aim. Did van Fraassen intend to show that sometimes explanations, which would ordinarily be unsatisfactory, are really incorrectly interpreted good explanations enmeshed in odd unnoticed circumstances? Van Fraassen recognizes that the explanations in the barometer and flagpole examples are not good explanations. So his purpose could not be to show that they are actually good. He does want to show that sometimes explanations of the “bad-looking” form, or explanations that run backwards, are really good. And he wants to attribute that “sometimes” to alterations in the backdrop of context. Here his emphasis on context re-emerges. I take him to have two aims, with regard to this example, a strong aim and a weaker aim. The strong aim is the resolution of the problem of the asymmetries of explanation, and the weaker aim is the support of the importance of context and its permeation into unexpected areas of explanation. I am convinced that the presentation of his example fails to achieve the stronger aim and achieves at best the weaker.

(IV) Salmon and Kitcher provide a further reason to take the account’s example as inadequate; this one relates back to the objection to their interpreted meaning of the fable, in (III), but is nevertheless worth independent mention. Its solution, with the overall solution, will be found below. After an often repeated suggestion of constraining the relevance relation formally, the two assert that it should be “easy to contrive a maximally well-founded question <Pk, X, R> such that the proposition of the shadow length [without the dependence on the chevalier’s intentions] will be the core of a perfect answer” (Kitcher and Salmon p. 328). This goal is achieved by assigning for the R relation:

A relation that holds between A and <Pk, X> just in case there is a D-N argument that derives Pk from A plus additional premises in K(Q). ([With constraints on the R appropriate to] ensure that the only available D-N arguments are those which invert the usual order of explanatory derivation) (Kitcher and Salmon, p. 328).

They conclude, “so van Fraassen’s theory allows explanations which correspond to those D-N explanations which intuitively ‘run the wrong way’” (Kitcher and Salmon, p. 328).

Van Fraassen’s account must provide a response adequate to deny this claim (in addition to addressing the above concerns) in order to successfully defend against the problem of asymmetries, again. Although no such solution appears
explicitly in his writing in response to the asymmetries his account does provide one implicitly. We will examine this within the discussion to follow.

We have, I think, gathered ample reason to suspect that the Tower and the Shadow example is a less than successful attempt on behalf of the account to solve the problem of asymmetries. Fortunately, I believe, van Fraassen’s account provides a more insightful, and effective, response to the problem.

Assessing the Account: Keeping it All in Context

In Salmon and Kitcher’s essay the central role context plays in van Fraassen’s theory goes largely unremarked upon. I take this to be an error.

Although the formal mechanisms of the account will do the bulk of the work in distinguishing explanations from non-explanations, much of the thrust of van Fraassen’s ideal is located in the relation of relevance. The relevance relation is an as-yet-undefined relation formally speaking, that is nevertheless specifically defined. The relation of relevance, R, is not a placeholder for “anything you like here”; although the “respect-in-which a reason is requested” is not subjectable to a standard formalization (or even a standard constraint). To expect it to be is to remove the subjective context-dependence upon which van Fraassen insists. This context-dependence is that which defines his account, and which will provide his response to the asymmetries problem.

Here, again, is a clue to van Fraassen’s interpretation of the asymmetries problem: “...the asymmetries of explanation result from a contextually determined relation of relevance...” (van Fraassen p. 130). He felt this would be proven if he could find an example of context reversing an asymmetry, and that would display that his context-based system solved the problem. As such, I believe his only error was in saying too little in his section on asymmetries. His system has, implicitly, the mechanism to handle the asymmetries; it needs only be made explicit.

“The asymmetries of explanation result from a contextually determined relation of relevance,” because without that relation [and the other context-based factors (the contrast class for instance)], explanation qua explanation would not exist at all. Where there is no “respect-in-which a reason is requested,” van Fraassen asserts, there is no explanation. (See van Fraassen page 130, and much of chapter 5, for further discussion of this belief.)

If the asymmetries of explanation are the result of the relevance relation, then the awareness of that relation, as van Fraassen’s account has brought about, and its conscious use, should result in the needed mechanism of distinction. When is a given explanation a good explanation? And, when is it a non-explanation? For this, look to the relevance relation, contrast class, and
background knowledge.

Let us return to one of the traditional examples:

(Q1) Why (is it the case that) there is a storm (in contrast to there not being a storm)?

The topic, Pk: “there is a storm”; the contrast class, X, {there is a storm, there is not a storm}; the relevance relation? Does the questioner, in this case, implicitly ask for: a cause (a physical cause) or causes, at least one of which, sufficient to bring about the storm, occurred and brought it about? Possibly, even likely, the answer is yes. If so, certain specified “atmospheric conditions” would be a direct answer (B) to (Q1); we will call this answer simply (Al). There are other respects in which a question of that form (Q1) can be asked, but to that question, with that relevance relation, (Al) is the answer.

Would the response (A2): “because the barometer level fell” serve as the core of a direct answer to (Q1)? It may be that the topic is true, thus, it alone of the members of X, is true, and it favors the topic, and is a part of the background knowledge, and implies the topic and the negation of the other members of X; but, does it stand in the appropriate R relation to the topic and contrast class? Is it a cause, sufficient to bring about the occurrence of a storm? No, in failing to be in the appropriate relevance relation to the question (Q1) as it is being asked, (A2) fails to answer it; (A2) fails to provide an explanation for the topic (the explanandum) of the question.

This looks like a successful response to the second of our examples, above; what about the first? Will a similar treatment produce the necessary distinction again?

(Q2) Why (is it the case that) the shadow of the flagpole is 10’ long (rather than 5’, 6’, 7’ . . . long)?

(Q3) Why (is it the case that) the flagpole is 25’ high (rather than 15’, 20’ . . . high)?

