

# PRAXIS

Volume 1, Issue 1—Spring 2008

## Contents:

Interview with Ronnie  
de Sousa

Probability and  
Certainty  
Jonny Blamey

The Internal  
Constraint on  
Humanitarian  
Intervention  
Neb Dobos

Metaethical Moral  
Relativism  
Alexandre Erler

The Nature of Applied  
Mathematics  
Davide Rizza

The Role of Affective  
States  
Amanda Taylor

A Problem With  
Defining Testimony:  
Intention And  
Manifestation  
Mark Nicholas Wales



MANCHESTER  
1824



- 2**      **INTERVIEW WITH RONNIE DE SOUSA**
- 15**     **PROBABILITY AND CERTAINTY:**  
Jonny Blamey, Kings College London
- 34**     **JUSTIFYING HUMANITARIAN INTERVENTION TO THE PEOPLE WHO PAY FOR IT:**  
Ned Dobos, University of Melbourne
- 52**     **METAETHICAL MORAL RELATIVISM AND THE ANALOGY WITH PHYSICS:**  
Alexandre Erler, Lincoln College, Oxford
- 69**     **THE NATURE OF APPLIED MATHEMATICS: REMARKS ON FIELD'S VIEW:**  
Davide Rizza, University of Sheffield
- 88**     **THE ROLE OF AFFECTIVE STATES IN PHILOSOPHICAL INQUIRY:**  
Amanda Taylor, University of Durham
- 105**    **A PROBLEM WITH DEFINING TESTIMONY: INTENTION AND MANIFESTATION:**  
Mark Nicholas Wales, University of St Andrews

**EDITORS**

Vasco Castela (Manchester, UK)  
Paula Satne Jones (Manchester, UK)

**ART EDITOR**

David Alexander Smith (Goldsmiths, UK)

**ADVISORS**

Peter Goldie (Manchester, UK)  
Ronnie de Sousa (Toronto, Canada)

**REFEREES**

Laurens van Apeldoorn (Oxford), Johan Blok (Groningen), Phil Brown (Manchester), Glen Carruthers (Adelaide), Elena Casetta (Turin), Dan Cavedon-Taylor (Birkbeck), Richard Clarkson (Durham), Christine Clavien (Lausanne), Dave Cochran (St. Andrews), Andy Cochrane (Leeds), Mathilde Cohen (Columbia/Paris X), Thomas Connor (Manchester), Richard Corrigan (UC Dublin), Jack Darach (King's College London), Marcello Ghin (Paderborn), Raphael Fiorese (Monash), Chloe Fitzgerald (Manchester), Alfredo Gaete (Manchester), John Guelke (Manchester), Hans de Haan (Erasmus), Luke Jerzykiewicz (Clark), Kile Jones (Glasgow), Carolyn Kane (NYU), Christos Kyriacou (Edinburgh), Jouni-Matti Kuukkanen (Durham), David Lawton (Manchester), Siu-Fan Lee (King's College London), Jean-Paul de Lucca (Malta), Bertha Manninen (Arizona State), Daniel Pointon (King's College London), Jean-Philippe Ranger (Ottawa/Paris IV), Michael Rush (Manchester), Glenda Satne (Buenos Aires), Jelle de Schrijver (Ghent), David Alexander Smith (Goldsmiths), Michael Sollberger (Lausanne), Stéphane Symons (Leuven), Geoff Stevenson (Manchester), Michael-John Turp (Durham), Lee Walker (Manchester)

## **INTERVIEW WITH RONNIE DE SOUSA**

Ronald de Sousa grew up in Switzerland and in the United Kingdom. He obtained a BA from Oxford and a PhD from Princeton. Based at the University of Toronto since 1966, he has lectured in more than twenty countries on the emotions, human rationality, philosophy of biology, philosophy of sexuality, ethics and aesthetics. He is currently an Emeritus Professor at the Department of Philosophy of the University of Toronto and a Fellow of the Royal Society of Canada.

**Praxis: How do you see the relation between philosophy and science, in terms of their respective tasks?**

**Ronnie De Sousa:** There is, of course, no simple answer. Here are four tries.

(i) Like religion, philosophy aims at a large scale vision of the place of humans and consciousness in the universe. But like science and unlike theology, it aims to establish conclusions on the basis of rigorous argument grounded ultimately in either necessary truths or empirical evidence.

(ii) Science makes progress on any given question when enough people can agree on what would count as answering it. Until then, questions are “philosophical”, which means there lacks an agreed method for tackling them. This, by the way, may account for the pejorative tone in which “It’s a philosophical question” is often uttered. When methodological consensus arises, a science branches off. Mathematics, Physics, Biology, and Psychology have all been bred by philosophy before establishing themselves as sciences; not so, as yet, Aesthetics and Ethics.

(iii) A third answer, favoured by many philosophers, is that philosophy deals with conceptual issues and necessary truths rather than empirical ones. That view has often been used to insulate philosophy from science, particularly in those areas that

concern normative issues. But since Quine's "blurring of the supposed boundary between speculative metaphysics and natural science" in his famous 1951 "Two Dogmas of Empiricism", the distinction is no longer sharp, and I myself side with those philosophers who seek once again to merge philosophy with science. I am encouraged in this by the thought that only religion beats philosophy, in the deplorably high ratio of certainties claimed to agreement secured.

(iv) This explains, incidentally, a cultural cleavage among philosophers, between those that *read aloud* their papers at conferences and those who give *talks*. For the latter, philosophy aims at the condition of science; like scientists, therefore, they want to get across *ideas*, and ideas can be explained in many different ways, so that others can in turn take them home and explain them in their own words. The former, by contrast—especially those that pride themselves (like Peter Sellers' shrink who gets to the strip club before his patient by "following him very fast") on being "*post-modern*"—regard philosophy as a branch of literature. Hence, having no actual ideas to convey, they need to get the wording just exactly right.

### **Emotions are a hot topic in both philosophy and science today. Why?**

Historically, philosophers have been very interested in emotions. Aristotle, the Stoics, Descartes, Spinoza, Hume, to name some, all had important things to say about the emotions. But emotions are theoretically messy; they were avoided, for a few decades, by analytic philosophers intent on making things neat. In psychology, the dominant ideology of behaviourism in the first part of the twentieth century and the focus on cognitive mechanisms in the second consigned emotions to the role of disrupters and spoilers. Only recently have methods become available in neuroscience to look into specific mechanisms in the brain that are responsible for emotions. (William James, the first anglophone philosopher to get interested in emotions in the 19th Century, had argued that there were no special brain centres

devoted to emotions.) As so often in science, then, what is done is driven by what can be done—by the opportunities afforded by technology.

Fashions in philosophy are perhaps less accountable. But what now attracts some of us to the study of emotion is the very feature that keep others away: sheer messiness. To study emotions, you need to think about every domain of philosophy from ontology through semantics, philosophy of mind, ethics, and rationality theory to aesthetics. The rapprochement of some philosophers with science is also a factor. Exciting results have poured in from scientific work. Though many still resist it, assimilating it to naive sociobiology, evolutionary psychology has thrown fresh light onto the genealogy of social and moral emotions. And how could philosophers resist the challenge thrown out by the findings of neuroscience? To cite just one example: there has been intensive study of the notorious “trolley problem”, in which one is asked whether one should do something that would result in the death of one innocent person in order to save five. Responses vary depending on the emotional cost of actually pushing someone to their deaths, as opposed to merely causing a mechanism to send the trolley down one line instead of the other. What does this tell us about the place of emotion and principle in ethics? And what philosophical moral are we to draw from the fact that subjects with frontal lobe lesions respond in a more consistently utilitarian way to both scenarios? Whatever the answer, a philosopher concerned with ethics will be drawn in.

**On your website, you list 10 things you believe in. One of the items on the list reads “We are animals, and everything worth knowing about us derives from biology”. Yet, in your latest book, *Why Think?*, you argue that rationality and language “allow the individual to offer a more or less anarchic resistance to the impersonal destiny embodied in the vestigial teleology of natural selection”. How is this second claim compatible with the first, and with your sympathy with evolutionary psychology?**

The answer to this could be put in the following slogan: *the truth of existentialist metaphysics is a fact of biology*. To explain this, I need a preamble about the meaning of *possible*.

Maynard Smith and Szathmáry (v. *The Origins of Life*, 1999) list eight “critical transitions” of evolution. Their list includes the formation of the first autocatalytic system that marks the beginning of life itself; the codification of information in DNA; the cell; sexual reproduction; and so forth. The last item is *language*. Each of these transitions marks a threshold at which the number of available possibilities is raised by many orders of magnitude. Each threshold, in other words, *creates a vast range of new possibilities*. In a strictly logical sense, of course, one might claim that possibilities don’t evolve: whatever is logically possible is necessarily possible. Possibilities, like all self-respecting philosophical truths, are eternal. But there is a more interesting and important sense in which not every logical possibility is a possibility *for* something or someone at a particular time: possibilities for a given person—or species—can begin (and perhaps cease) to exist. In that sense, new possibilities were indeed created by each of the critical transitions of evolution. Call these *potentialities* rather than possibilities, if you like, in tribute Aristotle’s distinction between different senses in which a baby or an adult “can talk”. It is possible in the fullest sense for a normal adult to talk; for an infant, it is possible in that she is expected to become an adult. We could add that it’s not *logically* impossible for insects to talk, as in the fable of the ant and the grasshopper; but logical possibilities aren’t the kind that can come into being.

Now to explain why our capacity to resist or even “transcend” the “vestigial teleology” of nature is itself *derived* from our biological nature. Note first that the teleology of nature is “vestigial” because it doesn’t involve any genuine goals. Instead, nature behaves *as if* its sole goal was replication. Its tool in achieving that goal is the diversification of forms, both at the genotypic and at the phenotypic

level. Language is just the latest of those increasingly ingenious tools. As it happens, however, that tool can be turned against nature's goal itself. For once we start to describe, to argue about, and to reason, we discover that our own goals need have nothing to do with those of nature. That is what it means to have escaped, in reward for tasting the fruit of the Tree of Knowledge of Good and Evil, from the prison that was Eden.

It's still true, of course, that *most* of what we do and think is done from motives we neither know nor understand. Yet deliberation yields a huge range of possible new values independent of the teleology inherent to life. (It's curious, incidentally, that the Vatican, while so keen on our spiritual freedom from biological determinism, rejects the very emblem of that freedom—the capacity, as Monty Python put it, to “put a little rubber thing on the end of my John Thomas.”)

We are animals, but we talk. Hence our relative freedom from the bare, stark goals of life. That freedom is part of our biology. Thus we become, in the crucible of gossip, self-conscious thought, reason-giving, and debate, the first beings for whom “existence precedes essence”.

### **Are there moral emotions? If so, how do they relate to value?**

Some philosophers restrict the term “moral emotion” to those that foster social cohesion. The main candidates are then compassion, guilt, and anger. The first appears to motivate altruistic acts; the second pair represent apt responses in those who offend and those offended against. In accordance with the commonly held doctrine that emotions imply action tendencies, it's assumed that people will be moved by these emotions to do things that reinforce social bonds. Among other emotions, some, such as love and grief, can be made out to be derivatively important for morality, in that love selects targets of altruistic acts and grief is a

sort of warrant of love. Still others, such as jealousy, are sometimes considered in the same light, though I am more inclined to think of it as a pathology of love, driven by biological imperatives inimical to the development of the best potential for human intimacy and social life. (On the pathology inherent in the ideology of love, as exemplified by wife-killers, see *In the Name of Love*, the forthcoming book by Aaron Ben-Ze'ev).

There are also thoroughly nasty emotions, such as envy, spite, and despair, which have been held to be redeemed neither by the motives they might induce, nor by any intrinsic pleasure they bring, nor by any contribution they make to correct ethical evaluations. As implied by my answer to the previous question, there is nothing paradoxical about the idea that some emotions might be irredeemably nasty: we need only suppose that their personal and social drawbacks are outweighed by objective biological ends with which they are associated.

A radically different perspective, broadly speaking aesthetic, starts from the fact that if there were no emotions, nothing would matter. Considered in that light, the connection between emotions and values is intrinsic and fundamental. But the values in question are not necessarily moral: whether a value is “moral” depends on what *ought* (not necessarily morally!) to count as moral. Aristotle’s version is *inflationary*: if all human potentialities contribute to making life worth living, none is irrelevant to “ethics” in the broadest sense; moral education then consists in learning to feel the right emotions to the right extent on the right occasions. All emotions are then moral or immoral, but none is amoral.

By contrast, a *minimalist* view of ethics restricts moral concerns to the prevention of harm and injustice. On that view, to accept that every emotion is in some sense tied to a value is merely to say that for every emotion there is a standard of appropriateness. I call that standard an emotion’s “formal object”,

and it is specific to each distinct emotion. Since not every value is moral, neither is every emotion. An emotion may be appropriate to its formal object, without thereby being one that a moralist ought to recommend. (Similarly, a good economist or a good thief isn't necessarily a morally good person.)

A moral minimalist, then, can regard all or most emotions as apprehensions of value, without denying that some emotions are amoral and others even immoral.

### **In what sense can one say that emotions are rational?**

Rationality, as I understand it, is in the business of maximizing the likelihood of success. Its various forms depend on the relevant mode of success and failure appropriate in different types of cases. Whenever we assess something as more or less rational, we therefore have in mind — or perhaps should bring to mind—some norm or standard of success. In the case of epistemic rationality, any way of acquiring belief, whether by intuition, perception or inference, can be assessed as more or less likely to result in true belief. For practical rationality, what the success of any particular action might be depends, of course, on the goals the action sets itself; since those tend to be complex and any action's outcomes are partly unpredictable, practical rationality is more difficult to define. So is the rationality of desire, since the incompatibility of desires doesn't follow from the incompatibility of their objects. Emotions involve all three of beliefs, action tendencies, and desires, in addition to their respective formal objects as appraisals. The appropriate standards of rationality for emotions are therefore the most complex of all. For what is it for an emotion to be "successful"? Four domains of assessment are relevant. First, an emotion might be said to be rational, in a derivative sense, as the beliefs that constitute its cognitive content are more or less epistemically rational. Second, you could think of assessing an emotion's rationality as derivative from the rationality of its attendant desires: and here it is crucial to remember that mere incompatibility

of *states desired* does not make the *desires* inconsistent. Third, one might assess an emotion in terms of practical aims which it overtly or covertly pursues. A fit of anger may function to affect another's behaviour, and its practical outcome could be judged accordingly. That too is a derivative sense of rationality, though it can be of major importance in human interactions. (Thus the deterrent effect of an emotional outburst will have its full effect only if the opponent is convinced that you are mad enough to act out of spite.) Most interesting and difficult is the fourth domain, which alone measures emotional rationality intrinsically rather than as it derives from component parts of emotions. This is the *axiological* domain of values apprehended by our emotional responses. Considered as part of that fourth domain, then, the rationality of an emotion will be gauged in terms of the intrinsic quality of the specific value that it apprehends, considered not as a means to any goals, but in terms of its adequacy to its object. And that adequacy is not unrelated to its various functions and components, but it's not reducible to them. From the biological point of view emotions, like everything else in our mental functioning, must owe their nature directly or indirectly to some practical function; but values transcend function, as gastronomy transcends nourishment and sex transcends reproduction.

**Do all emotions have formal objects, and is it by their formal objects that we can distinguish them?**

Talk of formal objects is often deemed suspect: when I try to explain it I often feel a certain resistance. But I regard it as important. So let me try and make the idea as plain as I can.

It's easiest to begin with formal objects of states that are not emotions: belief and desire. That sets up the case for emotions, with the interesting stuff stemming from both analogies and disanalogies with the cases of belief and desire.

The root idea is that some mental states can be distinguished in terms of their aim or point. I mean that in the sense in which Frege seems to have used it when he wrote “The word ‘true’ indicates the aim of logic as does ‘beautiful’ that of aesthetics or ‘good’ that of ethics” (Frege, G. 1956. “The thought: A logical inquiry.” Trans. A. Quinton and M. Quinton. *Mind* 65:289). The aim or point of a state is the state’s characteristic conditions of success as a state of that kind. So when someone *believes that p*, that state of belief is successful, in its own terms, iff *p* is true. If someone *desires that p*, the success of that state lies not in the truth of *p*, but its goodness or desirability. So the formal object of a state is given by the most trivial answer possible to the question ‘Why do you S O’ for a state S and its intentional object O. Example: “Why do you believe that p? —Because it’s true!” “Why do you want *x*? —because it’s good!”. In other words, the notion of formal object rests on the assumption that there is a systematic analogy between what is wrong with believing something false and what is wrong with desiring something undesirable. Although the truth of *p* would *satisfy* the desire, there is nothing odd about desiring what is not the case. There is something odd, however, in desiring what is not desirable.

Admittedly, the first case is a less controversial than the second. And the case of emotions may seem even murkier: but that’s precisely why it might be interesting. The tendency to contrast emotions with belief and desire, as if they were all on the same level, suggests that if there is a single formal object for each of belief and desire, there will be a single formal object for emotions. So what is philosophically interesting about the formal object of emotion is precisely that *there isn’t one*. If one asks the same question of any emotion: what are the conditions under which it is somehow in the most trivial sense *right* for the circumstances? What immediately becomes apparent is that there is no standard answer that applies equally to all emotions. The trivial answer to “Why are you angry at *X*?” is “Because *X* did me wrong.” The trivial answer to “Why are you surprised at *O*?” is “Because *O* was

unexpected”; to “Why are you afraid of O?”, “Because O is frightening”. And so forth. In each case, the property *F* trivially attributed to the object of the emotion in question is its formal object (and the emotion is appropriate in the specific circumstance iff the object itself actually has property *F*.)

From this it is immediately apparent that emotions are indeed distinguished by their formal objects, but that this doesn’t actually tell us anything very interesting about each emotion that we didn’t know. It does, however, tell us that no there is only a formal but no substantive *general* answer to the question of what makes an emotion appropriate.

### **Could scientific enquiry show that this view of formal objects was not correct?**

The notion of formal object as I’ve described it falls squarely in the category of conceptual analysis. So it’s a good example of something that won’t be upset by a simple counterexample or pinpointed scientific finding. That is because it belongs to the systematic level of theorizing about emotions; but it is also partly for the more boring reason that this view of formal objects needn’t insist that *all* emotions have one. Some things which we commonly call emotions, notably euphoria and depression, simply don’t have any standards of appropriateness. Some people interpret this as a reason for not including them in a list of emotions. I don’t have strong feelings about this, though it seems to me on the whole that they have enough in common with emotions to warrant the term, despite lacking formal objects (and objects of any kind, for that matter, unless “the world as I see and feel it” is an adequate object). As neuroscience explores the brain concomitants of experienced emotions, we may gather new reasons for thinking that some emotions are more “basic” than others, in the sense that—according to the research of Panksepp and LeDoux, among others—there is a consilience of neural circuits, hormonal activity, behavioural tendencies and phenomenology that is specific to each. In these cases, the formal object would relate to the function of those syndromes: to understand

in functional terms what the frightening is, is also to understand what the emotion of fear is. (Depression, I am led to infer, doesn't have a "biological function"; if it is found to have one, then I'll have to withdraw the observation that depression has no formal object.) If, as some philosophers and psychologists have claimed, there is actually no such thing as a basic emotion, and no such natural kind as emotions in general, then the fact that emotions in general don't share a formal object will seem a lot less interesting. And since the role of formal objects in the philosophy of emotion is a conceptual, structural one, its validity as a concept stands or falls with the use that can be made of it in organizing our thinking about emotions.

**Tell me about three pet hates of yours - three stubborn presuppositions in philosophy that need to be challenged, and why?**

By definition, all interesting presuppositions are stubborn, until they are unmasked, and then they are no longer presuppositions, but theses to be defended, attacked, or dismissed. So the question can't really be posed in such a way as to be answerable in a non-trivial way. Of course there are people who believe things I don't think are true, but if someone is aware of them they are no longer "presuppositions."

If you think that's a pretext, you're probably right. The real reason is that I can't think of three pet hates that wouldn't be obvious and banal—such as, for example, the obvious and the banal. And yet perhaps I can just descry one thing that might be, at least in my own mind, in transition between a presupposition and a thesis. It's the role in philosophy of *temperament*. Some recent research seems to indicate that liberals and conservatives have distinctly different personality profiles (in terms of standard dimensions of personality theory, particularly conscientiousness, openness and neuroticism), which themselves are starting to be seen as linked through such neurobiological factors as serotonin transport to genes. (See, eg., *New Scientist*, Feb.

2 2008) In much the same way, I suspect that on some deep divisions of philosophy people's views are determined not by arguments but by temperament. This is nicely summed up in a quote a friend recently sent me from Natalie Clifford Barney: "A scholar's heart is a dark well, whence deeply buried emotions bubble up in the guise of arguments." No argument will ever be persuade a consequentialist to become a deontologist, or vice versa. We should pay more attention to the philosophical implications of that fact.

Let me add just one pet peeve. I often wonder, on reading this or that journal article, about the *further point* of analytical exercises in philosophy. Placing one's neat analytical effort in the framework of a broader philosophical problem is perhaps discouraged by the urgency of the publishing imperative; so the system is to blame rather than any particular philosopher's cast of mind. But while I endorse the ideal of rigour in philosophy, I think we should remember that the point of doing philosophy is to change our vision of the world as a whole—which doesn't preclude seeing it, for good or for ill, as deeply fragmented.

### **What is the meaning of life?**

Here is one place where philosophical temperament is the key to each philosopher's preferred answer, or preference for not answering. For my part, I am bemused by most people's reluctance to accept the essential randomness of life and fortune, which seems to me to be the core of the curious survival of religion in supposedly civilized countries. It's merely silly, of course, to think events in my life can't have a meaning for me unless Life has a meaning in itself. But there's more to it than a silly mistake: and that more is a matter of temperament, and perhaps of genes. My own temperament leads me to feel it would be supremely oppressive to think of life as having a "meaning", or to believe that what the fabled mathematician Paul

Erdős called the Supreme Fascist had *a plan for me*. But many, quite obviously, have exactly the opposite stance: they just can't believe that this or that event in their own life was not significant, meant to be, *about them*—like the airplane passenger who announced that now he really believed that God existed because in a crash in which several others were killed, *he* had survived. It is a step to wisdom is to recognize, with the existentialists, that “the fault, dear Brutus, is not in our stars, but in ourselves...” But a further step is to see that fault and merit are not in ourselves either but in the awesome churning of chance and necessity, in the emergence, out of randomness, of patterns unrepeatable and uniquely precious, including ourselves and our friends. In the words of Wallace Stevens, “Death is the mother of beauty”:

*Shall she not find in comforts of the sun,  
 In pungent fruit and bright, green wings, or else  
 In any balm or beauty of the earth,  
 Things to be cherished like the thought of heaven?  
 Divinity must live within herself:  
 Passions of rain, or moods in falling snow;  
 Grievings in loneliness, or unsubdued  
 Elations when the forest blooms; gusty  
 Emotions on wet roads on autumn nights;  
 All pleasures and all pains, remembering  
 The bough of summer and the winter branch.  
 These are the measures destined for her soul.  
 But that's just what it is in my temperament to feel.*

## PROBABILITY AND CERTAINTY

JONNY BLAMEY

KINGS COLLEGE LONDON

### *Abstract*

Probability can be used to measure degree of belief in two ways: objectively and subjectively. The objective measure is a measure of the rational degree of belief in a proposition given a set of evidential propositions. The subjective measure is the measure of a particular subject's dispositions to decide between options. In both measures, certainty is a degree of belief 1. I will show, however, that there can be cases where one belief is stronger than another yet both beliefs are plausibly measurable as objectively and subjectively certain. In ordinary language, we can say that while both beliefs are certain, one belief is more certain than the other. I will then propose second, non probabilistic dimension of measurement, which tracks this variation in certainty in such cases where the probability is 1. A general principle of rationality is that one's subjective degree of belief should match the rational degree of belief given the evidence available. In this paper I hope to show that it is also a rational principle that the maximum stake size at which one should remain certain should match the rational weight of certainty given the evidence available. Neither objective nor subjective measures of certainty conform to the axioms of probability, but instead are measured in utility. This has the consequence that, although it is often rational to be certain to some degree, there is no such thing as absolute certainty.

### *The No Certainty Principle and the Maximum Likelihood Principle*

Many Bayesians hold to the principle that a belief in a predictive empirical proposition should never be certain, or nearly certain, unless it is logically entailed

by the evidence (Good 1950, p. 36; Kaplan 1996, p. 93). Another principle is that you should proportion your belief to the evidence, and if the evidence consists in observed relative frequencies then you should believe a prediction to the degree that matches the relative frequencies, or to put it another way, to a degree that makes the observed frequencies retrospectively maximally likely (Fisher 1922, pp. 310 - 312; Popper 1974, p. 168; Good 1950, p.77).

These two principles, the ‘no certainty principle’ and the ‘maximum likelihood principle’, clash when all observations to date have been uniform, in other words, when the relative frequency has been 1 : 0. For example, if I have seen many ravens, all of which have been black, and my lover tells me that she has made a gift of a raven for me that is hidden in a casket, then the maximum likelihood principle tells me to be certain that the raven in the casket is black.

Predicting that the probability of being black given being a raven is 1 makes my previous observations most likely, so I should predict with probability 1 that the raven in the casket is black. However, the “no certainty principle” does not allow this. Since being black does not follow logically from being a raven, we should never assign probability 1 to black given raven. If we did, then it would be impossible, according to Bayesian conditionalization, to ever change our belief (Howson & Urbach 1989, pp. 284- 287; Williamson 2000, pp. 204, 218-220; Kaplan 1996, pp. 91 – 93). This clash of principles can be seen as a problem for statistics and mathematics, but it is clear that if we consider the philosophical problem of scepticism, we get this same duality of advice. It seems on one hand that we should be certain of many things, but on the other that we should be certain of nothing.

### ***Consumer and Producer Accounts of Belief***

These two principles are input driven, born out of a priori reasoning on the nature of evidential support. If we use Ruth Millikan’s (1993, p. 126) terminology, they are

both *producer* principles. Certainty is defined as a relation to a class of evidential support. If we are to resolve the dispute in a non circular fashion we need a consumer account. What is certainty *for*? In what way is certainty consumed by our mental economy? What does certainty *do*?

In decision theory, the output of degree of belief is usually defined in terms of the odds of an indifferent bet. In this sense, degrees of *confidence* are measured by probabilities. It is assumed that the more confident you are that a proposition is true, the greater the odds at which you would be prepared to bet. Certainty, on this view, is the highest degree of confidence. Certainty is when you would bet on  $p$  whatever the odds, and refrain from betting against  $p$  whatever the odds.

If odds are the only measure of subjective degree of belief, then once a subject is indifferent between a bet at 1 : 0 odds there is literally no way of being more certain, the measure not allow it, thereby giving certainty the character of infallibility. But as several people have pointed out in different contexts, once this threshold has been transcended, it is still possible to be more or less certain (Millikan 1993, p. 253; Williamson 2000, p. 99).

If a man publicly claims that his parrot is two hundred and fifty years old I would bet at any odds that he was speaking falsely and at no odds that he was telling the truth. In the consumer sense, in terms of output, it seems I am certain that his parrot is less than 250 years old. The maximum likelihood principle seems to endorse this, since assigning probability 1 to parrots dying before they reach 250 years old is the best way to account for the fact that I have never seen nor heard of any parrot approaching that age. But when a second man claims that his great grand father is still alive at the age of 350 years old, I find that I am even more certain of the negation of this second proposition. Yet how is it possible to

be even more certain than certain? Can there be something more probable than probability 1?

### *Ramsey's Consumer Measure*

Frank Ramsey in 1926 wrote a remarkable paper called 'Truth and Probability' where he presented a consumer theory of subjective probability. He compares degrees of belief to time intervals in physics. He argues that 'the degree of belief is just like a time interval; it has no precise meaning unless we specify more precisely how it is to be measured.' (Ramsey 1926, p. 63). He goes on to argue that in order to have a producer account of probability, we must have a precise measurement for what it is to have the corresponding degree of belief in those propositions, i.e. a corresponding consumer account.

After considering the problems associated with using the lowest odds at which a subject is prepared to bet on  $p$  as a measure of the subject's degree of belief that  $p$ , namely, the subject's reluctance or over eagerness to bet and the diminishing marginal utility of money, Ramsey developed a measure for utility based on preference. The subject's preferences between hypothetical worlds are ordered on an interval scale that I will indicate using capital letters A, B and C. Ramsey's genius was to transform the ordinal scale into an interval scale by using the concept of an ethically neutral proposition. An interval scale is one where, though zero and the unit is arbitrary, once the unit is decided, the interval between any two values is commensurable with the interval between any other two values. With an interval scale of utility, capitals 'A', 'B'...etc. can represent the utility to a subject of a set of worlds between which he is indifferent.

An interval scale can quantify the ratio of  $A - C$  to  $A - B$ . Time is also measured on an interval scale. For example, because time is measured on an interval scale, we can state clearly that the time difference between 2pm and 4pm is twice the

time difference between 5pm and 6pm. This is important in utility terms because it means we can give an exact value to the ratio of the utility difference between losing a bet and not taking the bet to the difference between winning and losing the bet.

### *Ramsey's measure for degrees of belief*

Having established a method for measuring utility, Ramsey (1926, p.75) defined the following precise measure for a subject's degree of belief that p:

If the subject is indifferent between options:

1. A for certain, or
2. B if p and C if  $\sim p$ .

Then the subject's degree of belief that p is equal to:

$$\frac{A-C}{B-C}$$

He then demonstrated from this basic definition and constraints of consistency over preferences that subjective degrees of belief should obey the basic laws of probability.

### *A Consumer model of Certainty*

On Ramsey's model, the consumer measure of certainty becomes the disposition to be indifferent between options 1 and 2 whenever  $A = B$ , since if one is certain that p the choice of options 1 and 2 becomes a straight choice between A for certain and B for certain. This is the prescription to act whenever there is any advantage at all to acting if p is true over not acting, regardless of the disadvantage of acting if p is false. C, which is the loss if the belief is false, has disappeared out of the equation. But C is still relevant to decision making.

Since we are looking for another dimension of measurement for degree of belief, we can take advantage of this hidden variable and use it to measure the level of

certainty when the degree of belief is at probability 1. Certainty is associated strongly with knowledge, and recent theorists have linked the truth of knowledge ascriptions to what is at stake in the situation. (Stanley 2005; DeRose 2002; Millikan, 1993 pp. 252 - 255). What is at stake in the situation can be defined precisely as the utility difference  $B - C$ , which is the difference in utility between acting if  $p$  and acting if  $\sim p$ . The more certain you are, the larger this value can be before electing not to act.

So we have got our output theory, or our consumer interpretation. I am more certain that the man is less than 350 years old than that the parrot is less than 250 in that the stakes on a bet would have to be a lot higher on the man before I would decline from accepting any odds. We can call the willingness to accept a bet at any odds a *guarantee*. Formally a guarantee can be defined as indifference between:

Option 1, A for certain, and

Option 2, B if  $p$  and C if not  $p$ , whenever  $A = B$ .

The amount of the guarantee will then simply be the difference between B and C, where C is less preferable than B. This amounts to C being a penalty for being wrong. At low stakes I would happily *guarantee* that the great grandfather is not 350yrs old (or is dead), in that I would be indifferent between option 1, not losing anything either way, and option 2, a penalty of £100, if the grandfather was 350 yrs old, and nothing, if he was younger than 350yrs old.

Since I am certain that the man is not 350yrs old, the two options have the same expected utility. I don't expect to have to pay £100 at all. The same goes for a £100 on the parrot being under 250yrs old. But as the penalty increases in value, I may opt out of the parrot guarantee sooner than I would opt out of the man guarantee. Doubts as to whether parrots could in fact live to incredible ages would begin to creep in at lower guarantees than doubts that there could be a 350yr old man. We

could in this fashion give an exact utility measure of my dispositional certainty in a proposition. It would amount to the precise threshold of loss at which I would cease to be indifferent between the two options. So at low stakes I am certain of both, but there is a range of stakes such that I would be uncertain that the parrot is not 250yrs old, but would remain certain that the man was not 350yrs old. The certainty threshold would be a utility measure and would not conform to the axioms of probability.

***Producer Theory for Probability: ratio of success to failure in experience***

What we need now is a producer interpretation, or an input theory. The most general and simple input theory for probability is that the degree of belief that an event type E will have property P is equal to the ratio of observed events type E with property P to observed events of type E with or without the property. If one in three men in your experience has been bald, the probability you should assign that the next man you meet is bald is 1/3. If all the men you have met have a heart, then the probability that the next man you meet will have a heart should be 1. This is a naïve frequency interpretation of probability. It is easy to show that these figures will conform to the axioms of probability.

There are many interesting questions as to how to type the predicted event, what to count as positive instances and so on, but the general principle remains the same. We can see that, as a general input theory for all epistemological issues, it is inadequate, for the very simple reason that it gives no extra weight to the sheer size of the sample from which the prediction is made. It also gives no extra weight to the quality of the sample, and the relevance of the sample to the event in question. For example, this input rule would prescribe the same actions whether i) I had only met six men in my life, two of whom were bald; ii) I knew that world wide one in three men are bald; or iii) I knew the next three men I was going to meet, but not in which order, and knew exactly one

of them to be bald. Intuitively these three samples are epistemically different, though each would recommend a degree of belief  $1/3$  that the next man I meet will be bald.

### ***Size Relevance and Depth of Experience Important***

So, when thinking about the degree of support an evidence sample lends to a prediction, it is not merely the ratio of cases that is relevant. There is also the size and relevance of the sample. In general, the broader and deeper one's experience, the better one's prediction. So a producer rule for certainty is that the greater the experience, the greater the certainty. Although the ratio of parrots under 250 years old to parrots in my experience, is equal to the ratio of men under 350 years old to men in my experience, (they are both 1 : 1); it is clear that my experience of men is greater in terms of relevance, depth and sheer number. If the input function for probabilistic degree of belief is the just a function of ratio, then there is room for non probabilistic variation in degree of certainty.