In the case of (Q2), the topic is “the shadow of the flagpole is 10’ long;” and there is an appropriate contrast class. What is the relevance relation? The respect in which we want a reason, might be “why . . . when it is only the early part of the afternoon?” A math instructor might ask the question (Q2) of her students; a physics instructor might do likewise. The answers to these questions, and any others we might dream up for (Q2), will likely at least refer to the
flagpole in some way. And what of (Q3)?

For (Q3), topic and contrast class, and etc., are as expected; but, the relevance relation is a little harder to imagine; it is not often that a person asks “Why?” a flagpole is a certain height. The most likely R relation here is one that requests a reason in terms of design, or accident. For instance, “it had to be high enough to display the flags to the visitor’s gate across the square,” or “no particular reason, the poles are just cut to that size.” It is possible in an imagined context that it: blocks a view from a window, or, was ordered at 23’, or, was supposed to serve as a sundial as well, or, represents a new era in flagpole dimensions. But, regardless, the respect-in-which an answer is requested will not admit of: “because the shadow is 10 feet, and the sun is so high and etc.” (A3), unless the question specifically requires such an explanation (perhaps a math or science instructor did ask it of his/her students); and, in that instance, that explanation would be an explanation, not a non-explanation.

This point brings us forward a rather large pace in the discussion. As I find these two examples convincing (that this may in fact be a solution to the problem of asymmetries and not a startling coincidence), I am content to move on.

Is it correct and in the spirit of solving the problem of asymmetries, ever, to respond to (Q3) above with an answer/argument/explanation including: the length of the flagpole’s shadow, the elevation of the sun, and the principles of optics? Or is that sort of response a non-explanation automatically? The reason for this question should be apparent. The final examples of the paragraph above produce just that response, under van Fraassen’s system; if it is never an appropriate explanation, that should not occur. And yet, it makes sense that it does.

On a meteorology exam, an instructor provides a diagram for the students. The diagram pictures computer screens with indefinite satellite data on hot/cold fronts, a pictured television screen with a weather man, gray puff-clouds and fat water droplets, and a barometer with an arrow indicating the level as dropping. If the exam question was “Why is there a storm?” A student might be expected to answer, “because the barometer level dropped.” This is a healthily contrived context, but there is a message in it.

The reason that the drop of barometer level, and the shadow length, and red shift and other such responses are not explanations, is not located in their nature as objective objects to be categorized. It is located rather, in the respect-in-which an explanation is being sought. As such, sometimes, they are explanations. This does not “fly in the face” of asymmetry, rather, it resolves it, by identifying the heart of the problem itself.
Hempel’s model, and a number since then, proved unable to distinguish non-explanations from their parallel good explanations, because the distinguishing “features” were not there (objectively) to be found by the systems. If the asymmetries of explanation have always been the result of our (as “the explainers”) unconscious assumptions of relevance relations of certain kinds (most often, I would guess, causal), than self-conscious application of these assumptions should indeed resolve the problem of them, providing a distinguishing mechanism.

Section V raised a variety of concerns with van Fraassen’s account. These promissory notes are now being settled. In (I) we asked, “Why does van Fraassen think that giving us an example of an asymmetry being reversed by a change in context, will show either, (i) that an explanation is not an explanation unless it is context-determined, or (ii) that his context-based account of explanation thereby solves the asymmetries problem?” In (II), Wesley Salmon raised the similar concern, that although van Fraassen’s account might deal well with surprising instances of good explanations running the wrong way (as in the Tower and the Shadow example), this occurrence provides no indication that his account can pick-out or block those instances of bad explanations running backwards. In (III) we addressed and dismissed a claim from “van Fraassen on Explanation,” about van Fraassen’s aims in presenting the Tower and the Shadow example, considering two variations on the aim we had only begun to wonder about in (I). The strong aim we took to be solving the asymmetries problem, the weaker aim we took to be reinforcing the importance of context and its “permeation into unexpected areas of explanation.” In section (IV) we recognized Salmon and Kitcher’s frustration with the lack of formal constraints on the relation of relevance. And we noted a couple of formulas they suggested for infecting a van Fraassen-style explanation with the Hempelian explanation’s asymmetry weakness through that tender spot.

These questions [with the exception of (IV), to be addressed below] have been answered, not with the Tower and the Shadow example, but with the account of explanation itself. Van Fraassen’s emphasis on context, although mentioned and then largely ignored by Salmon and Kitcher, is justified; context is central to his claim to have resolved the problem of asymmetry. In answer to (III), van Fraassen’s example achieves the weaker aim (indicating the importance of context), while his account as a whole, achieves the stronger (addressing asymmetry as arising from previously unacknowledged contextual influence thus solving that problem). In answer to (I), his example at least points to (i), indicating once again that explanation is essentially . . . contextual;
his account, once again handles (ii) resolving asymmetries if he is correct about their nature. In answer to (II), his example might indicate that the account can handle good explanations running the wrong direction; his account indicates the machinery it uses to distinguish the good explanations from the bad. Although there may be further concerns with van Fraassen’s account of explanation and his claimed resolution of the problem of asymmetries, Salmon and Kitcher’s concerns presented here with the exception of the loosely-related (IV), treated below, are revealed to be unproblematic.

From here, two thoughts are in order, requiring us to sidestep for a moment. Firstly, in the above “meteorology exam” example, I am not completely certain that it is appropriate to label the expected answer “an explanation” of the storm (and likewise, I am not sure that van Fraassen would deem it appropriate to do so; Wesley Salmon’s remarks, (Salmon p. 127) indicate I believe, that he [Salmon] would not, and that he takes van Fraassen to share that view of it). In that specific and contrived context, maybe, the term could be used. Similarly, in a mathematics text, I might expect to see, “the elevation of the sun and the length of this shadow explain/describe this pole’s height,” but I am not convinced that that is an appropriate use of the term. If so, is van Fraassen’s project propelled back into troubled waters with regard to asymmetry? I would say, with qualification, no; even if this is the case, one may not answer (Q3) with (A3), except under the dubious context of mathematical or scientific calculation, and so, this “backward” response will not trouble us in “standard” occurrences of the question. However, as I am not certain of the mechanism that distinguishes the calculation-related Why-questions from the “regular” questions, I am not confident in this response.