### ***Another Input Measure beyond Probability***

A fruitful place to look for such a measure is the cost of acquiring the evidence. Consider a repeatable experiment with two possible outcomes defined as successful or unsuccessful. The experience set can be expressed as the ratio of the number of successful trials to the total number of trials. When all the trials are successful we can predict with certainty (in the minimal sense) that a future trial will be successful. The degree of certainty can be defined simply as the number of trials. To convert this into a utility measure we can multiply the number of trials by the cost of performing a trial. In a case as simple as this, it is easy to see the logic. If the prediction of a trial demanded a high threshold, in that the penalty for being wrong was higher than the cost of performing more trials, then it would seem sensible to perform more trials. Whereas if the penalty for being wrong was less than the

cost of having performed the number of trials in experience, then it would seem sensible to go ahead and risk being certain. A similar line of thinking has been pursued by Good (1967):

“Our conclusion is that, in expectation, it pays to take into account further evidence, provided that the cost of collecting and using this evidence, although positive, can be ignored. In particular, we should use all the evidence *already* available, provided that the cost of doing so is negligible. With this proviso then, the principle of total evidence follows from the principle of rationality.”

In general Good is a proponent of the no certainty principle. His conclusion therefore does not apply to certainty. If we do not ignore the cost of “collecting and using the evidence” then collecting evidence will only pay if the advantage given by the information outweighs the cost. To make this decision we must have a value for the advantage given by the information. In an obscure note by Ramsey (1991, p. 285) entitled “Value, or weight of knowledge” Ramsey gives a formula for the value of knowledge that you already know, in other words the value of finding out that same information in relation to the expectation of advantage of an unknown proposition. If we take it that our current knowledge gives probability 1 to a prediction, we can use Ramsey’s formula to decide on the value of our current knowledge and whether it is adequate for the risk inherent in the prediction. The “value of knowledge” will of course be relative to interests. Given that there is no knowable difference between that which we know and that of which we are certain, (Ramsey 1991, p. 82) we now have both pieces of the jigsaw: an input measure for the value of certainty and an output measure in terms of the magnitude of the loss in error.

I know of a lot more ages of death of men than of parrots, and the cost of significantly increasing that knowledge is inestimably high with regard to men, but lower with

regard to parrots. (I could look into the life spans of parrots and increase my knowledge quite easily). So my certainty threshold *should* be as it is, higher with regard to the man than to the parrot. The literature on the interest relativity of knowledge ascriptions (Stanley 2005; DeRose 2002) would be much clarified by this model. That S appears to know that  $p$  at low stakes, but does not know that  $p$  at high stakes given equivalence of evidence would be explained by the certainty threshold lying between high stakes and low stakes. The high stakes situation would make seeking further evidence cost effective.

### ***Certainty Threshold Useful for Warrant Transmission***

The idea of a cut off stake size for certainty is especially pertinent when transmitting warrant through testimony. It is often the case that a more experienced and knowledgeable person will want to transmit the epistemic status of their prediction without having to say exactly what their experience consists in. For this reason there are many words in natural language which give an indication of whether the cut off stake size has been reached. Certainty is the most basic of these. It has often been pointed out that certainty is intuitively incompatible with less than 1 probability. The most famous illustration of this point is the lottery paradox whereby it feels unnatural and wrong to say that it is certain that a ticket will lose before the draw, however unlikely the ticket is to win (Kaplan 1996; Hawthorne *et al* 1999; Kyburg 1961). Yet certainty is on a sliding scale, there is nothing infelicitous in talking of increasing or decreasing levels of certainty. This paradox also applies to the term “sure”. Other words in this ball park are “knowledge” and “outright belief”. While not so clearly gradable, (Jason Stanley<sup>1</sup> denies that knowledge is gradable at all), both seem at once to denote a probability 1 and yet to be sensitive to non evidential features of the practical situation. Both Williamson and Millikan have noticed this link between probability 1 and variation due to stake size.

---

<sup>1</sup> This was from a talk he gave in St Andrews and London in 2006 called *Knowledge and Certainty*.

“Outright belief still comes in degrees, for one may be willing to use  $p$  as a premise in practical reasoning only when the stakes are sufficiently low” (Williamson 2000, p. 99)

“Unless the stakes are high, we do not require of the person who “knows that  $p$ ” because he has been told by so-and-so that he be ready to place extremely high stakes on the truth of  $p$ ” (Millikan 1993, p. 253)

The way to look at it is this: if you were in a betting shop and had to decide on whether to bet on  $p$ , there are two questions to be asked. “At what odds should you bet on  $p$ ?” and “how much should you bet?” These questions can be translated as: “how probable is  $p$ ?” and “how certain is  $p$ ?” The first figure will be a ratio that conforms to the axioms of probability calculus. The second will be an integer value with no absolute limit. Both these values have their place in ordinary talk, but in science people only use probability (of course other features of the data can be represented in some cases by statistical conventions like standard deviation, but in cases where the probability of an event is 1, standard deviation doesn’t have any meaning).

### ***Probabilism, second order Probabilities***

An alternative to the view I am putting forward here I shall for ease of expression call “Probabilism”. Probabilism is the view that any variation in prescriptive or descriptive degree of belief is a variation in probability that should conform to the axioms of probability calculus. A probabilist could object that the increase in the sample size simply increases the probability of the generalisation which issues the prediction. As the sample increases in size, Bayesian dynamics will ensure that the probability of the generalisation will tend towards 1 without ever reaching 1. This has the consequence that no predictions are certain since no generalisation is ever certain.

For example, suppose I have only ever seen and heard of black ravens. Each extra black raven increases my degree of belief in the hypothesis that all ravens are black. The more ravens I have seen, the greater my belief in this universal generalisation. Any singular prediction will have a probability conditional on the probability of the generalisation being true. This has the consequence that the probability of a prediction is always lower or equal to the probability that the generalisation is true. More formally, if the evidence “E” is counted as known such that E consists in a large number of sightings of black ravens and no sightings of other ravens; whereas the generalisation “H” is the hypothesis that all ravens are black, then  $(H|E) < 1$  by the no certainty rule.

If we know of a raven but we do not know its colour, then it seems that given E we should predict with a high probability that the unobserved raven is black (RB). What is  $P(RB | H)$ ? Surely it is the  $P(RB + H)/P(H)$  (Kaplan 1996 p.48). But in this case the  $P(RB | H)$  is less than or equal to  $P(H)$ . Of course RB might be true but H false. But in this case it is hard to see what relevance your degree of belief in the generalisation that all ravens are black has to your prediction that the raven will be black. The generalisation is redundant.

### ***Probabilism is Counterintuitive***

Probabilism is counterintuitive since a singular prediction is a lesser commitment than a generalisation. I am intuitively a lot more certain that the next person I meet has a heart than I am that all living humans have a heart. Secondly the generalisation is likely to be tailored to the prediction rather than the other way around. If I tell you that I am certain that that pill won't kill you I am not necessarily committing myself to a generalisation that all pills won't kill all people, or even that no pill of that kind ever kills anyone. Some aspirins may kill some people, some pills that find their way into aspirin bottles may kill anyone. What I am guaranteeing is that *this* aspirin won't kill *you*.

***Probabilism fails to distinguish between, probability 1 at low levels of certainty and probability of less than 1 at high levels of certainty***

I may be certain that if you go on this bungee jump you won't die on the grounds that a hundred people have been on this bungee jump and none of them have died. A probabilist says I shouldn't predict with probability 1 that you won't die on this bungee jump. I should instead leave some room for a possible increase in certainty. Let's say that I should predict with probability 0.999 that you will not die. But suppose there is another bungee jump where a thousand people have jumped, one of whom died? Should I predict that this bungee jump is *no more dangerous* than the one that was used less, but that has had a perfect safety record? What of a bungee jump where a hundred million people had jumped and one in every thousand had died? In this case I would be reasonably certain that the chance of dying was one in a thousand. In the first case the 0.999 safety figure is cooked up, plucked out of the air. Much better to say that the probability of dying is 0, *as far as we know*.

***Threshold Theories***

A plausible line of argument that a probabilist might give is that the term "certainty" as used by the folk is to be interpreted as a Bayesian degree of belief over some threshold less than 1 (Kaplan 1996, p. 93 – 101). But this is also problematic for the Bayesian who considers the degree to which evidence confirms a theory (T) to be the increase in probability that the theory is true conditional on the evidence. This has the consequence that the results of a scientific experiment have a probability 1. If the result of a scientific experiment was (e), and the prior probability of  $(T | e) = n$ , then the new probability of T should be n. This has the implicit assumption that if (e) is scientific evidence, then the probability  $(e) = 1$ . A reason to reject any threshold theory is that it would seem peculiar if the results of scientific experiments had a higher probability than certainty since even scientific observations are fallible. So if the results of scientific experiments have a probability 1, then propositions of which we are certain must have

a probability of at least 1. But there can be no probability higher than 1. Further more if we allow observations to have a probability of 1, and yet the threshold of certainty to be less than 1, we get an alarming failure of modus ponens.

Suppose for the sake of reductio that the threshold for certainty is 0.99. Suppose that  $P(h|e) = 0.99$  and is therefore certain and  $P(e) = 0.99$  and so certain. Then the probability (h) could be less than probability (e). For example if  $p(h) = p(h + e) = 0.9801$ . But this is absurd, since it has the consequence that (e) could be certain and the probability of (h|e) also certain and yet (h) be less than certain. For example let (h) be that Jones is mortal and (e) be that Jones is dead. Let us say that (e) is certain and (h|e) is certain and this means that they both have probability 0.99. Suppose also that it is certain that Jones is dead. It is possible, with the right prior probabilities, that the posterior probability that Jones is mortal should be 0.9801, and therefore not be certain. The absurd consequence is that it can be rational to be certain that Jones is dead, and to be certain that if Jones is dead then he is mortal, but to be not certain that he is mortal. These arguments are conclusive that if certainty implies any probability at all it implies probability 1. Probabilism therefore cannot explain varying degrees of certainty. This leaves variation in cut off stake size due to sample size and quality as the only account available.

### ***Generalisations are counterfactual supporting***

Yet another reason to reject probabilism is that the generalisations themselves must be counterfactual supporting since we can make predictions of counterfactual events of which we are certain. This has to be the case otherwise we could not use predictions in order to avoid doing actions with dangerous or fatal consequences. If I shoot myself in the head then I am certain that I will die, therefore I will not shoot myself in the head. Because generalisations have to be counterfactual supporting, the population over which they are quantified is infinite. Since any data set is going to be finite, the subjective probability that the generalisation is true can

never raise above 0 since the generalization can never be confirmed. (Popper 1974, p.63) Also, if the rationale behind the increase of probability of the generalisation due to increase in sample size is that the bigger the sample, the more likely it is to accurately represent the total population, then the rationale fails for infinite populations. (Strawson 1956, p. 253 -255) This is because the constitution of any sample of any size has a zero probability of matching the constitution of the total population since any number divided by infinity is zero.

### ***Probabilism as an Output Function is Insensitive to Risk***

So from the input side there is plenty of reason to endorse an independent measure of certainty which is a function of the size and quality of the sample, rather than to try and translate any increase in epistemic standing as an increase in a single probability figure. It is not hard to find criticisms of probabilism on the output side as well, and these can easily be tackled by having an independent certainty scale. The focus of the criticisms is that subjective probability is insensitive to the subject's attitude to risk and to the diminishing marginal utility of money. The most famous criticism is from the Allais problem (Allais 1953). Allais showed that intuitively non gambles are more valuable than gambles even when the expected utility is the same.

### ***Solution to Allais Problem through degrees of Certainty***

My strategy is to use  $B - C$ , as a separate measure of certainty.  $B$  is the value of an action conditional on  $p$  and  $C$  is the value of the same action conditional on  $\sim p$ . Think of  $B - C$  as representing what is at stake in assuming  $p$  with regard to a specific action. On my view there is no need to assume that the rational degree of belief will remain the same as  $B - C$  varies in magnitude. The assumption that the magnitude  $B - C$  has no affect on the rational degree of belief is demonstrably false, since if  $B - C$  is 0 then the truth or otherwise of  $p$  becomes irrelevant to the choice of whether to act or not, and the subject's degree of belief that  $p$  becomes

undefined by her choice of actions. This is a purely mathematical result since whenever the denominator is 0, the value of the fraction is undefined. The Allais problem exploits this and presents a series of options one of which has  $B - C$  as zero. It seems intuitively rational to favour the option where  $B - C$  is zero, especially when  $B$  and  $C$  are themselves of great positive utility. (Anecdotally Savage found himself preferring the sure gain over the higher expected utility, going against his own theory. However, instead of amending his theory, he amended his intuitions.) Seeing this mechanism it is easy to come up with Allais-type problems for subjective probability. Given the choice of either:

Option 1: WIN £1 million if two sixes are thrown on a fair die and WIN £1 million if two sixes are not thrown; or

Option 2: WIN £50 million pounds if two sixes are thrown on a fair die but LOSE £100 000 if two sixes are not thrown.

It would be a reckless person who chose option 2 even though the expected utility is much higher. Please ask yourself if you would really view it as more rational to go for a  $1/36$  chance of winning £50 million and a  $35/36$  chance of losing £100 000, over a certain million?

Also, if one could bet as much as one liked on the single toss of a coin at odds of 11:20 in one's favour, how much should one bet? Surely not every penny to one's name. Commentators are split on these issues. Ramsey (1926, p. 72), Friedman & Savage (1948) and Jeffrey (1970, pp. 161 - 2) were persuaded that these problems can be solved by recourse to a robust utility theory. Ramsey added the value  $A$  to counterbalance the subject's attitude to bets. He presumed that his utility theory would do the rest.

Others think these problems are decisive and that subjective probability theory should be seriously modified or abandoned (Kaplan 1996, p. 172; Allais 1953). These

latter tend to ignore the solution inherent in adding the counterbalancing value  $A$  for not accepting the bet. I am suggesting a middle ground. This is to drop the assumption that the rational bettor should accept the same range of odds whatever the magnitude of the difference between the loss and the gain. This assumption is counterintuitive, as the Allais problem shows. It is also demonstrably false since when  $B - C = 0$  then the bettor is rational to accept any odds.

Furthermore, it is not argued for by a Dutch book argument since there is no way a cunning bettor can have simultaneous bets at different stake sizes, since this would be in effect to have a single bet with the sum of the various stake sizes. (If I bet you £10 that it is raining and £100 that it is not raining, both at 1 : 1, I am not making two bets; I am making one bet of £90 that it is not raining). A final consideration in favour of dropping this assumption is that the truth of knowledge ascriptions seems to be both sensitive to practical interests and yet be negated by sub unitary probabilities. If probability itself varies with stake size, then the probability necessary for knowledge can be invariant at 1 while the knowledge ascriptions themselves vary with stake size.

### ***Conclusion***

The conflict between the maximum likelihood principle and the no certainty principle can be resolved if we consider the practical prescriptions given by certainty. One can rationally be certain when the stakes are low but become less than certain when the stakes get higher. The exact point at which one ceases to be certain is the *degree of certainty* as opposed to the degree of belief measured as a probability function. The degree of belief is the exact odds at which one would cease to bet. The degree of certainty is the exact stakes at which one would cease to bet at any odds.

The degree of certainty so defined is an output function, it is certainty seen from a consumer point of view. If the consumer needs two dimensions, then the producer

must produce two dimensions. Degrees of belief can be produced by considering the ratio of successes to failures in experience. Degrees of certainty can be produced by considering the size and relevance of the experience. The no certainty principle can then be seen to be referring to the maximum degree of certainty.

Given the nature of value and experience, there is no absolute certainty so defined, so the no certainty principle is correct. The no certainty principle is only incorrect in so far as it suggests that we should never assign a probability 1 to a singular prediction. With these two dimensions we can give a normative account of the interest relativity of knowledge ascriptions. We can also concede to the sceptic that we can never be absolutely certain of anything. But since it is only the kind of certainty that is relative to our interests that is necessary for knowledge, we do not have to concede to the sceptic that we don't know anything. We are and should be certain of many things which are of little or no importance, but of the things that really matter it is wise to keep an open mind.

### ***References***

- Allais, Maurice, 1953. "Le comportement de l'homme rationnel devant le risqué: critique des postulats et axiomes de l'école Américaine", *Econometrica* 21: 503 – 46
- DeRose, Keith, 2002. "Assertion, Knowledge, and Context", *Philosophical Review*, Vol. 111, No. 2. Apr. pp. 167-203
- Fisher, R.A., 1922. "On the Mathematical Foundations of Theoretical Statistics", *Philosophical Transaction of the Royal Society. A*, 222: 309 – 368
- Friedman, M & Savage, L. J., 1948. "The Utility Analysis of Choices Involving Risk" *Journal of Political Economy* 56, pp 279 -304
- Good, I.J., 1967. "On the Principle of Total Evidence", *British Journal for the Philosophy of Science*, 17

Hawthorne, James and Bovens, Luc., 1999. "The Preface, the Lottery and the Logic of Belief", *Mind*, Apr., Vol.108, pp. 430

Jeffrey, Richard C., 1970. "Dracula meets Wolfman: Acceptance VS. Partial belief", Swaine, M. ed., *Induction Acceptance and Rational Belief*, D. Reidel Publishing Company, Dordrecht Holland

Howson, C & Urbach, P, 1989. *Scientific Reasoning the Bayesian Approach*. Open Court Publishing Company

Kaplan, Mark, 1996. *Decision Theory as Philosophy*, MIT University Press

----- 2006. *Coming to terms with our human fallibility*, Epistemic Value Conference and Pre-Conference Workshop August 18th-20th, Stirling Management Centre

Kyburg, Henry, Jr. 1961. *Probability and the Logic of Rational Belief*. Wesleyan University Press

Millikan, Ruth, 1993. *White Queen Psychology*, MIT University Press

Ramsey, F. P., 1926. "Truth and Probability" *Philosophical Papers*, ed. D.H. Mellor. 1990. Cambridge University Press

----- F. P., 1991. *Notes of Philosophy, Probability and Mathematics* Ed. Galavotti, M. C. Bibliopolis - Edizioni di filosofia e scienze

Stanley, J., 2005. *Knowledge and Practical Interests*, Oxford University Press

Strawson, P. F., 1956. *Introduction to Logical Theory*, Methuen

Williamson, T., 2000. *Knowledge and its limits*, Oxford University Press

## **JUSTIFYING HUMANITARIAN INTERVENTION TO THE PEOPLE WHO PAY FOR IT**

NED DOBOS

UNIVERSITY OF MELBOURNE

### ***Abstract***

The practice of humanitarian intervention, which involves one state (or coalition) intervening militarily into another state in order to prevent abuses of human rights, raises a plethora of ethical and political issues. How is foreign intervention to be reconciled with state sovereignty? Is intervention a threat to international peace and stability? Are alien values being imposed on the target society? Each of these questions has been thoroughly explored by both philosophers and jurists. But the notion that a state infringes the rights of its own citizens by waging war to defend the human rights of foreigners has received relatively little attention. The only thorough philosophical exploration of this problem to date – Allen Buchanan’s “The Internal Legitimacy of Humanitarian Intervention” (1999) – is the focus of this paper. My aim here is not to offer a resolution to the “internal” problem, but simply to show that Buchanan misrepresents it in several important respects. First, Buchanan understates the strength of this objection: it is much more resilient than he gives it credit for. Secondly, he overstates its scope, or the range of humanitarian interventions to which it applies. And finally, Buchanan is mistaken to think that whether or not an intervening state fulfils the rights of its own citizens can be determined prior to, and therefore independently of, the question of whether its actions are consistent with the rights of the society targeted by the intervention.

### ***Introduction***

The practice of humanitarian intervention, which involves one state (or coalition) intervening militarily into another state in order to prevent abuses of human rights,

raises a plethora of ethical and political issues. How is foreign intervention to be reconciled with state sovereignty?<sup>1</sup> Is intervention a threat to international peace and stability? Are alien values being imposed on the target society? Each of these questions has been thoroughly explored by both philosophers and jurists. But the notion that a state infringes the rights of its own citizens by waging war to defend the human rights of foreigners has received relatively little attention. Since the concern here is with the citizens of the intervening state, rather than the rights of the target government or the intended beneficiaries of the action, we can call this the *internal* objection to humanitarian intervention. The only thorough philosophical exploration of this problem to date – Allen Buchanan’s “The Internal Legitimacy of Humanitarian Intervention” (1999) – is the focus of this paper. My aim here is not to offer a resolution to the internal problem, but simply to show that Buchanan misrepresents it in several important respects, providing an inaccurate account of 1) its strength; 2) its scope, and; 3) its priority.

For Buchanan, the internal objection issues from a particular understanding of the social contract that binds a government to its citizens, which he calls the “discretionary association” view. On this view, citizens empower their government to act as an agent or trustee *for the sole purpose of promoting their interests*. They submit to the authority of their state and relinquish a portion of their earnings in tax in return for this service. The contract says nothing about using public resources to vindicate moral ideals or to defend human rights in foreign countries (1999, p. 74). Thus a government that prosecutes a humanitarian intervention tramples on the rights of its own taxpayers in the process, by breaching the terms of its contract with them.

---

<sup>1</sup> Whether or not foreign intervention necessarily compromises state sovereignty depends partly on how one defines intervention. According to some definitions, intervention is something that occurs without the consent of the target state. According to others, a state can invite humanitarian intervention into its territory. On the latter definition state sovereignty can obviously be left in tact by foreign intervention.

Buchanan attempts to undercut this objection by dismantling the account of the social contract on which it relies. But I shall argue that the discretionary association view is more resilient than Buchanan gives it credit for. Moreover, Buchanan is mistaken to think that consistency with the terms of the social contract is sufficient for internal legitimacy. Obstacles to intervention arise also out of the state's distinct moral contract with its armed forces, which Buchanan does not consider. It is in this sense that he understates the strength of the internal objection.

But at the same time he seems to overstate its scope. Apparently, the discretionary association view makes internally legitimate humanitarian intervention "impossible" (Buchanan 1999, p. 73). If we accept this version of the social contract, every humanitarian intervention constitutes a breach of the citizens' trust. I will argue that this is true only if we adopt an outdated and implausible definition of "humanitarian intervention". Once this error is corrected, the internal objection will be seen to apply to a rather narrow range of all humanitarian wars – possibly to none at all.

Finally, Buchanan is mistaken to attribute logical and moral priority to the internal over the external dimension of legitimacy. The question of external legitimacy is simply the question of whether an intervention wrongs the inhabitants of the target country and/or their government. Buchanan seems to suggest that this question can be addressed only after the internal question has been settled. I shall argue that, on the contrary, we cannot determine whether an intervening state honours the rights of its own citizens without first determining whether its actions are externally just.

### ***Strength***

If the discretionary association view introduced above is an accurate account of the social contract, then it turns out that any state which prosecutes a humanitarian intervention infringes the contractual or fiduciary claims of its citizens. In order to

show that it is possible for a humanitarian intervention to be internally legitimate, then, it is necessary to rebut the discretionary association view. Buchanan's paper is devoted in large part to this task. Due to limited space, I will provide only a crude sketch of Buchanan's argument below.

Buchanan builds his case around the premise that we are each obliged, independently of our prior transactions and undertakings, to defend human rights when we are able to do so at no significant cost to ourselves. From this he infers that people in the world's most affluent countries will, at least sometimes, be obliged to make their collective resources available for the defence of human rights beyond their borders. Acknowledging this, however, seems to make the discretionary association view untenable. For if a community is morally obliged to finance an intervention, its members cannot possibly transfer all of their money to an agent, and then bind that agent to a contract which absolutely forbids her to use that money for the purposes of humanitarian intervention. We simply have no right to do this with our resources given our obligations, and so a contract to this effect must be invalid. Were it otherwise, individuals and groups would be able to shun their moral obligations simply by empowering an agent to act on their behalf (Pogge 2002, p. 126). The upshot of this is that a humanitarian intervention cannot breach any valid terms of the social contract, and thus cannot breach the citizens' trust, where these citizens are duty-bound to financially support the intervention.

Unfortunately, Buchanan's argument is unlikely to convince those advocates of the discretionary association view who are most in need of convincing: the political realists.<sup>2</sup> Buchanan acknowledges that a high-cost proviso comes attached to whatever duties we owe to oppressed foreigners. We are only obliged to finance humanitarian operations where we can to do so without sustaining too high a

---

<sup>2</sup> Due to limitations of space I will not provide a thorough exposition of political realism here, but will simply focus on certain aspects of it. For a detailed introduction to political realism, see Tim Dunne and Brian T. Schmidt (2001).

cost or taking too great a risk. Buchanan simply assumes, (understandably), that the citizens of at least some countries will, at least sometimes, be in a position to render assistance to foreign nationals without great risk or sacrifice. But the realist is notorious for his denial of this very premise.

Proponents of the realist theory of international relations are characterised by their fierce opposition to the intrusion of morality into foreign policy. This opposition is based on the belief that a government jeopardises the security of its people by pursuing moral goals and ideals in its dealings with other states. This belief, in turn, stems from a Hobbesian view of the international arena. Because there is no overarching world government capable of enforcing moral standards in international relations, it is reasonable to expect that states will relentlessly seek “power after power” to bolster their own security, says the realist. In this setting, any state that diverts resources away from this project and towards international philanthropy exposes its citizens to prey by allowing potential aggressors to acquire a relative advantage in terms of strategic positioning and resources that can then be translated into a military advantage. Thus every state must devote itself purely to self-aggrandizement in order to avoid making its citizens vulnerable to attack.

The following passage from prominent realist George Kennan illustrates this attitude well:

Whoever looks thoughtfully at the present situation of the United States in particular will have to agree that to assure these blessings to the American people is a task of such dimensions that the government attempting to meet it successfully will have very little, if any, energy and attention left to devote to other undertakings, including those suggested by the moral impulses... of its citizens. (1985-6, pp. 206-7)

Though Kennan does not appeal to the anarchical nature of the international system in this passage, the implication is clearly that a state cannot ensure the security of its people and still have time and money left over to defend human rights abroad. It follows that a state which does pursue moral ideals must compromise the security of its own citizens in the process.

If the realist is right about this, then the high-cost proviso admitted by Buchanan is perpetually in effect. Even if we sometimes have a *prima facie* duty to make our resources available for the defence of human rights abroad, this duty need not ever be discharged, since the anarchical nature of the international system means that acting on it always includes great risk.<sup>3</sup> Buchanan's first mistake is thus to assume that the discretionary association view cannot withstand the admission of a *prima facie* positive duty towards foreigners. The realist will simply grant the existence of this duty, but add that the high-cost proviso consistently exempts states from the requirement to fulfil it.

Buchanan's second mistake is to assume that the only internal constraints on altruistic war are those that arise out of the social contract, such that consistency with the terms of this contract is sufficient for internal legitimacy. Buchanan says nothing of the distinct moral contract between the state and its armed forces more specifically, to which any military operation must also conform. Martin L. Cook, Professor of Philosophy at the US Airforce Academy, sketches an account of this contract that closely mirrors the discretionary association view:

The military contract obliges military personnel to run grave risks and to engage in morally and personally difficult actions. They do these things on the basis of the implicit promise that the circumstances under which they must act are grounded in political leadership's good faith judgment that the

---

<sup>3</sup> By "anarchic" I simply mean absent an overarching power capable of enforcing rules of peaceful cooperation.

defense of sovereignty and integrity of the nation (or, by careful extension, the nation's vital interests) require their action (2000, p. 62).

The crucial point here is that soldiers implicitly agree to fight and die for their country and its interests, *not* for the human rights of foreigners. If this is correct, then a state's fiduciary obligations are transgressed by humanitarian intervention even when its taxpaying citizens are obliged to foot the bill. And it is important to note that Buchanan's argument against the discretionary association view cannot simply be reformulated and reapplied here. "Distributive justice may require your money, but it cannot require your life" (Luban 2002, p. 94). Thus we cannot say that, just as taxpayers are naturally obliged to finance humanitarian wars under certain circumstances, soldiers can be naturally obliged to fight in them.

Having said this, the obstacle presented by Cook's account of the soldier-state contract can I think be overcome. Humanitarian intervention has become one of the primary functions that militaries serve, at least in the West. Some of the most recent engagements of the Australian Defence Force, for example, include Operation Astute, aimed at restoring peace in East Timor, Operation Anode, geared towards a similar end in the Solomon Islands, peace-keeping operations in Sudan, and of course the "rehabilitation and reconstruction" of Iraq.<sup>4</sup> One cannot join the ranks of some such force, in full awareness of its history and policy of humanitarian intervention, and then protest that this was not what he or she signed up for.

And even if we accept the position outlined by Cook, it does not seem to yield a blanket prohibition of all humanitarian wars. Consider the following two means of carrying out an intervention: 1) By "riskless" or "post-heroic" means, characterised by reliance on long-range or high-altitude bombing and the absence

---

<sup>4</sup> To be sure, some of these do not technically count as humanitarian interventions given the prevailing definition, according to which an intervention necessarily takes place against the will of the target state. (Australian forces have been invited by the governments of Timor-Leste and the Solomons).

of ground forces, such that military casualties are avoided entirely. 2) Using a special expeditionary force comprised of volunteers. If either of these conditions were met, humanitarian intervention could be carried out consistently with the intervening state's fiduciary obligations toward its armed forces. Troops deployed on a riskless humanitarian mission could still remonstrate that "this was not part of the deal", but in this case their dissent would have lost much of its force. What is problematic is "ask[ing] someone who volunteered to die for America to die for Father Aristide", not asking a soldier to drop bombs from a position of perfect safety (Kaus 1994, p. 61).

Thus I do not mean to suggest that the internal objection is insurmountable, only that Buchanan understates its strength. The cogency of his case against the discretionary association view rests on a premise that is likely to be rejected by the most stubborn advocates of this view: namely that states can devote substantial resources to the promotion of human rights abroad without putting their own citizens at risk. Apart from this, Buchanan is mistaken to think that by rebutting the discretionary association view, the internal objection is done away with.

### *Scope*

Although Buchanan understates the strength of the internal objection, at times he seems to overstate its scope. The discretionary association view is said to make "internally legitimate humanitarian intervention impossible", as if to suggest that all such wars are liable to the charge of constituting a breach of the social contract. But this is simply not the case.

Let us grant for the sake of argument that the discretionary association view, and the corresponding account of the soldier-state contract, are correct. It follows that a government has the authority to wage war only in pursuit of the national interest – that is, only with the intention of advancing the security and wellbeing of its own

citizens. What this means is that humanitarian intervention is consistent with the state's fiduciary obligations only on the condition that it is motivated by national self-interest. But this, some would say, is logically impossible. Humanitarian intervention is by definition driven by humanitarian sentiments. Any intervention that is motivated by national self-interest is therefore disqualified.<sup>5</sup>

This is very much in line with the way that humanitarian intervention has traditionally been defined, but the view that mixed motives do not necessarily disqualify a military action from being classified as humanitarian is gaining wider acceptance. For Nicholas Wheeler, as long as the political and economic motives do not obstruct the achievement of the humanitarian objective, their presence does not affect the character of the action (2000, pp. 37-39). Tony Coady agrees: "It may indeed be too much to require that the humanitarian reasons should constitute sufficient conditions for the intervention, but they should at least be necessary and prominent" (2002, p. 11). Surely this *is* too much to ask. The traditional definition has the counter-intuitive implication that Vietnam's intervention into Cambodia, which put a stop to Pol Pot's savagery, was not a humanitarian intervention since one of the main reasons it was waged was to prevent massive numbers of Cambodian refugees from spilling over onto Vietnamese territory.

Some would even go so far as to say that humanitarian motives need not be in the mix at all. Teson emphasises a distinction that is all too often overlooked in the literature on humanitarian intervention: that between *intentions* and *motives*. An intention covers the contemplated act; "what the agent wills to do". By contrast a motive is a further goal that one wishes to accomplish with the intended act (Teson 2005, p. 5). Suppose that I rescue a drowning kitten solely because I expect to receive a reward from its owner. To be rewarded is my *motive*, but my *intention* (the act contemplated and

---

<sup>5</sup> Bhikhu Parekh, for instance, argues that intervention aimed at protecting the national interest is "an act of self-defence and not a case of humanitarian intervention". The latter by definition involves the state acting "in a more or less disinterested manner". (Parekh 1997, p. 57).

carried out as a means to my end) is to rescue the cat. The important thing to notice is that the character of the action is determined by the character of the intention, not that of the motive, (it is, after all, still an act of *rescue*). Teson extrapolates that a government that deposes a tyrant ultimately for self-interested reasons, say to secure continued access to his country's natural resources, is still engaged in humanitarian intervention as long as there is an intention to unseat the tyrant and with this to liberate his people as a means to this end (1988, p. 115).

Now where humanitarian intervention *is* motivated by national self-interest, it seems to be perfectly consistent with the terms of the social contract as delimited by the discretionary association view, and with the corresponding soldier-state contract. If the state is empowered for the sole purpose of obtaining optimal conditions for its citizens - ensuring their security and freedom, improving their standard of living - then if a course of action adopted by a government is geared towards this end, it falls squarely within the perimeters of the state's mandate. But then how can the fact that it is *also* expected to benefit foreigners, as a means to, or a side effect of this, possibly make a difference? Surely this additional effect cannot mean that the action is no longer one that the state is authorised to pursue. Likewise a soldier cannot sensibly protest that he did not enlist to fight for the national interest in cases where this is to be promoted via the provision of humanitarian assistance, or where foreigners are also to be advantaged.

One might of course retort that even where intervention is expected to serve the national interest, there will often be some alternative course of action that would prove even more domestically advantageous, and this may very well be true. But when we empower a statesperson, we authorise him to use the state's resources in whichever way best serves the national interest *in his judgment* (Pennock 1979, p. 316). The state is not merely a delegate whose role is to transmit the direct instructions of its principals or authors. It is what Joel Feinberg refers to as a "free

agent”, defined as “an expert hired to exercise his professional judgment on behalf of, and in the name of, his principal” (1968, p. 675).