It may help to respond to this issue to note that in much of this discussion I have not addressed firmly van Fraassen’s intentions that the answer-explanation be a causal one. Those example answers given above, are, in different senses causal. In the case of the meteorology exam, for instance, the explanation being sought might appropriately be characterized: “Of the above possible reasons pictured here, which one is appropriate to read as giving a reason sufficient to claim that there is an expected storm.” Or in other words, “. . . which one best explains the proclamation of a storm’s approach by a properly trained meteorologist.” This is a causal explanation, in a round-a-bout sort of way; so, it may warrant the term “explanation.”

The second thought, I’d like to consider briefly, is somewhat more optimistic. Wesley Salmon and Philip Kitcher, and others, are distressed by the lack of formal constraints on the relation of relevance. It has been suggested
to me\(^\text{(4)}\) that the relevance relation is formally constrained. If it is the case that the relevance relation is independently generated at each instance of an explanation request by the context in which that request arises, then the formal constraints on that relevance relation are so generated as well, by the context at that time. I find this possibility very intriguing; and, if explored, it seems to me that this might prove to be a formal defense of van Fraassen’s account.

These discussions bring us back to section (IV) of Section V of this paper; in which Salmon and Kitcher claim to have provided, through van Fraassen’s system, an explanation which they contrived which “runs the wrong way,” by asserting a relevance relation which holds “just in case there is a D-N argument that derives \(P_k\) from \(A\) plus additional premises in \(K(Q)\)” (Kitcher and Salmon p. 328). We are now in a position to see what is wrong with this assertion, although the response to it is a little hard to follow. Recalling that the relevance relation is not imposed at random, but rather defined by the question, it is apparent that imposing a contrived relevance relation (a perfectly justifiable test of an account) will automatically impose a specific question. Let us try to imagine what that question would look like; take:

\[(Q3) \text{“Why (is it the case that) the flagpole is 25’ feet high (rather than 15’, 20’... high)?”}\]

Now try to imagine the question itself, the abstract entity behind the sentence, as Salmon describes it (Salmon p. 102). In order to capture the assigned relevance relation above, what would that more fully expressed question have to be? Something like, I imagine, “Why (would we say) (that is, it the case that) the flagpole is 25’ high (rather than 15’, 20’... high”) if we were looking for a Hempelian-type explanation.” This question might come up in a philosophy text or class, in fact, it likely has. Moreover, to this question, the answer which “runs the wrong way,” “does not,” in fact run the wrong way, it is a perfectly adequate explanation, and is serviceable in critiquing the Hempelian account. This is not the only question that might be expressed which requires that particular relevance relation, although I can not immediately imagine any others. The point at issue is this: whenever a question is asked in which the response is defined by that contrived relevance relation, the question itself is defined by that relevance relation too; in other words, there will never, I think, be an invented question that really means, “Why (straightforwardly, causally, with no reference to D-N explanations) is the flagpole yea big?”, which will require a D-N model response attributing length-of-shadow, the
This being the case, I feel that van Fraassen’s understanding of explanation has addressed this “section (IV)” problem; when the answer is modified to be D-N restricted, so too is the question, thus any such answer will be an explanation (a good, and not a non-one) to its, appropriately interpreted, question.

In their collaboration, Salmon and Kitcher provide another example of the type of move applied in section (IV), let us apply this solution discussed to it for confirmation:

Consider the simple relation of derivation. This relation holds between A and <Pk, X> just in case there is a (first-order) derivation of Pk from A plus additional premises in K(Q) (Kitcher and Salmon, p. 327).

After defining this relation they set up a “famous Hempelian example” as follows:

Let Pk be the proposition that Horace is bald, and R be the relation of Greenbury-School-board-derivation that holds between A and Pk just in case A is a conjunction of propositions one of whose conjuncts is the proposition that Horace belongs to the Greenbury school board, Pk is derivable from A, and there is no conjunct in A that could be deleted while still enabling Pk to be derivable from the result (Kitcher and Salmon p. 327).

They define the appropriate contrast class, and give an A, which is a conjunct of Horace being a Greenbury School Board member, and all GSB members being bald. Their claim is that A is a perfect answer to a question addressing Horace’s baldness. They go on to claim that:

...most (if not all) of the examples of non-explanatory arguments that Hempel hoped to exclude-both those he succeeded in debarring and those which have caused persistent problem for the theory of D-N explanation-give rise to a corresponding “relevance” relation that van Fraassen ought to exclude. (Kitcher and Salmon, p. 327).

If they are correct in this assertion, van Fraassen’s account is in trouble, but, I do not believe that they are. To apply our proposed solution from above, we must ask: what is the question (behind the question) that is really being asked, as defined (in part) by that unique and contrived relevance relation?
Once again, we have a D-N derivation requested in the meaning of the question, so our question relates to Hempel’s model. It looks as though it (possibly) asks something about “What would be the Hempelian explanation for the baldness of Horace, which would begin with a conjunctive premise and . . . ?” In other words, the response that gives an explanation as the conjunct A above, would be providing, accurately and advisedly an explanation to that question. Once again, van Fraassen’s model provides us with an explanation where Salmon and Kitcher do not expect there to be one. If the two are persistent in assigning values to the relevance relation, they are assigning them in both the answer, and the question. I believe that, in disregarding the contextual-relevance, the role and function of the relevance relation, and the contextual restraints that define it, they continually skip over this important point.