And it is important to emphasise that the specifics of *how* the intervention will promote the national interest are immaterial in this connection. Neo-conservative political realist Charles Krauthammer proposes that if and only if humanitarian intervention is “strategically necessary” does the intervening government act within the limits of the authority vested in it by its people. “To intervene solely for reasons of democratic morality is to confuse foreign policy with philanthropy”, writes Krauthammer. “And a philanthropist gives out his own money. A statesman is a trustee. He spends the blood and treasure of others.” But “strategic necessity” is an ambiguous term needing delimitation. Krauthammer offers no clear definition, but gives us some examples of what he has in mind. If an intervention will “defend the perimeter or sea lanes or chromium mines” of a state, it passes the test. On the other hand if the decision to intervene is a response to “international law, world public opinion, and the public sentiments of our (America’s) allies”, it does not (1985). These are not encompassed by the term “strategic necessity”.

Furthermore, Krauthammer appears to suggest that there needs to be a threat to the national interest that intervention would avert for this criterion to be met. “We should commit ourselves only in those places where our interests are *threatened*” he writes (1985). And again:

We will support democracy everywhere, but will commit blood and treasure only in places where there is a strategic necessity – meaning, places essential to the larger war against the existential enemy, the enemy that poses a global mortal threat to freedom (2004).

With this, Krauthammer effectively rules out intervention that would

*positively improve* conditions in the country doing the intervening. For example, intervention that would secure future access to a valuable market or resource from which a state already profits is covered, but intervention that would tap in to *new* resources and markets is not. It is unclear whether Krauthammer sees this implication of his remarks, and whether he would be willing to accept it.

If Krauthammer is simply concerned about intervention conforming to the terms of the state's domestic contracts, I do not see any reason or justification for these restrictions. Why must the national interest be served by the containment of a threat? And why can it not be advanced by improving the state's relations with - and winning it the respect of - its friends, allies, and even its enemies? After all, developing a positive international reputation is advantageous in terms of national security and also tends to facilitate a country's economic flourishing. Furthermore, as Joseph Nye observes, "if a state can make its power legitimate in the eyes of others, it will encounter less resistance to its wishes" (1990, p. 182). Responsiveness to international law and world public opinion would therefore seem to be recommended by an enlightened, far-sighted national self-interest.<sup>6</sup>

Humanitarian intervention thus need not be "strategically necessary" in Krauthammer's sense. As long as it is expected to advance the national interest, it conforms to the terms of the social contract *even on the discretionary association view*. The internal objection cannot be brought to bear against every humanitarian intervention, then, but only against those that are *disinterested* or unadulterated by national self-interest. At times Buchanan does seem to realise this, referring specifically to "pure" humanitarian intervention, rather than humanitarian intervention *per se*.<sup>7</sup>

---

<sup>6</sup> Krauthammer might of course retort that any benefit a state gets from improving its reputation is likely to be outweighed by the disadvantages associated with prosecuting a humanitarian intervention that is not strategically necessary. Again, however, it is the role of the state *qua* free agent to determine how best to promote the national interest.

<sup>7</sup> Buchanan 1999, p. 78

But if the internal objection applies only to “pure” humanitarian interventions, then there is an acute possibility that it has no purchase on reality. Two varieties of political realism can be distinguished: normative and descriptive. The former, on which we have focussed so far, says that states *ought* to look only to the national interest in their dealings with other states. According to the descriptive realist, states *cannot*, by their very nature, act on motives other than the national interest. The pursuit of security, wealth and power underlies all political action, despite appearances. On this view, a state waging war without sincerely believing that this is the best way to promote the national interest in the circumstances is about as likely as one of the scientific laws of nature failing to obtain.

If the descriptive realist is right, then the internal objection advanced by the normative realist is purely academic. Since all wars are motivated by national interest, we can rest assured that the terms of the social contract will be honoured. And even if the descriptive realist is excessively cynical, we can still be confident that the range of cases that are liable to the internal objection is narrowly circumscribed, given that a large proportion of interventions can be reasonably presumed to be motivated by the national interest. Far from rendering internally legitimate humanitarian intervention “impossible”, then, the discretionary association view at most rules out a subclass (and a potentially very small subclass) of all such interventions.

### ***Priority***

The problem of internal justification “remains after all the familiar questions [about humanitarian intervention] are answered satisfactorily” says Buchanan. Yet it “precedes” all of these familiar questions “because unless it can be answered the other questions do not arise”. (1999, p. 72). What Buchanan seems to be saying here is that just because an intervention does not wrong the target society, its government, or the international community more generally (or just because it is *externally* legitimate, which is what most of the familiar questions are concerned with) – it

does not follow that it is internally legitimate. This is quite right. But the notion that internal legitimacy can be determined *prior* to, and therefore independently of external legitimacy, is wrongheaded. The case could be made that an externally unjust war is *ipso facto* internally unjust. Therefore internal legitimacy cannot be addressed before the question of external legitimacy has been settled.

We have seen that where citizens are morally obliged to finance an intervention, their government cannot be said to wrong them by using public resources to prosecute that intervention. Where citizens are *not* under any such obligation, however, they are forced to make a sacrifice that they are within their rights to refuse, and this does give them grounds for complaint. Now obviously we cannot be morally obliged to facilitate our government's immoral actions. A state that prosecutes an externally unjust intervention therefore necessarily compels its taxpayers to make a sacrifice that is not required of them.

But this is just one of the harms that a wrongfully intervening government could be said to inflict on its constituents. The victims of unjust war have a claim to reparations against the party responsible for that war. A state could be said to wrong its citizens by putting them at risk of one day having to compensate for its military misadventures.

On top of all this, the case could be made that a government damages its citizens *morally* by invading a foreign country without justification. By electing a government, individuals empower an agent to act in their name and on their behalf. What bearing does this have on their moral status in the event that this government wages an (externally) unjust war? Joel Feinberg tells us that an agent's acts "are binding on [his] principal as if he had done them himself" (1968: 675). Taking this at face value would seem to imply that where an agent acts in a harmful manner within the scope of his specified duties, his guilt simply transfers to his author as if the author had himself engaged in the harmful conduct. But this will not do. Moral guilt is not

the sort of thing that admits of being transferred from one party to another, since one of its necessary components is a guilty mind or “mens rea”.

But there are other ways for citizens to be morally compromised by the wrongdoing of their government. A number of contemporary philosophers subscribe to the notion of “moral taint”: the idea that an individual can be blemished by immoral actions for which he is not responsible, simply in virtue of his relationship to the guilty party. Writes Gregory Mellema:

The wrongful acts of one’s spouse, sibling, or co-worker sometimes appear to strike at our moral integrity in a way that goes far beyond embarrassment or damage to our reputation. We feel a type of moral involvement in these acts that is hard to explain with the categories and concepts of traditional moral philosophy (1997, p. 80).

In addition to this, an elected government that goes to war unjustly could be accused of putting its people in a position where they are liable to *become* culpable or morally guilty, thus exposing them to even more serious moral damage. Consider a bank manager that hires a teller to deposit customers’ cash into their accounts. Shortly after being employed the teller begins to steal from clients, putting part of every deposit into his own pocket. Despite being aware of what his employee is up to, the manager does not fire the teller or take steps to prevent any further theft, say due to laziness or indifference. The manager is certainly culpable for his employee’s crime, not simply because he was in a position to stop it (as a regular customer might also have been), but because he had independently assumed a duty to do so by employing the guilty party.

One might say that democratically active citizens are in an analogous position: they assume a responsibility for the actions of their government by electing it. If an

elected government is using its power to violate the rights of some third party (the society targeted by an unjust intervention), then the authors of this government have a duty to revoke its authority or to at least take steps to prevent further wrongdoing. Should they fail to do so, those members of the collective who are responsible for the failure – most obviously those who vote to re-elect the warring government - become complicit in their government's crimes. By exposing its citizens to moral damage in this way, an unjustly intervening government could plausibly be said to harm them.

We have, then, a number of reasons to think that internal legitimacy cannot be addressed prior to external legitimacy. The former seems to presuppose the latter, or to have it as a necessary condition. What this means is that the internal dimension of legitimacy does not have the logical and moral priority that Buchanan attributes to it.

## **Conclusion**

By turning the spotlight away from the intended beneficiaries of humanitarian intervention and towards the men and women whose sacrifices make it possible, Buchanan has disinterred an issue that has remained well and truly buried in the scholarly debate around altruistic war. For this he is to be commended. My aim here has simply been to offer a more accurate depiction of the problem. I have suggested that the range of cases in which the internal objection arises to begin with is much narrower than Buchanan lets on. But at the same time, where it does arise, it seems to withstand Buchanan's rebuttal. It is not enough to show that we are obliged to make our resources available for the defence of human rights when we can do so without taking too great a risk. To convince those in need of convincing, it is also necessary to show that such circumstances might actually arise in the real world. Furthermore, the fact that an intervention is consistent with the fiduciary rights of the taxpayer does not guarantee that it leaves the fiduciary rights of the soldier intact. By saying nothing of this, Buchanan reduces military personnel to

just another kind of public resource whose expenditure needs to be justified only to those who pay their wage. This is wrongheaded. Finally, Buchanan is mistaken in attributing precedence to the internal over the external dimension of legitimacy. The internal status of a humanitarian intervention cannot be determined before its external status is well and truly settled.

### **References**

Buchanan, A. 1999, 'The internal legitimacy of humanitarian intervention' *The Journal of Political Philosophy*, vol. 7, no. 1, pp. 71-87.

Coady, C.A.J. 2002, *The Ethics of Humanitarian Intervention*, Peaceworks 45, Washington: United States Institute of Peace.

Cook, M.L. 2000, 'Immaculate war: Constraints on humanitarian intervention', *Ethics and International Affairs*, vol. 14, pp. 55-65.

Draper, K. 1993, 'Fairness and Self-Defense', *Social Theory and Practice*, vol. 19, No. 1.

Dunne T. and Schmidt B.T. 2001, 'Realism', in John Baylis and Steve Smith (eds.), *The Globalization of World Politics*, 2<sup>nd</sup> edition, New York: Oxford University Press.

Feinberg, J. 1968, 'Collective responsibility', *Journal of Philosophy*, vol. 65, pp. 674-88.

Kaus, M. 1994, 'TRB from Washington: Somalia syndrome', *The New Republic*, vol. 211, no. 16, pp. 4/61.

Kennan, G. 1985-6, "Morality and Foreign Policy", *Foreign Affairs*, vol. 64, pp. 205-218.

Krauthammer, C. 1985, 'When to intervene: What's worth fighting for?', *The New Republic*, vol. 192, pp. 10-13.

----- 2004, 'Democratic realism: an American foreign policy for a unipolar world', Irving Kristol Lecture, American Enterprise Institute for Public Policy Research. [http://www.aei.org/publications/pubID.19912,filter.all/pub\\_detail.asp,\(2/9/07\)](http://www.aei.org/publications/pubID.19912,filter.all/pub_detail.asp,(2/9/07))

Luban, D. 2002, 'Intervention and civilization: Some unhappy lessons of the Kosovo war' in P. de Greiff and C. Cronin, (eds), *Global justice and transnational politics: essays on the moral and political challenges of globalization*, Cambridge Mass: MIT Press.

Mellema, G. 1997, *Collective responsibility*, Amsterdam; Atlanta: Rodopi.

Montague, P. 1981, 'Self-defense and choosing between lives', *Philosophical Studies*, 40.

Nye, J. 1990, 'The changing nature of world power', *Political Science Quarterly*, vol. 105, no. 2, pp. 172-92.

Parekh, B. 1997, 'Rethinking humanitarian intervention', *International Political Science Review*, vol. 18, no. 1, pp. 49-69.

Pennock, J.R. 1979, *Democratic political theory*, Princeton: Princeton University Press.

Pogge, T.W. 2002, *World poverty and human rights: cosmopolitan responsibility and reforms*, Cambridge: Polity Press.

Teson, F.R. 1988, *Humanitarian intervention: An inquiry into law and morality*, Dobbs Ferry, N.Y.: Transnational Publishers.

Teson, F.R. 2005, 'Ending tyranny in Iraq', *Ethics and International Affairs*, vol. 19, no. 2, pp. 1-20.

Wheeler, N. 2000, *Saving Strangers: Humanitarian Intervention in International Society*, New York: Oxford University Press.

----- 2003, 'Humanitarian intervention after September 11, 2001', in A. Lang (ed), *Just intervention*, Washington DC: Georgetown University Press.

## METAETHICAL MORAL RELATIVISM AND THE ANALOGY WITH PHYSICS

ALEXANDRE ERLER

LINCOLN COLLEGE, OXFORD

### *Abstract*

This paper deals with a specific version of metaethical moral relativism, known as “speaker-relativism”. It starts by explaining the position, focussing on the views of two prominent contemporary relativists, Gilbert Harman and James Dreier. Both authors draw an analogy between ethics and modern physics: just as Einstein showed that judgments about time or mass were always relative to a specific frame of reference, Dreier and Harman argue that “absolutist” judgments about moral rightness or wrongness need to be reinterpreted as relative to some particular moral system. They also claim that this analogy allows us to salvage ordinary moral talk. I consider a number of possible objections to their argument, beginning with one, concerning the possibility of moral disagreement, which I think can be successfully answered, and then presenting two criticisms that I take to be more problematic for the relativist. I argue that despite its initial appeal, Harman and Dreier’s suggestion regarding our use of moral language seems to be a source of confusion in certain cases of moral disagreement, and does not appear able to preserve specifically *moral* normativity – which leads me to conclude that it is best viewed as a variant of an error theory about morality, rather than as the distinct metaethical position it purports to be.

### *Metaethical Moral Relativism*

Briefly put, metaethical moral relativism (henceforth MMR) is the view that the truth or falsity of moral judgments, or their justification, is always *relative* to some

particular moral code. (In what follows I shall focus on the most challenging view, the one that takes truth and falsity to be relative.) The view comes in several different versions. The moral code that is relevant for the truth or falsity of the judgment in question can be said to be that of a *group* of people, for instance the community to which the person making the judgment belongs; or the moral standards of one particular individual (e.g. the speaker herself) can be singled out as relevant. On the other hand, assuming that it is the moral outlook of one particular person that is relevant, a distinction should also be drawn between theories that make moral judgments relative to the *speaker's* outlook, and those that tie them to the convictions of the *agent* whose action is being judged. In the former case, we get what is often called “appraiser-relativism” or “speaker-relativism”, whereas in the latter case we get “agent-relativism”.<sup>1</sup>

No matter what particular form of the theory they endorse, metaethical moral relativists (henceforward, relativists) agree with error theorists like John Mackie that we have no good reasons for believing in the existence of “absolute” moral values. (See Mackie, 1977.) However, contrary to the latter, they do not conclude that in the absence of such values, all the substantive moral evaluations we make are, as far as we can tell, purely and simply false.<sup>2</sup> As my rough statement of MMR shows, relativists believe that there is such a thing as moral truth – they just claim that it is relative. Now what, more precisely, do they mean by that? In what follows, I will focus on two prominent modern-day relativists: Gilbert Harman and James Dreier. Harman and Dreier subscribe to slightly different forms of MMR. Dreier explicitly describes himself as advocating speaker-relativism, in a form that has

---

<sup>1</sup> A similar distinction can of course also be drawn for theories that make moral judgments relative to the convictions of a *group*: we would then get “speaker’s group relativism” and “agent’s group relativism”, respectively.

<sup>2</sup> “Substantive” moral evaluations, as I understand them, typically refer to judgments of the form: “X is right/wrong”, as opposed to “It is not the case that X is right/ wrong”; or “S ought/ ought not to...”, as opposed to “It is not the case that S ought/ ought not to...”. Yet as we shall see, relativists seem to want to make the class of substantive moral judgments more inclusive.

affinities with noncognitivism. He proposes to understand moral terms on the model of indexicals, like “here” or “now”. According to this proposal, “ ‘x is good’ means ‘x is highly evaluated by standards of system M,’ where M is filled in by looking at the affective or motivational states of the speaker and constructing from them a practical system” (Dreier, 1990, 9). Since the motivational states in question are not cognitive states, Dreier’s view incorporates a noncognitivist element without nevertheless qualifying as a form of noncognitivism (given that for him moral statements do possess truth-value). Harman’s position, on the other hand, may be closer to speaker’s group relativism, since he thinks that moral judgments (at least those that are adequately phrased) involve an implicit reference to moral standards that are shared by the speaker and his audience. However, his position still implies that the *speaker’s* moral standards are relevant to the truth of or falsity of his statements, and he also acknowledges the possibility that the group with which the speaker stands in implicit agreement might include only himself (see Harman, 2000, 8-19; 36). In what follows, I will thus focus on speaker-relativism, and treat the analogy which both Harman and Dreier appeal to as a defense of that particular version of MMR (I shall henceforth use “MMR” to refer to speaker-relativism).

I have said that relativists believe they are able, contrary to error theorists, to salvage talk of moral truth, even when substantive moral judgments are concerned. Here it might be asked what exactly they mean by that. As we have just seen, Harman and Dreier think that an adequately phrased moral judgment will involve a reference to the speaker’s set of moral standards. Let us consider a judgment which seems to fit that description: “according to my own moral convictions, abortion is morally wrong”. If I really do happen to endorse a set of moral standards that prohibits abortion, then, according to the relativist, my judgment will be correct. Yet such a claim cannot allow the relativist to distinguish himself from the error theorist. The latter will readily agree that, under the conditions just stipulated, my judgment is correct – but this is because the judgment in question does not constitute a genuine

moral judgment. It is more adequately described as a “sociological” judgment concerning a particular person’s moral standards and its implications, which any observer, whatever his moral convictions, might appropriately make. Therefore, if the relativist wants to salvage the notion of truth as applied to substantive moral evaluations, he cannot content himself with pointing out that judgments of the sort just mentioned can, when the relevant conditions are met, be true.

However, this is not all that authors like Harman and Dreier have to say. First, they would presumably insist that the kind of judgments I have described as “sociological” are, strictly speaking, the best we can get, given antirealism about values, and thus that we should agree to treat them as genuine, substantive moral judgments.<sup>3</sup> To make this point more plausible, they might stress the fact that since every true moral judgment worth describing as “substantive” by their standards will be made in conformity to a moral code that the speaker himself accepts, every such judgment will carry normative weight for him, i.e. it will give him a reason to behave as that moral code prescribes. Secondly (and this may seem paradoxical given what I have just said, but the paradox will be resolved in subsequent sections of this paper), they think that MMR allows us to rescue the notion of moral truth even for “absolutist” judgments of the kind “abortion is wrong” – which everyone will acknowledge to be a substantive moral evaluation. How can they make such a claim? This is where the analogy between ethics and contemporary physics comes in.

### *The Analogy Between Ethics And Physics*

Harman and Dreier both draw an analogy between MMR and the theory of Relativity in physics. Here is one way of presenting it. Before Einstein put forward his Special Theory of Relativity, it seemed obvious that the passing of time was a

---

<sup>3</sup> Dreier, for instance, talks about “toning down” the aspirations of moral judgments. See Dreier, 2006, 261, and also Harman, 2000, 38.

perfectly linear and homogeneous process. Whether you were at rest or traveling very fast in a futuristic vehicle wasn't supposed to affect in any way the pace at which time elapsed. Yet Einstein's theory precisely demonstrated that common sense was mistaken in this respect. The answer to the question of how many years have passed since George W. Bush first became president of the USA will be different for us, who spent the whole period busy with our earthly affairs, and for an astronaut who had been traveling all the while and at very high speed in his spaceship. In fact, fewer years will have elapsed for the astronaut than for us, since as Einstein has taught us, time passes more or less quickly depending on how fast one is traveling. Thus there is no 'absolute' answer to the question. Any answer will have to be relative to a particular frame of reference – relative to us, seven years have elapsed since Bush was first elected, whereas only three or four years might have elapsed for the astronaut in his spaceship. Relativists like Harman and Dreier propose to understand moral judgments on a similar model. Even though, as we have seen, they agree with error theorists that we have no good reasons for believing in the existence of objective values that might validate "absolutist" moral judgments, they maintain that such judgments can still be true if they are understood as being *relative* to a particular moral code, in the same way as judgments about time, in order to be correct, need to be made relative to a particular frame of reference. Thus for the relativist, it is no more necessary to declare that absolutist moral judgments are all false than it is to say, if someone claims that seven years have elapsed since Bush first became president, that that person is wrong if he fails to specify a frame of reference to which his judgment might be tied. Indeed, it is perfectly legitimate to regard the relativizing clause "for us on the surface of the earth" as implicit in that judgment. Similarly, says the relativist, many moral judgments framed in absolute terms can legitimately be regarded as true if they are reinterpreted as relative to the relevant moral outlook: for instance, "abortion is wrong" will count as true if its relativistic translation,

“abortion is wrong relative to the set of moral standards that I accept” is true (see Harman, 2000, 37-38, and Dreier, 2006, 261-62).

### ***Problems With The Analogy***

*What about the possibility of moral disagreement?*

Though the analogy drawn by Dreier and Harman might seem appealing, it raises a number of difficulties. I will begin with a possible objection which I think can be successfully answered, before considering others that I take to be more problematic. Harman and Dreier propose to reinterpret absolutist moral statements by making them implicitly relative to the moral standards of the speaker. Now doesn't their view have the undesirable implication that moral disagreements actually do not occur, despite our very strong intuition that they do?

I think that the relativist should not have too much difficulty replying to that objection. He can agree that when we do mean to make moral judgments in an “absolute” manner, genuine moral disagreements are possible. However, this need not worry him, for the suggestion made by Harman and Dreier is a *reformative* one; they do not pretend to be merely describing the way we ordinarily use moral language. It is only once moral statements have been suitably relativized that people with different moral outlooks won't be disagreeing with each other anymore.<sup>4</sup> And neither need this be a problem for the relativist: he can remark that the same would happen if we and the astronaut, in the example given above, initially made our claims about the amount of time that had elapsed since G.W. Bush was first elected in absolute terms (thus disagreeing with one another), and then came to realize that Einstein's Special Theory of Relativity was true. We

---

<sup>4</sup> Though two people sharing the *same* moral outlook might still disagree as to the exact implications of the set of moral demands they both accept.

would then suitably modify our claims and the disagreement would disappear – but the fact that it leads to the removal of such disagreements is certainly not something we can hold against Einstein’s theory. And so it also goes, the relativist might say, for MMR. True, the impossibility of coherent disagreement between speakers with different moral outlooks is something we might find more difficult to accept than the kind of relativism about time implied by Einstein’s Special Theory of Relativity, given the importance of the pretension to absolute truth in moral talk,<sup>5</sup> but the relativist might still insist that we ought to accept such a consequence, for the only alternative – sticking to absolute moral statements and rejecting the relativist’s reform – will lead us to conclude with error theorists that all substantive moral judgments are false, a conclusion the relativist will say we should avoid.

### *The loss of moral normativity*

Paul Boghossian has raised what I think is a more serious objection against MMR. He himself observes that – as I have mentioned before – the relativistic translations that authors like Harman and Dreier propose for ordinary moral statements seem to turn them into what I have called mere “sociological judgments”. For Dreier, as we have seen, “X is good” means “X is highly evaluated by the standards of (moral) system M”. Harman holds quite a similar position. He thinks that morality depends on an agreement in intentions between different agents, each of them intending to act in keeping with a particular set of principles, provided that others intend to do so too. Harman’s translation of an absolutist moral statement like “it would be wrong of me to lie” would thus be something like “it would be wrong of me to lie given my intention (which you share) to act in keeping with set of principles M, which prohibits lying”. However, when taken by themselves, Harman and Dreier’s

---

<sup>5</sup> Assuming of course, as relativists do, that ordinary moral statements of the sort ‘it was wrong of him to break his word’ are to be given a cognitivist analysis.

relativistic statements do not seem to carry any normative weight whatever. Any non-relativist might accept them, even people who reject M.

True, since we are dealing here with speaker-relativism, the moral code that will be relevant to the relativistic translation of an absolutist moral statement will always be one the speaker himself accepts, and as I have said earlier, the relativist might insist that the speaker's adherence to M is sufficient to yield normative force, at least for *him*, to the judgment that X is good, or highly evaluated, by M's standards – in other words, it gives him a reason to treat X as good. Yet even if we concede that point, another worry awaits the relativist: is the normativity thus bestowed on the relativistic translation really the specific kind of normativity required if it is to count as a genuine moral statement, namely *moral* normativity? This leads us to Boghossian's challenge to MMR. The moral evaluations made by the supporter of MMR are supposed to be relative to a moral code that he accepts. But how, Boghossian asks, could the relativist still endorse a particular moral code, since he is committed, just like the error theorist, to the claim that there are no absolute moral facts? This seems to imply that no moral code can be correct, and how could the advocate of MMR go on subscribing to some moral code while believing the propositions constituting that code to be all false (see Boghossian, 2006, 25-27)?<sup>6</sup>

If we assume that only a belief in the existence of objective moral values might provide adequate justification for accepting a given moral code, then Boghossian's objection would seem to deal a fatal blow to MMR. However, Dreier and Harman would not agree with such a claim. As we have seen, Dreier identifies the particular "system"

---

<sup>6</sup> I shall ignore here the alternative claim that the moral principles that enter into particular moral codes are not false, but rather involve *incomplete* propositions, a bit like the proposition "Tom is taller than...". Indeed, I take such a suggestion to be rather implausible. The judgement that slavery is wrong, for instance, seems to us complete as it stands, and we seem to understand perfectly well what it means. The claim that we are actually mistaken in our impression, and that all the people who have been making such judgments so far were only under the delusion of making full-fledged moral claims, would seem, in the absence of very strong arguments, rather difficult to accept.

to which moral judgments are supposed to be relative with a set of *motivational attitudes*, i.e. *noncognitive* states. Harman's intentions to keep an agreement are also noncognitive, motivational states. Now doesn't that remove the worry raised by Boghossian, since we certainly do not need to believe in the existence of absolute moral facts in order to have such motivational attitudes?

I am not convinced that such an answer can suffice. Moral systems typically involve one or more principles of the sort: "lying is wrong" (though these principles need not have such a simple and general form; as Hare has suggested, they may also be of extreme complexity).<sup>7</sup> Now the motivational attitudes Dreier refers to presumably involve being motivated to act in keeping with such moral principles. But then we will be entitled to ask whether the propositions in which those principles find expression are true or false. Clearly, they cannot be true absolutely speaking, or else MMR would be false. And it does not seem they could themselves be true in a relative manner, for if we said they were true in relation to still other principles, the same question would arise at this more fundamental level. Of course, we might say that the propositions in question are true in relation to themselves (after all, any proposition logically follows from itself), but this would be perfectly trivial, and would not give us any reason to accept a principle, or set of principles, we did not already subscribe to. It thus seems we ought to assume that all such propositions are false, and that Boghossian's challenge to MMR holds good.

I suppose Dreier might reply that what matters for the truth of a specific moral statement is not whether it follows from some true moral principle, but whether it is consistent with the speaker's set of motivational attitudes. For example, the speaker's judgment that some particular act of lying is wrong will be true if this person is reliably motivated to act in keeping with the general principle according to which lying is wrong, even if that principle does not state anything

---

<sup>7</sup> See Hare, 1981.

like an objective moral truth. Yet such a reply could not alleviate our initial worry. While the fact that I am motivated to act in keeping with a certain moral code arguably does endow the corresponding specific statements (e.g. “it was wrong of you to lie to him”) with some sort of normativity, it does not appear to give us *moral* normativity, as opposed to the broader kind of normativity that one might attribute to motivational states in general insofar as they provide reasons (broadly understood) for action – what Bernard Williams called “internal” reasons (see Williams, 1981, 101-13). Perhaps Dreier might argue that since the affective attitudes in question involve a motivation to act according to what are unambiguously *moral* principles, the normativity they carry is indeed moral normativity. But he would then seem to be putting forward a *new* way of understanding moral normativity, rather than describing our ordinary understanding of that notion, which appears to involve, as Richard Joyce has argued, a *categorical* dimension: moral reasons and imperatives are supposed to apply to us regardless of the conative attitudes we contingently happen to have (see for instance Joyce, 2001, 157). Now one cannot show that relativistic moral statements deserve to be regarded as substantive moral claims by arguing that such statements do preserve moral normativity, if the notion of moral normativity one appeals to is not the standard one, but a new proposal.

Another possible option for Dreier would be to adopt a noncognitivist analysis of moral principles based on a minimalist theory of truth, holding for instance that to say that the principle “lying is wrong” is true, is simply to express one’s disapproval of lying. He would then not be committed to an error theory about moral principles. But we might then ask why Dreier, if he chose to take that route, does not extend his noncognitivist analysis to specific “absolutist” moral statements like “it was wrong of you to lie to him” – which of course would amount to abandoning MMR in favour of a full-fledged noncognitivism.

Harman, as we have seen, grounds morality in an agreement in intentions between different people regarding the ways they are to behave towards one another. He explains why people have such intentions by reference to the notion of “moral bargaining”: intending to observe a certain “agreement” about our ways of treating each other, provided that others similarly intend, brings mutual benefit (think for example about the agreement not to do harm to each other). That doesn’t mean, however, that Harman’s theory reduces moral motivation to self-interest. It may be that the most effective way of promoting the interests of people who intend to observe a particular agreement is precisely for them to develop, through habituation, genuine non-self-interested concerns for others (cf. Harman, 1977, 148-50). Yet in spite of this, Harman’s position seems in no better shape than Dreier’s when it comes to preserving moral normativity. Even if, not shying away from sophistication, we were to translate the statement “it would be wrong of me to lie” as “it would be wrong of me to lie given my intention to act in keeping with set of principles M, which prohibits lying, an intention wholly founded in a genuine concern for truthfulness”, it seems that we would lose the specific sort of normativity (moral normativity) that was present in the initial statement. This suggests that we cannot retain that kind of normativity while at the same time avoiding absolutist moral talk. And if we cannot retain it for relativistic moral statements, it then seems that such statements have no claim to be considered “substantive” moral judgments.

Here, two possible options for the relativist *à la Harman* (though not necessarily desirable ones) would be, first, to endorse ethical egoism, claiming that the intention to adhere to a particular “agreement” because this furthers one’s own interests is sufficient to count as a moral reason, but this is not the stance taken by Harman. Moreover, ethical egoism would seem to make moral relativism dispensable: if it goes against my interests to lie under circumstances C, then it is simply (morally) wrong for me to lie in such a case, and there is no need to add a relativistic clause to that statement. The second option would be to declare that we get *moral* normativity

as soon as we have an *agreement* in intentions, without implying that the agreement must necessarily be in the interests of all the parties involved. But obviously such a suggestion would be too crude, for it would not allow us to distinguish, for instance, between moral judgments and judgments of etiquette. And even if this proposal could be improved upon so as to pick only the right kinds of agreement in intentions, it would still run into the problem we mentioned before regarding nonstandard conceptions of moral normativity. Since it does not appear possible to show that relativistic moral statements retain moral normativity in the standard sense, I think that the relativist should agree to say with the error theorist that all substantive moral judgments are false, and present his main claim within the framework of the error theory, as a proposal for reforming our flawed ways of speaking.

#### *A misleading use of moral language*

Finally, another problem, which specifically concerns the analogy between ethics and physics, is that the way authors like Harman and Dreier use the analogy seems to be a source of confusion when two people subscribing to different moral codes make incompatible absolutist moral judgments. As we have seen, the point of the analogy with physics is to support the relativist's claim that we can salvage much of ordinary moral practice and talk even while accepting antirealism about values, by showing that there is nothing wrong, as such, with making moral judgments phrased in absolutist terms. Just as in the case of physics, what matters is that we don't make the false background assumption that such judgments are "absolutely" true. Whenever we use them, we should understand them as involving an implicit relativizing clause. A possible source of concern here is that this seems to imply that two conflicting absolutist moral judgments can both be true at the same time. Suppose Tom accepts a set of moral standards that prohibits the use of violence to defend oneself against an attack, whereas John adheres to a moral code that regards

that practice as perfectly legitimate. Tom says that it was morally wrong of Bob to resort to violence when a young hoodlum attacked him in the street; John says that Bob did the right thing. According to Harman and Dreier's proposal, both Tom and John are right in their judgments. Harman himself explicitly acknowledges that MMR has that paradoxical implication:

According to meta-ethical relativism, there can be conflicting moral judgments about a particular case that are both fully correct. The idea is that two people with different moralities might reach conflicting moral judgments concerning a particular case – for example, one saying the agent was morally right, the other saying the agent was morally wrong – where both opinions are correct. (Harman, 2000, 24)

Now it might seem that Harman's contention is simply untenable, because it involves a straightforward contradiction: two incompatible judgments like those of Tom and John cannot both be true at the same time. A number of authors have faulted speaker-relativism precisely for leading to a logical contradiction (see for example Lyons, 1976). Yet Harman would reply that his view does not have such a consequence, because, as we have seen in section 2, the notion of "truth" to which he appeals when he says that two conflicting absolutist judgments can both be true at the same time is a special one: he thus writes that "we should count a nonrelativistic moral judgement right if the corresponding relativistic moral judgement is right made in relation to the morality accepted by the person making the nonrelativistic moral judgement. The relativist can intelligibly suppose that really conflicting nonrelativistic moral judgements are both right in this sense" (Harman, 2000, 38).

The first problem I see with such an answer is that this appeal to a special notion of truth would seem to considerably diminish the force of the relativist's claim to salvage the notion of truth with respect to substantive moral judgments. Strictly

speaking, it is *not* the conflicting absolutist moral judgments of Tom and John that are both true together, but only their relativistic translations. So, even if we agree to reinterpret absolutist moral judgments in the way Harman and Dreier suggest, it still seems that we won't have salvaged the notion of truth for such judgments in any significant sense. Even if the relativist insisted that the relativist translations of absolutist moral judgments should be treated as substantive moral judgments, since strictly speaking they are the most we can get, this would at most allow him to make a purely verbal point against the error theorist. Indeed, the latter has no problem agreeing with the relativist that such judgments can, literally speaking, be true. He will just deny, for the reasons already given, that they really constitute substantive moral judgments. And no matter what stance one takes on this particular issue, the main thrust of the error theory is to deny that any absolutist moral judgment can, strictly speaking, be true. The analogy between ethics and physics used by Harman and Dreier does nothing to show that such a claim is incorrect.