The asymmetry problems, and this lesser problem, just above, both come from ignoring the contextual framework in which, van Fraassen says, explanations are inextricably a part. The solution, in both instances, is awareness of that context in a thoroughgoing way.

Conclusion

In response to Wesley Salmon and Philip Kitcher’s objections as discussed, I am content to say that these concerns are (preliminarily) addressed. The example chosen by van Fraassen, although inadequate as a defense, is not representative of the full range of the account, which has within it the means of addressing the problem.

Given the above discussion, we should now have a handle on van Fraassen’s account, sufficient to address flagpoles and barometers and red-shifts and a variety of other common and contrived examples of explanation’s, somewhat puzzling, nature. I believe that van Fraassen’s admonition to consciously apply and be aware of the relevance relation (and other contextual factors of requests for explanation), from which stem the problem of the asymmetries of explanation, provides the distinguishing mechanism required of his account to differentiate these good explanations from non-explanations, resolving that problem finally. As such, his account of explanation may be seen as a serious contribution to the philosophy of explanation, and earns it the title: “pragmatic theory of explanation.”

End Notes

A Van Fraassen shares that criteria; he takes a resolution of the problems of the asymmetries and rejections to be essential to a competitive theory of explanation.

B There is an important argument-answer distinction to be drawn here. Salmon
and Kitcher acknowledge that van Fraassen rewrites the language of argument in explanations as answers. Although they approach these issues as though the two may be treated equivalently, I am not sure that van Fraassen would agree. As this should not interfere with this discussion, I will not dispute it here.

C Salmon and Kitcher emphasize in their essay that this is one interpretation of van Fraassen’s example and so do not claim that this is necessarily the one he intended.

D By Alfred Turnbull, who helped me to work through some of these ideas.

Bibliography


A Socratic Unity of the Virtues

Wendy Lynn Clark
Lewis and Clark College

I. Conflicting Unities

In some places in the early dialogues, it appears that Socrates conceives of the virtues as different names for the same thing; in other places, he seems to believe that the virtues are definitionally distinct parts of a single thing; and in still other places, Socrates seems to believe in a hierarchy of virtues, such that some virtues can be classified beneath more encompassing virtues. How are we to understand these apparently conflicting passages? The two points that seem to be particularly at odds are Socrates’ suggestions that the virtues are united through wisdom (Prt. 350c, 361b-c) and his belief that the virtues need separate definitions (Euthphr. 12c-d, Lach. 190a-199e).

This paper will examine two of the leading theories regarding the unity issue. I will examine their strengths and weaknesses, and in the end, I will suggest a theory that has greater interpretive power. I believe a unified theory of the virtues is possible if we look at the Greek text more closely.

II. The Identity Thesis

The remarks that Socrates makes at Prt. 349b-c have led some scholars to believe that he means the strongest possible unity of the virtues, the identity thesis. This thesis, as it has been formulated and defended by Terry Penner, states that an instance of one virtue is at the same time an instance of all of the virtues. In the Prt. passage, Socrates says:

I believe the first question was this: Wisdom, temperance, courage, justice, and piety—are these five names for the same thing, or is there underlying each of these names a unique thing, a thing with its own power or function, each one unlike any of the others? You said that they are not names for the same thing, that each of these names refers to a unique thing, and that all these are parts of virtue, not like the parts of gold, which are similar to each other and to the whole of which they are parts, but like the parts of a face, dissimilar to the whole of which they are parts and to each other, and each one having its
own unique power or function (349b-c).

The parts of a face are different not only in reference but also in meaning: The eyes have a different appearance and function than the nose. Although both may be necessary to make up a face, one would not talk about the eyes in terms of smelling nor the nose in terms of seeing. They are functionally, definitionally distinct. The parts of a gold bar, however, are the same: They have the same properties of malleability, yellowness, and ductility in addition to having the same definition. Since Protagoras takes the view exemplified by the parts-of-face analogy, and Socrates spends the rest of the dialogue in argument with him, Penner argues that Socrates affirms the identity or parts-of-gold analogy. According to Penner’s identity thesis, Socrates believes that the virtues are one thing.

There is plenty of evidence for this reading in the Prt. For example, in one part of the argument, Socrates and Protagoras agree that for every one thing, there can only be one opposite. Socrates goes on to secure Protagoras’s agreement that wisdom and folly are opposites. By the end of the argument, they have concluded that temperance and folly are also opposites. At this point, Socrates asks:

Then which of these propositions should we abandon, Protagoras? The proposition that for one thing there is only one opposite, or the one stating that wisdom is different from temperance and that each is a part of virtue, and in addition to being distinct they are dissimilar, both in themselves and in their powers or functions, like the parts of a face? (333a).

Since they refuse to abandon the first proposition, Socrates asks, “Wouldn’t that make wisdom and temperance one thing?” Protagoras grudgingly agrees.

Perhaps the most obvious objection to this reading is that if Socrates seriously believes that all of the virtues name a single thing, his view disintegrates into incomprehensibility. Piety is just not the same things as courage; to say that something is pious is not to say that something is courageous. Imagine the statement “temperance involves moderation in eating” is translated into “courage involves moderation in eating.” The sentence no longer has the same meaning. It is incomprehensible to imagine that temperance is identical in meaning to courage, or vice versa.