Secondly, I do agree that when two people share the same moral outlook, it might seem a legitimate and useful shorthand to make moral judgments without attaching to them any relativist clause, even if antirealism about values is true. Indeed, such a clause might then be considered implicit, just as the claim that seven years have elapsed since Bush first became president can be understood without distortion as bearing an implicit relation to the relevant frame of reference. However, it seems that the situation is not the same when people *differ* in their moral outlooks. To say to two people who are making incompatible nonrelativistic moral judgments that they are both right is only likely, in the absence of any further explanation, to puzzle them, and not to resolve the disagreement. If he is to make himself understood, the relativist will have to explain that by this he means that each of them is right according to his own particular moral convictions; and if the speakers insist that they themselves meant to say more than that when making their judgments, the

relativist will have to reply that their pretension to absolute truth is unwarranted. Now since moral judgments seem to involve such a pretension to a degree that is not found in judgments about physics (a novice who first learns about Einstein's Special Theory of Relativity is unlikely to want to continue making judgments about mass or time in an absolute manner, but it seems more difficult to many of us to abandon absolutist moral talk), it would seem more advisable for the relativist to call for a partial reform of moral talk, and to demand that at least in cases where people with different moral outlooks are facing each other, we make it clear that their absolutist judgments are both false, and should be supplemented with a relativizing clause tying their judgments to their own particular moral standards. Such a reform has actually been called for by a number of supporters of moral relativism, notably anthropologists like Melville Herskovits, and it would avoid the defects that Harman and Dreier's proposal seems to have.

Of course, supporting such a reform would seem to blur much of the difference between MMR and the error theory. We have seen that error theorists agree with relativists regarding the ontology of values, and that they are also happy to concede that statements such as "cruelty is wrong for me", i.e. relative to the moral standards I accept, can perfectly well be true. They may perhaps even accept the idea that "absolutist" moral statements might be used as shorthand for their relativized versions between people sharing the same moral outlook. So if it endorsed the reform just suggested regarding our use of moral talk, MMR would appear more like a possible way of developing the error theory,<sup>8</sup> rather than a genuinely distinct metaethical view. Harman and Dreier's more conservative proposal makes MMR more distinctive, but as we have seen it seems to have misleading consequences,

---

<sup>8</sup> Not even necessarily the best way of developing it. Indeed, in some cases error theorists might still want, for practical purposes, to go on using absolutist moral talk even when their listeners do not share, or not yet, their own moral standards – if only when educating children. In such cases, a position like moral fictionalism, in the form defended for instance by Richard Joyce, according to which we ought to go on *pretending* that the substantive moral judgments we accept are true even though we know them to be false, might seem a more attractive option for the error theorist. (See Joyce, 2001.)

in that it is likely to elicit confusion when two non-relativists with different moral outlooks disagree with each other.

### ***Conclusion***

What emerges from the preceding discussion is that metaethical relativism, even when construed as speaker-relativism – a version arguably superior to its alternatives – does not live up to its promise, which was to salvage talk of moral truth despite the relativists’ commitment to antirealism about values. Indeed, it is only thanks to a kind of verbal trick that relativists are able to retain talk of truth for absolutist moral judgments. Strictly speaking, the only kind of moral truths they are successful in preserving is one error theorists can perfectly acknowledge. The conclusions drawn by Dreier and Harman from the analogy with physics are difficult to accept: it does not seem advisable to retain absolutist moral discourse in cases where people do not share a common moral outlook. The often-made suggestion that we ought then to replace moral statements phrased in absolute terms by statements of the kind “action X is right/ wrong *for me*” (at least in cases when we are having a honest discussion with each other, and our primary purpose is not persuasion), seems to be a better option for the relativist. Yet even such a version of MMR does not manage to preserve specifically moral normativity, as opposed to the more general kind of normativity that is peculiar to motivational states as internal reasons for action. Thus understood, MMR seems to be better described as a particular variant of the error theory, rather than as a rival metaethical view.<sup>9</sup>

### ***References***

Boghossian, Paul (2006). “What Is Relativism?”, in P. Greenough and M. Lynch (eds.), *Truth and Realism* (OUP, 13-37).

<sup>9</sup> I wish to thank my thesis supervisor in Oxford, Dr. Jonas Olson, the members of the institute of Philosophy at the University of Neuchâtel, Switzerland, and two anonymous referees, for their helpful comments.

Dreier, James (1990). "Internalism and Speaker Relativism", *Ethics* 101:1, 6-26.

---- (2006). "Moral Relativism and Moral Nihilism", in D. Copp (ed.), *The Oxford Handbook of Ethical Theory* (OUP, 240-64).

Hare, R.M. (1983). *Moral Thinking: Its Levels, Nature, and Point* (Oxford: Clarendon Press).

Harman, Gilbert (1977). *The Nature of Morality: an Introduction to Ethics* (New York: Oxford University Press).

---- (2000). *Explaining Value* (Oxford: Clarendon Press).

Joyce, Richard (2001). *The Myth of Morality* (Cambridge University Press).

Lyons, David (1976). "Ethical Relativism and the Problem of Incoherence", *Ethics* 86:2, 107-21.

Williams, Bernard (1981). *Moral Luck* (Cambridge University Press).

## THE NATURE OF APPLIED MATHEMATICS: REMARKS ON FIELD'S VIEW

DAVIDE RIZZA

UNIVERSITY OF SHEFFIELD

### ***Abstract***

In this paper I raise some objections to Field's characterization of applied mathematics, showing, by means of three examples, that it is too restrictive. While doing so, I articulate a different and wider account of applicability. I conclude with an argument supporting its compatibility with an anti-realistic view on the existence of mathematical entities.

### ***Introduction***

The problem of applicability has always been more or less implicitly present in current debates in the philosophy of mathematics, especially through the discussion of indispensability arguments. Only in relatively recent times, however, (see especially Steiner (1995, 2005)) has it begun to be considered a philosophically relevant one in its own right: in particular, it has been felt necessary to *account for* the applicability of mathematical theories, i.e. to explain how they can model empirical phenomena (or be used to describe them: in this connection see again Steiner (1995, p.137)).

I embrace this perspective in this paper and apply it to a very influential book in the philosophy of mathematics, namely Field (1980), because it has never been discussed in depth as a study of applicability, despite the fact that one of its aims is precisely that of accounting for it (Field 1980, p.6).

My main thesis is that Field's view, although containing important insights into the nature of applied mathematics, severely restricts its role and does so in a way that can be shown to be unrealistic in the face of the structure of concrete applications.

To clarify this claim, some preliminary terminology and a few observations are needed. Firstly, in the context of applied mathematics, I distinguish between qualitative theories, which directly describe possibly idealized empirical domains, and the mathematical theories that are applied to them, which do not and may contain primitives or defined concepts that have no empirical counterpart (e.g. in affine geometry, taken as the geometry of space-time, no operation is defined on points, whereas in its analytical version the addition of coordinates is defined). Here ‘qualitative’ may be understood as synonymous with ‘nominalistically acceptable’ in Field’s sense<sup>1</sup>.

With this distinction in place, Field (1980) can be said to maintain that mathematical theories enter applications as:

- a) Devices for shortening proofs, i.e. carrying out deductions from qualitative theories more simply and quickly (e.g. within a numerical framework).
- b) Languages wherein qualitative theories may be formulated<sup>2</sup>. Here Field claims that platonistic scientific theories (ones formulated in mathematical terms) can always be replaced by qualitative counterparts, in which the use of mathematical terms is avoided.

Both (a) and (b) are based on the idea that mathematics plays a role in applications, which is subordinated to the availability of a qualitative theory. For in (a) deductions already possible within a qualitative theory can simply be translated and more expeditiously carried out within a mathematical, typically numerical context. On the other hand, (b), given Field’s claim, implies that a mathematically formulated theory can be accepted as empirically meaningful only insofar as it is possible to isolate its qualitative content.

---

<sup>1</sup> A qualitative theory is nominalistically acceptable because it has idealized empirical interpretations and thus may count as an ‘empirical’ theory. The same holds for all nominalistically acceptable theories considered in Field (1980) (even though some of them, as affine geometry, have mathematical interpretations, e.g. a numerical one: indeed this is needed, according to Field, for the applicability of mathematics to be possible, insofar as he bases it on the provability of representation theorems).

<sup>2</sup> Think of classical mechanics, where e.g. Newton’s second law of dynamics is formulated as an algebraic equation between numbers (if we take the components of the vector equation).

While I recognize the fundamental importance of qualitative theories to illustrate the way mathematics is related to a qualitative content, I also believe the role played by mathematical theories in applications should not be restricted to (a) and (b). In many cases, the use of mathematical theories is not conceptually subordinated to the availability of a qualitative theory but, on the contrary, directs and organizes its emergence.

In particular, mathematics is used as:

- 1) A guide to the qualitative characterization of an empirical domain;
- 2) A means of analysis of the structure of empirical data;
- 3) A source of (metatheoretical) explanations for the connections between a qualitative theory and a mathematical one.

As can be seen, (1) to (3) differ from (a) and (b). It remains to show that they correspond to the actual employment of mathematics in applicative contexts. The next sections are devoted to doing this and articulating a characterization of applied mathematics richer and stronger than Field's and based on these examples.

This raises an ontological issue, since Field describes applicability in such a way as to make it consistent with its nominalistic project while, on the other hand, it is not obvious whether a different and stronger account of applicability like the one I propose can be retained together with mathematical antirealism. In the last section of the paper I show that it can.

### *The nature of physical magnitudes*

My first illustration of how mathematical theories can guide the development of qualitative theories in sense (1) of section 1 is extensive measurement, i.e. the measurement of additive magnitudes (lengths, masses etc.). This is often formalized by means of an axiomatic theory, which provides a qualitative characterization of magnitudes, and entails a metatheorem to the effect that any model of the theory

(any system of magnitudes) can be embedded into the additive, positive reals (arithmetical addition on the reals is just the numerical interpretation of an additive operation on magnitudes like, e.g., the juxtaposition of rods in the case of length measurement). This means that all systems of extensive magnitudes modelling the axioms have a measurement scale on the real numbers.

The interesting fact concerning this example is that, if one allows for the possibility of infinitely large domains of magnitudes, a proof that a real measurement scale exists requires the assumption of a condition of finite comparability between magnitudes, called Archimedes' axiom<sup>3</sup>. It is this axiom which governs the process of approximation to the 'true' numerical measure of a magnitude and ensures that it is a uniquely determined real number. The classical theory of magnitudes, as given in e.g. Huntington (1902), uses, instead of Archimedes' axiom, the condition of Dedekind completeness<sup>4</sup>, which implies it and forces the structure of a system of magnitudes to be structurally identical to the linear continuum of the reals (something similar happens to lines in the affine geometry used by Field to describe space-time in Field (1980)). In principle this result is delivered by the Archimedean axiom alone because, in presence of the other assumptions of the theory of extensive measurement, it ensures the existence of a unique Dedekind completion for any model of this theory (i.e. a Dedekind complete extension of the model).

The crucial point is that neither Dedekind completeness nor Archimedes' axiom is assumed on the basis of the observable behaviour of magnitudes. The reason is that Dedekind completeness is untestable, being a statement about infinite sequences of magnitudes, which would require an (unfeasible) infinitely precise measurement

---

<sup>3</sup> This says that if, according to some empirical ordering, magnitude  $a$  is smaller than  $b$ , there exists, for some  $n$ ,  $na$ , the result of empirically adding together  $n$  copies of  $a$ , such that  $na$  is greater than  $b$ .

<sup>4</sup> This condition states, for real numbers, that any increasing sequence of reals which is bounded above (i.e. such that there is a fixed real greater than any element of the sequence) has a limit, i.e. a number that bounds it above and is the least number to do so. It is easy to see how the same assumption can be rephrased for extensive magnitudes.

procedure to be checked; on the other hand, as long as experimental practice is restricted to dealing with finite domains of magnitudes, Archimedes' axiom can be dispensed with in the proof of the existence of a numerical scale.

Because of this, Dedekind completeness and Archimedes' axiom are *theoretical* assumptions concerning magnitudes. Strictly speaking, they cannot be considered extrapolations from experiment. In particular Archimedes' axiom imposes a condition on infinitely large domains of magnitudes (or, equivalently, their asymptotic behaviour) which is automatically satisfied by sufficiently structured domains when their size is finite.

To see a way in which finiteness makes the Archimedean condition dispensable<sup>5</sup>, consider the elementary case of mass measurement performed with an equal arm balance. In this context an ordering on masses is defined by looking at whether the balance is or is not in equilibrium (it may be assumed that the balance be infinitely precise, since imprecision doesn't bear on the point under discussion): thus, given objects  $x$  and  $y$ ,  $x >_m y$  means that, when  $x$  and  $y$  are put on the distinct pans of the balance,  $x$  descends while  $y$  ascends. A weak ordering on masses (i.e. a transitive and connected binary relation) can be defined by similar observations. In addition, it is possible to put several objects on a single pan of the balance: if they are called  $x_1 \dots x_n$ , then the sum  $x_1 +_m \dots +_m x_n$  is their physical addition, whose mass equals the mass of any object which equipoises  $x_1 \dots x_n$  when they are all on one of the balance's pans. Order and addition on masses defined on a domain  $M$  of physical objects determine a structure  $\mathbf{M} = \langle M, \geq_m, +_m \rangle$  abstractly describing an experimental measurement setting: the experimental procedures applied to this setting always give rise, if actually performable, to finitely many comparisons between concatenations of masses.

---

<sup>5</sup> In this connection also see Suppes (1969, p.4–8).

Given a first order language for  $\mathbf{M}$ , this finite number of comparisons can be denoted by a finite set of atomic formulas, which might be read, if a scale of measurement existed, as linear inequalities. The possibility of finding real measures for the elements of  $\mathbf{M}$  (supposing it to be finite) is then reduced to the existence of a solution of a system of linear inequalities in the positive reals (this also means that  $\geq_m$  and  $+_m$  can be respectively interpreted as order and addition on the positive reals). This way of setting up a scale of measurement does not require an Archimedean axiom. The reason can be briefly illustrated as follows: consider a set of axioms  $A$  for extensive measurement, e.g. the one in Huntington (1902) mentioned above, and call  $A'$  the axiom system obtained from  $A$  by removing completeness (so that no Archimedean condition can be proved from the remaining axioms) and adding the commutativity of the binary operation  $+_m$ .

It has been shown in Adams et al (1970)<sup>6</sup> that the existence of a structure-preserving mapping from  $\mathbf{M}$  into the positive, additive reals (i.e. the existence of a real solution for a system of linear inequalities generated by a structure of the type of  $\mathbf{M}$ ) is equivalent to  $\mathbf{M}$  satisfying  $A'$ <sup>7</sup>.

Thus, whenever we want to measure a finite extensive structure, we can make use of an algorithm, providing the solution of a system of linear inequalities, which does not rely on the usual technique of successive approximations based on Archimedes' axiom.

This result shows that, while the real, Archimedean continuum may provide an important heuristic model to study extensive measurement, its choice is by no means

---

<sup>6</sup> The original result, for which cf. Adams et al (1970, p.397–398), is formulated in a different manner and for a different axiom system (namely, a modified version of Suppes (1951)). However  $A'$  entails the statements of that axiom system and, on the basis of this, my statement of the result is essentially the same as the original one.

<sup>7</sup> Since the existence of a structure-preserving mapping as the one above also entails that all scales of measurement for  $\mathbf{M}$  are related by a positive constant factor, this result itself is independent, for finite extensive structures, of Archimedes' axiom.

forced upon us by experiment. It therefore appears plausible to look at certain non Archimedean structures extending the reals (i.e. structures which may be viewed as enlargements of the reals, containing infinitely large or small numbers, which violate Archimedes' axiom) and examine the possibility of whether they may provide an alternative numerical model guiding the abstract conceptualization of extensive magnitudes.

It turns out that, following this approach, systems of magnitudes which are structurally different from the reals but admit a measurement scale on them can be described. Their particular qualitative characterization is suggested by a purely numerical fact, namely the relationship between the real numbers (more precisely, a structure like the real field or, possibly, the full real structure) and their non Archimedean elementary extensions, which may be obtained by means of the compactness theorem of first order logic. An application of the latter makes it possible to determine, given a suitable real structure, an extension thereof containing infinitesimals<sup>8</sup>. If a sufficiently rich real structure is considered (e.g. the ordered field structure, including addition and multiplication) it is possible to 'add' to its domain infinitely large and small numbers (i.e. numbers that are respectively greater than any positive integer or smaller than its reciprocal in the ordering extending '>' on the reals). In particular real numbers may be surrounded by other numbers that are infinitely close to them.

While (i) any two real numbers have a finite difference, and thus a sequence tending to one of them as its limit must at some point finitely diverge from the other, (ii) in the elementary extensions of the reals there may be sequences that have only an infinitesimal divergence and thus end up being infinitely close to the same real limit.