A second objection to the identity thesis is that Socrates appears to have a different opinion about unity across dialogues. In the Lach., for example,
Socrates refutes Nicias’s definition of courage because his definition is too similar to the definition of the whole of virtue. Nicias defines courage as the knowledge of what is to be feared and hoped for, but wisdom is the knowledge of good and evil. Socrates ends the dialogue with Nicias by telling him, “we have not discovered, Nicias, what courage is” (199e). It has been argued that the text is neutral here, and that Socrates is waiting for Nicias to come to the realization that in reality, wisdom and courage are the same thing. However, a close examination will reveal that the text is not neutral. Socrates does not say that they should give up one definition or the other. He flatly states that they have not discovered what is courage.

There is a further problem for the identity thesis. According to it, an instance of one of the virtues is actually an instance of all of the virtues. Thus, an instance of justice is also an instance of piety, wisdom, courage and temperance. However, this does not seem to make any kind of sense. There would be no need for different definitions, and there would really be no need for different words. They could all be melted down into one word: virtue. Finally, Socrates declares that the god is all-wise. According to the identity thesis, this means that the god is also just, pious, courageous, and temperate. However, when does the god need to be temperate or courageous? In which circumstances? The identity thesis leads to some unsettling conclusions.

III. The Role of Wisdom

Another problem for those who would prefer to attribute a coherent philosophy to Socrates is the role wisdom plays in his dialogues. The role appears to change between the Prt. and the Lach. In the Prt., wisdom is a part of virtue, but in the Lach., wisdom is the whole of virtue. The leads Daniel Devereux to suggest that perhaps Plato represents Socrates’ own incoherence in the first dialogue and then attempts to resolve it in the subsequent Lach. and Men. dialogues.

Devereux believes that Penner, with some exceptions, was generally correct in his interpretation of the Prt.: Socrates indeed affirms an identity unity of virtue. The different virtues are all names for the same knowledge of good and evil, like the parts of a gold bar. In the Prt., wisdom is the same as temperance and as courage, and justice is nearly the same as piety; there is no suggestion that it is the whole of virtue (Prt. 349b). In the Lach., however, Socrates suggests that wisdom is the whole of virtue and that the different virtues require different definitions. He does not list wisdom when he enumerates the virtues (Lach. 198a). Additionally, his closing refutation of Nicias demands a definition of courage that is distinct from a definition of wisdom, the whole of virtue (199e).
Thus, in the *Prt.*, wisdom is merely a part of virtue, while in the *Lach.*, it is the whole of virtue.

The benefit of this incoherence thesis is that it sidesteps the problem of instantiation. Because Socrates’ earlier view was confused, we need not be concerned with whether or not the god need be courageous. Instead, we can take the *Lach.* to be a more comprehensible dialogue. According to it, it is possible for an action to be just that is not courageous, and a courageous man need not be a pious man unless there is something to be pious about. A real objection to Devereux’s thesis, however, is its very premise. His reading renders Socrates’ theory of the virtues to be incoherent. In the remainder of this essay, I hope to show that Socrates did have a coherent conception of virtue and that any incoherence is due to a translation issue. To understand his theory, it is important to understand the ancient Greek words *dunamis* and *ergon*.

**IV. Dunamis versus Ergon**

A *dunamis* is a power, capacity, or ability to do a certain task. An ability for carpentry, for example, is an appropriate *dunamis* for making a set of shelves. In the *Prt.* when Socrates asks Protagoras about the parts-of-gold analogy, he is asking if Protagoras believes the virtues are united by the same *dunamis*. In the examples, the *dunamis* of the eyes is the ability to be able to see, whereas the *dunamis* of the nose is the ability to be able to smell. Socrates asks if the ability to act in a virtuous manner is brought about by the same ability: Are these abilities alike as parts of a gold bar, or are they like parts of a face, “each one having its own unique power or function [dunamin]” (349b, my emphasis).10

It is important to recognize that the ability to be courageous is not the same as actually being courageous. Socrates asks how the virtues are related, not what makes each virtue different and unique. He is asking for the *dunamis* of virtue. Protagoras says that the virtues do not have a single *dunamis*, but have different powers and functions. Socrates spends the rest of the dialogue showing him that he is mistaken. Courage is wisdom (350c), temperance is wisdom (333b), and justice and piety are nearly the same thing (331a-e). This is a unity of identity similar to the sort for which Penner argues. The difference is that Socrates does not claim that all instances of justice are instances of piety (in fact, the *Euthphr.* expressly denies this). Instead, Socrates claims that the virtues are united by their *dunamis*, that which gives them each and all the ability to be performed as virtues.

This single virtuous *dunamis* is wisdom. According to Socrates, one cannot act virtuously without wisdom. At *Prt.* 350c, Socrates says, “. . . these men who are so confident turn out not to be courageous but mad? And, on the other
side, the wisest are the most confident and the most confident are the most courageous?” The men who appear confident without wisdom are not courageous, but crazy. Socrates considers only wise men to be acting courageously.  

Devereux is right to note that the role of wisdom changes between the Prt. and the Lach. This is because the Lach. asks a different question about virtue. The question of the Lach. is not how the virtues are related, but what makes them distinct from one another. When Socrates asks Nicias what courage is, he does not ask for the dunamis. However, this is what Nicias gives him: the knowledge of future goods and evils (wisdom). Socrates agrees that this is necessary for acting courageously (as opposed to acting stupidly or recklessly) but he also says that this is not what they are looking for. He does not want to definition of virtue entire; he wants the definition of courage alone.

Though this is not explicitly stated in the dialogue, I believe that what Socrates is asking from Nicias and Laches is not a dunamis but an ergon. An ergon is the end product of a dunamis. In the case of the virtues, it would be a function. Julia Annas defines it well:

An ergon is what a thing does qua a thing of that kind. As applied to people referred to by what they do, it means ‘job’ or ‘work’. As applied to a natural kind like men, it means little more than ‘characteristic behavior’. But it is important to remember that ergon covers sculpting, the sculptor’s job, and the characteristic behavior patterns of species . . . a thing’s ergon is that which only it can do, or which it can do best (54).

In the context of the virtues, ergon would be used in the sense of work. It is that which makes the courageous man act courageously. His choice to examine the ergon of courage makes perfect sense in the Lach. because he is discussing the teaching of fighting in armor with two Athenian generals. When Socrates begins the discussion with Laches, he makes it clear that he is not interested in the whole of virtue, but only with that part of virtue called courage. He asks Laches to “state what courage is” (190d). Laches proves wholly unequal to the task, but he does come to the conclusion that courage in battle or in fighting passions, “whether by standing their ground or running away” (191e) is a case of wise endurance. Endurance in the face of passions or overwhelming fears, when it is produced with the knowledge of good and evil, may indeed by a partial definition of courage. Later, Nicias gets into trouble because he defines the product of courage with that which produces it: wisdom. Socrates objects
on the grounds that wisdom does not have the same *ergon* as courage.

How is wisdom, the *dunamis*, related to courage, one of its *erga*? A *dunamis* is an ability or a capacity. An *ergon* is a job or work that is done to produce something. Hence, being wise enables one to be courageous in a situation when courage is necessary. Wisdom brings about *erga* besides courage. It also brings about justice, piety, and temperance. These are different *erga* and they produce different results. Thus, a man can be courageous while not necessarily being pious at the same time. This reading avoids the problem of instantiation because the god need not be courageous while being all-wise. This also explains why *erga* are so difficult to define. Courage, for example, does not consist in a list of actions that are always courageous in every situation. Laches tries to define courage by giving such a list, and for every example, Socrates gives a counter-example where the opposite action produced the same results.\(^{13}\) This is because the application of wisdom, and then the enactment of an *ergon*, are different in different situations.

Thomas Brickhouse and Nicholas D. Smith give an excellent example of how the virtues are united, and their example maps well onto the distinction I draw between *dunamis* and *ergon*. Their example is triangulation. Both coastal navigation and surveying use triangulation to fulfill their tasks, the former to navigate a boat and the latter to map out terrain. However, an act of coastal navigation is not an act of surveying. They are different in definition as well as in reference; yet, both use the same type of knowledge. In a like way, both courage and temperance use wisdom in order to make courageous and temperate choices. However, the courageous act to rescue a child from a burning building is not the same as the temperate act of exercising daily to keep up a healthy physique. Though both require wisdom, they are different acts by their intensions and by their extensions.

To extend the analogy further, what is produced by coastal navigation will be different every time coastal navigation takes place. If one asks how many degrees to turn the rudder in order to make for the coast and the answer, after using coastal navigation, is two degrees, that does not mean that the answer will always be two degrees. The answer will depend upon the position of the boat to the coast line. In a like way, courage will not always consist in standing one’s ground. Sometimes, it will involve running away and tricking the enemy, showing weakness while springing a trap. Because the virtues do not consist in lists of actions but in applications, it is understandable why Socrates would say that virtue cannot be taught. If humans cannot teach wisdom, then students will not have the *dunamis*, or ability, to be virtuous.\(^{14}\) Nicias, Laches, and
Protagoras should renounce their claims to teach virtue and focus instead upon something more useful, such as philosophy.

**IV. Conclusion**

It makes the most sense to speak of the virtues as united by a single *dunamis*, not by a unity of identity. Socrates’ theory can be rendered coherent and comprehensible by reading the *Lach.*, *Men.*, and *Euthphr.* dialogues as asking for the definitions of the virtues as *erga*, not asking for the *dunamis* of virtue. The only problem is that wisdom, as Socrates laments in the *Ap.*, is only available to the god. We may occasionally get lucky and have true belief, but true belief, like the statues of Daedalus, is flighty and easily swayed by appearance. Perhaps it is only the pursuit of philosophy that can strengthen the measuring art, and set us on the path to virtuous action. As Socrates says, “The unexamined life is not worth living for a human being” (*Ap.* 38a).

**End Notes**

3. *Euthphr.* 11e-12e.
4. Penner formulates his general thesis in “The Unity of Virtue.”
5. Noted also by Vlastos.
6. Thanks to Nicholas D. Smith for help clarifying this analogy.
8. Penner argues for this reading in his article “What Laches and Nicias Miss—And Whether Socrates Thinks Courage Merely a Part of Virtue.”
9. *Ap.* 23a-b. “What is probable, gentlemen, is that in fact the god is wise and that his oracular response meant that human wisdom is worth little or nothing . . .”
10. From the beginning of the *Prt.*, Socrates makes it clear that he is asking for the power or function of the virtues, the *dunamis* involved. He asks for this specifically at *Prt.* 330a and at 333a, for example.
11. Socrates makes a distinction between foolish and wise endurance in *Lach.* 192c-194b.
12. In the *Euthphr.*, however, it is made explicit that Socrates is asking for an *ergon*. Regarding piety, Euthyphro tells Socrates that it benefits the gods that humans are pious the way that slaves are beneficial to masters—by a type of service. Socrates responds by saying the service of a ship-builder is to build ships, of a doctor is health, and so, “Then tell me, by Zeus, what is that excellent aim [*ergon*] that the gods achieve, using us as their servants?” (13e12). It is likely that Socrates asks for an *ergon* in the *Lach.* because the generals Nicias and Laches are not offering to teach virtue, but only an *ergon* of virtue: courage in the art of fighting in armor.
13. At *Lach.* 190e-191e, Laches defines courage as “willing to remain at his post”
in a war situation, and Socrates brings up an example where the Spartans won their battle while in retreat. The retreat was a tactical move, not a cowardly one, and therefore Laches could not call the Spartans cowardly.

14 That Socrates believes wisdom is not available to humans unless through divine dispensation is apparent from *Men*. 100b.


16 Many thanks to Nicholas D. Smith for his encouragement and critical insights, and without whom this paper would not have been possible. Thanks also to the Socrates seminar, David Puckett of the University of Memphis 2001 undergraduate conference, and Michael Brent and the participants of Pacific University's 2001 Northwest Regional Undergraduate Philosophy Conference for their comments and criticisms. Though none of them can be presumed to agree with this thesis, their arguments and insights strengthened my own.

Works Cited


**STOA: Vol. 1, Issue 1**

**Articles**

*Andrew Moore,* “Is Time Travel Logically Possible?”
Mentor: Prof. Timothy Williamson, University of Edinburgh

*Valerie McCool,* “The E-God Argument vs. the Ontological Argument”
Mentor: Dr. Aleksandar Jokic, Santa Barbara City College

*Matthew T. Kline,* “Rawls's Response to Hart”
Mentor: Prof. Burleigh Wilkins, University of California, Santa Barbara

*Adam Davidow,* “Gadamer: The Possibility of Interpretation”
Mentor: Prof. David Hoy, University of California, Santa Cruz

*Dominic Gregory,* “Senseless Communication”
Mentor: Dr. Tim Crane, University College London

*Miranda K. Smith,* “Divergent Views of Choice, Voluntariness, and Coercion in the *Nicomachean Ethics*”
Mentor: Prof. Richard Miller, University of Indiana, Bloomington

*James Walmsley,* “A Puzzling Footnote to the Godel-Hilbert Debate”
Mentor: Prof. William Newton-Smith, Oxford University

*Miroslava Andjelkovic,* “The Paradox of Inexact Ignorance”
Mentor: Prof. Milos Arsenijevic, University of Belgrade

~ Quidditas ~

“Problems of Time and Change”
*Prof. L. Nathan Oaklander*
University of Michigan-Flint
STOA: Vol. 1, Issue 2

Articles

Alice Bell, “Ought I To Worry About Being Dead?”
Mentor: Prof. Mark Walker, University of Birmingham, England

Rachel Walton, “Faith or ‘Superstitious Delusion’?: A Study of David Hume’s Critique of Miracles”
Mentor: Prof. Timothy Shanahan, Loyola Marymount University

Shannon Fullmer, “Refuting the Evil Genius Argument”
Mentor: Prof. Richard Montgomery, West Virginia University

Michael Mays, “Is it Good to be Noble? A Critical Examination of Friedrich Nietzsche’s The Genealogy of Morals”
Mentor: Prof. Joe Mintoff, University of Newcastle, Australia

Christopher Timpson, “Belief”
Mentor: Prof. Mark Walker, University of Birmingham, England

Heather Fotheringham, “How Long is the Present?”
Mentor: Prof. Robin Le Poidevin, University of Leeds

Timothy C. DiMarco, “Schelling’s Philosophy of Art”
Mentor: Prof. Predrag Cicovacki, Holy Cross College

Nicolas Paul Hare, “Knowing That One Knows”
Mentor: Prof. Timothy Williamson, University of Edinburgh

~ Quidditas ~

“Why Is There Something Rather Than Nothing?”
Prof. Robin Le Poidevin
University of Leeds
STOA: Vol. 2, Issue 1

Articles

Veronika Elfimova, “An Ethic of the Sexual and the Just”
Mentor: Prof. Laurence Nelson, Santa Clara University

Michael Ferreira, “Psychological Egoism”
Mentor: Prof. Barbara MacKinnon, University of San Francisco

Phil Jenkins, “Objectivity and Evolution in Ethics”
Mentor: Prof. John Longo, Hunter College

Jonez Kratz, “Realizing Self-Respect in a Rawlsian Society”
Mentor: Prof. Debra Satz, Stanford University

Cynthia Inda, “Rousseau and Freedom”
Mentor: Prof. J. Russell Muirhead, Harvard University

Mark Schroeder, “Relector Ergo Obligatus Sum: Christine Korsgaard and the Moral Normative Question”
Mentor: Prof. Jennifer Manion, Carlton College

Christopher Warren, “Nelson Goodman and the Value of Artistic Expression”
Mentor: Prof. Mark Walker, University of Birmingham, England

Andrew Fisher, “Rejecting Externalism and Internalism as a Response to Michael Smith’s ‘Moral Problem’”
Mentor: Prof. Mark Walker, University of Birmingham, England

Leif Holden, “Searle on Intentionality”
Mentor: Prof. Joseph White, Santa Barbara City College

Gordon Hesketh, “Defending Externalism”
Mentor: Prof. Timothy Williamson, University of Edinburgh

~ Quidditas ~

“Heidegger and Sartre: Radically Opposed?”
Prof. David Cooper
Durham University, England
The Center for Philosophical Education
Previous Conferences

Time, Tense and Reference
April 11-13, 1997
Participants: Hugh Mellor (Cambridge University), David Kaplan (University of California, Los Angeles), Nathan Salmon (University of California, Santa Barbara), Nathan Oaklander (University of Michigan), Mark Richard (Tufts University), Milos Arsenijevic (Belgrade University), Quentin Smith (Western Michigan University), Steven Humphrey (University of Louisville/UCSB)

War Crimes: Moral and Legal Issues
November 14-16, 1997
Participants: Margaret Gilbert (University of Connecticut), Richard W. Miller (Cornell University), James W. Nickel (University of Colorado), Michael Slote (University of Maryland), Jovan Babic (Belgrade University), Burleigh Wilkins (University of California, Santa Barbara), Angelo Corlette (San Diego State University), Kathleen Bruhn (University of California, Santa Barbara), Svetozar Stojanovic (University of Belgrade), Peter Digeser (University of California, Santa Barbara)

Violence, Responsibility and Ethnic Reconciliation
October 8-10, 1998
Participants: Richard Falk (Princeton University), Alfred Rubin (Tufts University), Slobodan Samardzic (Belgrade University), Virginia Held (CUNY), Aleksandar Pavkovic (Macquarie University, Australia), Burleigh Wilkins (University of California, Santa Barbara), Jovic Babic (University of Belgrade, Yugoslavia), Aleksandar Jokic (Portland State University)

The Nature of Consciousness
November 6-8, 1999
Participants: David Chalmers (University of Arizona), Michael Tye (Temple University), Tim Crane (University College London), Stephen Stich (Rutgers University), Brian Loar (Rutgers University), Quentin Smith (Western Michigan University), Robert Van Gulick (Syracuse University)
Duty and Virtue in Epistemology
November 12-14, 1999
Participants: Simon Blackburn (University of North Carolina), Keith Lehrer (University of Arizona), Robert Audi (University of Nebraska), Alvin Goldman (University of Arizona), Ernest Sosa (Brown University), Richard Foley (New York University), Christopher Hookway (University of Sheffield), Hilary Kornblith (University of Vermont), John Greco (Notre Dame University), Linda Zagzebski (University of Oklahoma)

Cosmology - In the Beginning
April 7 & 8, 2000
Participants: Richard Swineburne (Oxford University), Adolf Grunbaum (University of Pittsburgh), James Hartle (University of California, Santa Barbara), Graham Oppy (Monash University, Australia), Quentin Smith (Western Michigan University), William Lane Craig (Talbot School of Theology), Steven Humphrey (University of Louisville/UCSB)

~ Most Recently ~

Emotion & Cognition
April 27 & 28, 2001
“The Role of Emotion in Epistemic Evaluation”
Christopher Hookway, University of Sheffield
“Emotion, Memory, and Political Identification”
Sue Campbell, Dalhousie University
“Emotion and Moral Judgment”
Linda Zagzebski, University of Oklahoma
“The Role of Digust in Emotion”
David Kim, University of California, San Francisco
“Paradoxical Emotion”
Ronald DeSousa, University of Toronto
“Intellectual Desires”
Michael Stocker, Syracuse University
The Center for Philosophical Education - C.P.E.

was established in 1997 by the Department of Philosophy at Santa Barbara City College to foster a broader understanding of a deeper appreciation for the academic discipline of Philosophy within both the undergraduate and general populations. To these ends, The Center publishes STOA as a forum for celebrating and nurturing the philosophical growth of undergraduate students in addition to providing the general reader with access to a range of quality academic writings from novice philosophers. The journal has a particular interest in publishing undergraduate papers from a wide range of countries around the world, thereby fostering cooperation between scholars and students from a variety of backgrounds.

Through the publication of undergraduate philosophy papers submitted by professors, STOA recognizes and supports the vital pedagogical position of mentor. For many undergraduates, the finding of a mentor is often the pivotal intellectual experience in their undergraduate education. Both the student and the supervising faculty member will be credited in STOA. Finally, each issue of STOA will either contain a section titled, QUIDDITAS, in which a solicited paper by a professional, academic philosopher will provide an overview of a contemporary philosophical problem; or a section titled, SYMPOSIUM, in which a faculty member submits a group of student papers on a specific topic followed by his or her critique of those papers.

Submissions should be sent to the Center for Philosophical Education in care of the Executive Director. While there is no fixed length for submissions, only undergraduate submissions will be accepted and each submission must be made through a college or university faculty member who is formally recommending the student’s paper. Papers should be submitted double-spaced standard format and on disc. Accompanying each disc should be information regarding software used, including release (e.g. Microsoft Word 5.0) and type of computer (either PC or Apple Macintosh).

All inquiries regarding STOA can be made through the following:

Executive Director
The Center for Philosophical Education
Department of Philosophy, SBCC
721 Cliff Drive, Santa Barbara, California 93109-2394
tele: 805.965.0581 #2476
fax: 805.963.7222
e-mail: white@sbcc.net
**STOA: Vol. 3, Issue 1: Spring 2002**

The **Center for Philosophical Education**

**Articles**

*Nick Tosh*, “Science and Consciousness”  
Mentor: Prof. Marina Frasca-Spada, Cambridge University, England

*Abe Sprinsock*, “Introspection and Knowledge of Mental Content”  
Mentor: Prof. Alan Wallace, University of California, Santa Barbara

*Monica Dus*, “The Power of Intellect and Freedom”  
Mentor: Prof. Tammy Nyden-Bullock, University of the Redlands

*Richard Warfield*, “Must We Save Only the Phenomena?”  
Mentor: Prof. Daniel Rothbart, George Mason University

*Thomas Cochran*, “The Problem of Space”  
Mentor: Prof. Tim Crane, University College London

~ **Pacific University Undergraduate Philosophy Conferences ~**

**“Best Paper: Conference 2000”**  
*Richard P. Brown*, “James Sennet’s Epistemic Circularity and Soundness in Alvin Plantinga’s Ontological Argument”  
George Fox University

**“Honorable Mention: Conference 2000”**  
*Ian Hesketh*, “‘A Fish in Water’: Michel Foucault and Historical Change”  
Okanagan University College, British Columbia, Canada

**“Best Paper: Conference 2001”**  
*Anthony Reeves*, “Consciousness, Oppression and Resistance: A Hegelian Critique of Marx’s Assumption of a Proletarian Revolution”  
University of Portland

**“Honorable Mention: Conference 2001”**  
*Stephanie D. Pierce*, “Bas C. van Fraassen: Theory of ‘The Pragmatics of Explanation’, or Pragmatic Theory of Explanation”  
University of Washington

**“Honorable Mention: Conference 2001”**  
*Wendy Lynn Clark*, “A Socratic Unity of the Virtues”  
Lewis & Clark College