---

<sup>8</sup> A proof of the fact that elementary extensions of the full real structure contain infinitesimals (but the ordered field structure is sufficient to achieve it: as a matter of fact, because the field structure imposes the existence of multiplicative inverses, the existence of infinitesimals entails that of infinitely large numbers) may be found in Narens (1974, p.378– 379). An alternative, algebraic construction of elementary extensions of the reals using ultraproducts is very clearly described in Goldblatt (1998, p.23–27).

As long as the reals are used as a privileged model for the qualitative characterization of extensive magnitudes, the latter's behaviour conforms to (i): as a result, magnitudes can be identified with uniquely determined physical states, corresponding to the limits of certain sequences of approximations. On the other hand, (ii) makes it possible to think of magnitudes not as uniquely determined physical states, but as ones which are subjected to infinitesimal oscillations (it is sometimes assumed, in physics, that 'negligible', infinitesimal variations of magnitudes are possible: for example see Segel (1991)). It follows that the mathematical relationship between the reals and their elementary extensions, as briefly outlined by the contrast between (i) and (ii), suggests two alternative characterizations of the qualitative structure of additive magnitudes.

At the same time, it is noteworthy that (the possibility of doing so is rigorously shown in Narens 1974), if one thinks of magnitudes as oscillating within infinitesimal regions, it is still possible (as long as no infinitely large or infinitely small ones arise) to scale them on the reals, for infinitely close magnitudes can just be assigned the same real measure, as if their infinitesimal difference were negligible.

The results of ordinary measurement practice are compatible with a qualitative understanding of magnitudes which diverges from the classical one. But such a qualitative understanding has been suggested by the mathematical relationship between two numerical structures. This means, therefore, that numerical facts can be used to modify the intrinsic characterization of physical entities like lengths or masses or, equivalently, to integrate experimental data into different abstract frameworks.

### ***The chronological order of deposits***

Point (2) of section 1 qualifies mathematics as a means of data analysis. To justify this remark and explain exactly how (2) differs from Field's account of applicability,

I will use an example taken from archaeology. It is often the case that archaeologists need to solve what are called seriation problems. A seriation is the chronological ordering of different deposits within one or more distinct sites: this problem is relevant to the present discussion because it often happens that certain seriations are not known nor can be determined by chemical or physical tests. In this situation one needs to find a chronological ordering of deposits, starting only from information concerning the artefacts they contain. Because of this, a representational strategy like the one used in Field (1980) to describe applicability, is not available nor pertinent to the problem at hand. The reason why this is so is twofold: firstly, we cannot numerically label the deposits to obtain a numerical chronology, because we haven't got enough information to assign the appropriate labels; secondly, we primarily wish to determine a qualitative chronological order, rather than giving the numerical description of one.

Now, mathematics enters this problem not as a way of formulating a theory of order for deposits or as a way of expediting proofs concerning qualitative ordered structures, but as a way of processing archaeological data to generate a chronology. Mathematics works as an instrument guiding the identification of a qualitative structure. The details of how the seriation problem may be solved can be found in Shuchat (1984): here I restrict myself to outlining the mathematical strategy adopted, in order to clarify the nature of its application in this context. The seriation problem is in essence solved by giving a suitable mathematical presentation of the available data and, subsequently, by working on the abstract configuration of data thus obtained by means of relevant mathematical concepts and theorems. The data the archaeologist has at her disposal concern the number of deposits she wants to order and a classification into types of the artefacts occurring in those deposits<sup>9</sup>. Deposits and types are arranged into an incidence matrix, whose

---

<sup>9</sup> Deposits are identified with, roughly, 'points' in time at which they were formed, while types are identified with period of time during which they were in use.

entries  $a_{ij}$  are only 1 or 0: in particular,  $a_{ij} = 1$  if deposit  $i$  contains type  $j$  while  $a_{ij} = 0$  otherwise.

Under the assumption (h) that a given type be present in all deposits corresponding to the period through which it existed, it follows that the chronological ordering of the deposits, i.e. the chronological ordering of the rows of the incidence matrix  $[a_{ij}]$ , must give rise to a matrix whose columns do not contain any zero-entries lying between the 1-entries (otherwise (h) would be contradicted). A matrix satisfying this property is called a *Petrie matrix*, whereas a matrix  $M$  such that a permutation  $P$  of its rows generates a Petrie matrix is called a *pre-Petrie matrix*.

The problem of seriation is thus reduced to the mathematical problem of finding the necessary and sufficient conditions under which an incidence matrix is a pre-Petrie matrix: if it is not (in case, that is, (h) does not hold), the goodness of fit of a seriation can still be estimated, by finding a measure of how close an incidence matrix is to a pre-Petrie form.

Note, again, that the mathematical solution to the seriation problem is not based on the mathematical deduction of facts which could also be deduced from an associated qualitative theory: here we are not working with a general theory of finite orderings, but with a type of data configuration from which an ordering has to be extracted.

Once the seriation problem is formulated in terms of matrices, its solution comes from establishing a correlation between two mathematical theories, namely matrix algebra and the theory of networks (undirected graphs). This is because, multiplying  $M$  by its transpose<sup>10</sup>, one obtains a symmetric matrix  $S$ , such that  $s_{ij} =$  (number of types common to deposits  $i$  and  $j$ ).  $S$  is called a *similarity matrix*

---

<sup>10</sup> The transpose of  $A$  is obtained from  $A$  by interchanging its rows and columns.

because its entries tell us how many types occur simultaneously in any two deposits.

If there are  $n$  types, then  $d(i, j) = n - s_{ij}$  tells us how many types occur in one but not in the other deposit, i.e. gives us a measure of dissimilarity between them. This shows that there is an explicit relationship between the data configurations described by  $M$  and  $S$  and the qualitative content of the seriation problem: nonetheless, the solution to the problem – this is the crucial point – follows from a purely abstract analysis of the data configuration involved (none of the operations performed on it has a qualitative interpretation).

Using the dissimilarity function  $d$ , we can associate to an incidence matrix  $M$  a *path* through its  $m$  rows (given by the order in which the rows follow one another), to which the length  $\sum d(i, i + 1)$  is associated ( $i$  ranges over the  $m$  rows of  $M$ <sup>11</sup>). To any permutation of the rows of  $M$  there corresponds a path, whose length can be computed<sup>12</sup>. It is intuitively clear that, if the length of the path turns out to be minimal, we have minimized the dissimilarities between adjacent rows, thereby obtaining a chronological ordering. In particular it can be shown that a lower bound for the paths' length exists, which is attained if and only if  $M$  is a pre-Petrie matrix. If on the other hand  $M$  is not pre-Petrie, it is possible to work on the difference between the length of its associated paths and their lower bound: minimizing it leads to a 'best' approximation to a pre-Petrie matrix and, consequently, to a candidate solution for the seriation problem.

This method can be refined in several ways, for instance by using stochastic matrices instead of incidence ones (whose entries are the probabilities of finding

<sup>11</sup> Usually one adds an arbitrary 0-row to get the domain of  $i$  right and, when  $i = m$ , identifies  $m+1$  with the 0-row.

<sup>12</sup> This is actually a simplification of the original mathematical treatment, which involves Hamiltonian paths. I do not talk about them here because I'm only interested in presenting the mathematical strategy in an intuitive, informal fashion.

a certain type in a certain deposit) and provide a more realistic framework to describe chronologies. In any event, the mathematical procedure adopted to find a chronological ordering entirely relies on an examination of the mathematical properties of a data configuration. What is gathered from it is the linear arrangement of a qualitative structure.

***The continuous distribution of the gravitational potential***

To illustrate (3) of section 1, I consider a problem in Field (1980) and the way it is implicitly solved therein. The problem is to find the qualitative formulation of an analytical property of functions (thus a mathematical property). More explicitly, the ordinary treatment of the classical theory of gravitation (developed in terms of fields), that Field nominalises, involves the reference to a function  $f$  measuring the gravitational potential at any point in space and which may be defined through time as well. This is a numerical function from ordered quadruples of reals (the coordinates of space-time points) into real numbers (the values of the potential at those points).

The continuity of  $f$  can be defined by means of the usual Weierstrass  $\varepsilon$ - $\delta$  condition, stating that, for any point  $x$  in space-time, whenever a positive real  $\varepsilon$  is chosen, it is possible to find an open ball of centre  $x$  and radius  $\delta$ , such that all points of space-time in it have a value of potential that differs by less than  $\varepsilon$  from  $f(x)$ .

Because this definition involves a reference to the concept of distance (the radius of an open ball), it is not, as stated, expressible qualitatively within the framework Field adopts, for both the axiomatic geometry of space-time and the qualitative theory of the potential he uses lack, among their primitives, a notion of distance or congruence: their primitives are only order relations.

In order to express continuity in this context one has to make it independent of distance: that this is intuitively possible follows from the fact that the metrical

definition given above only involves, in essence, the idea of the closeness together of certain points. Closeness can be described in terms of order alone, by talking about arbitrary (and thus also arbitrarily small) intervals or regions, which may be defined in terms of order alone.

What Field does is to resort to a more general notion of continuity, which is formulated in the context of a mathematical theory, that of topological spaces. Here it can be proved that, under certain conditions (the so called first axiom of countability<sup>13</sup>), the metrical, analytical definition of continuity is equivalent to one formulated in terms of open sets, which is nonmetrical in the sense that it does not require a notion of distance in its formulation (a function  $f$  from a topological space  $S$  into a topological space  $T$  is continuous when, for any open set  $X$  of  $T$ ,  $f^{-1}(X)$  is an open set of  $S$ ).

Such nonmetrical definition can be formulated using only the primitives of affine space-time geometry and the (affine) theory of potential discussed by Field: thus, exploiting topological continuity, he can find a qualitative characterization of the mathematical property of continuity.

Topology works here as an abstract mathematical theory bridging a qualitative and a numerical theory. Semantically, any model of geometry or its real counterpart can be endowed with a topology and the availability of a topological notion of continuity makes it possible to interpret it within different models and, as a consequence, to connect different theories.

This semantic fact can be also understood as a relationship between the languages of the three theories involved, for the general definition of continuity in topological terms can be translated within the other theories by means of their primitives. At the same time, the assurance that this new notion of continuity is strong enough

---

<sup>13</sup> For the technical details see Pfanzagl (1968, Ch. 2).

to mirror the classical one is provided by a theorem of topology and therefore established deductively.

These remarks show that the importance of abstract mathematical theories in applications can be reduced to their deductive content, leading to the introduction of certain important logical connections between the relevant concepts. In other words, it is not their possible reference to abstract entities which makes the mathematics useful or conceptually crucial. I will get back to this point in the concluding section, in order to explain why my view on the applicability of mathematics, although different from Field's, remains compatible with an anti-realistic position concerning the existence of mathematical entities.

My main point for now is that, from the example above, one can see how mathematical theories work in applications as metatheoretical explanations for the connections between a qualitative theory and a mathematical one. The previous discussion has shown that the assumption of continuity, ordinarily made in the analytical treatment of the gravitational potential, is physically meaningful, because it can directly be described as a property of the qualitative variation of a magnitude in space-time.

This result, which is obtained at a metatheoretical level (because it involves the interaction of several theories and the interpretation of concepts from one theory into another), explains precisely why it makes sense to use the numerical property of continuity to describe an ordered physical geometry and a quantity varying over it: without it, the desired qualitative theory (one inducing continuity on the analytical representation of the potential) could not be formulated and thus we have a situation where mathematics is used to guide the formulation of a qualitative theory.

***Conclusion: Applicability and realism***

I take it that I have shown in the foregoing discussion how mathematical theories

can enter applications in a way that does not correspond to any of those highlighted by Field ((a) and (b) of section 1). Now I wish to emphasize the common aspects of the examples I have given and formulate precisely the reason why they are compatible with anti-realism about mathematical entities. By doing so, I intend to make apparent that a significantly richer role than the one identified by Field may be ascribed to the applicative use of mathematics, while this does not necessarily require to embrace some form of mathematical realism.

As already remarked, all the examples I have proposed describe a way in which a mathematical theory aids the characterization of a qualitative theory or a qualitative structure. The former case was illustrated in my last example, where it was possible to identify, by means of topology, a qualitative axiom corresponding to the analytical condition of continuity of the gravitational potential. The latter case (characterizing a qualitative structure) was described in my second example on archaeology, where the problem at hand was to identify a qualitative order. Finally, my first example involves, in a sense, both the aspects occurring in the other two, for it illustrates a change in the qualitative characterization of additive magnitudes but also an alternative basis for their numerical representation on the reals.

Thus all the examples I have proposed highlight the fact that mathematical theories (non standard analysis, the theory of networks and the theory of real vector spaces) are not necessarily subordinated to corresponding qualitative theories in applications, in the sense that they can legitimately be used only in presence of an independent and fully informative qualitative translation thereof.

In the case of extensive measurement a new qualitative theory is suggested *through* mathematics and the general strategy of Field (1980) is, after all, to take the standard analytical treatment of classical gravitation theory and look at the kind of qualitative structure it induces (i.e. the kind of axioms a physical domain must satisfy in order

for them to be analytically representable in the standard way). Thus, it is apparent that mathematics plays a heuristic role in the formulation of qualitative theories. This resembles (b), but it is crucial that in (b) we only have a given mathematically formulated theory which we then must try to nominalize, whereas e.g. in my first example we can clearly already have a qualitative theory and then modify it on the basis of certain mathematical results (the relation between the reals and their elementary extensions). The same is also true of my last example.

Finally, the discussion of seriation shows how one can reconstruct qualitative information simply by giving a mathematical presentation of a set of data and working on it within the context of a mathematical theory. Again, it is obvious that this process is attached to the original qualitative setting (indeed it must be), but its relevant feature is that it acts as a form of abstract data analysis and not as a mathematical translation of qualitative facts or as a deduction from a numerical representation thereof.

Because of these remarks, it follows that Field's characterization of applied mathematics is very restrictive and should be enlarged, if it were to cover a sufficiently wide range of situations and thus provide as comprehensive as possible an account of applicability. One worry might arise in this context, due to the possibility that the acceptance of the wider account of applicability I have outlined may imply a realistic commitment to mathematical entities, which obviously the nominalist wishes to avoid.

However, this worry can be dissipated by showing that the uses of mathematical theories I considered are compatible with mathematical antirealism. The reason why it is so has already been mentioned in the previous section.

It is, essentially, that mathematical theories are fruitful as guides in the construction or analysis of qualitative theories or structures because of the logical analysis they provide of these theories and structures, which, in turn, is articulated deductively

by their theorems. It is not their referring to abstract objects to be actually used, but rather their proving certain results which direct the way qualitative settings may be dealt with. At the same time, these results are motivated by the nature of the applicative problems at hand.

For instance, in my first example the relationship between the real numbers and their elementary extensions is based on viewing the former as measures or evaluations of magnitudes, rather than purely mathematical entities (there is no space here to elaborate upon this remark: let it suffice to say that numerical measures can essentially be constructed as indices encoding experimental information coming from observational reports, as shown in Niederée (1992)). The application of the compactness theorem illustrates how these measures can describe two different idealized processes of empirical approximation, involving limits or infinitesimal differences. Thus, the purely algebraic relation between the reals and their non Archimedean extensions takes its motivation from the need to characterize the asymptotic behaviour of empirical approximations. This is confirmed by the fact that Narens (1974) contains a theorem showing the embeddability of extensive magnitudes structures for which Archimedes' axiom is not assumed into elementary extensions of the reals, a result which parallels the embeddability of the Archimedean structures into the reals. Because the concept of representation is involved in both cases, a representational treatment of measurement in the style of Field (1980) goes through: however, it is clear that such treatment is subordinated to the analysis of the relation between the numerical representing structures (the reals and their extensions), which is used as a heuristic to reconceptualize the notions of magnitude and of measurement scale.

In the example from archaeology, on the other hand, the relevant theorems are those of matrix algebra and the ones giving bounds for certain networks (the paths associated to the incidence matrices and their permutations), motivated by the need

to analyze data configurations. Finally, in the case of gravitation theory the key result is a theorem of topology, establishing the equal strength of two definitions of continuity and motivated by the need to isolate the essential features of an analytical concept of quantitative variation, in order to make it expressible within a non-numerical setting.

It is through the above theorems that we can better understand how to deal with qualitative settings: as long as the focus is on them, i.e. on deductively obtained correlations motivated by ultimately empirical needs, it is not necessary to regard mathematics as carrying with itself an inescapable ontological commitment, because it is used essentially to direct model construction through proofs or to reflect upon the features of empirical models on the basis of data coming from them.

As a result, it is possible to keep together an anti-realistic view similar to Field's with the account of applied mathematics I have expounded here.

### *References*

Adams, E.W., R.F. Fagot and E.R. Robinson 1970. On the Empirical Status of Axioms in Theories of Fundamental Measurement. *Journal of Mathematical Psychology*, 7, p.379–409.

Field, H., 1980. *Science without Numbers*. Oxford: Blackwell.

Goldblatt, R., 1998. *Lectures on the Hyperreals*. London: Springer.

Huntington, E.V., 1902. A Complete Set of Postulates for the Theory of Absolute Continuous Magnitudes. *Transactions of the American Mathematical Society*, 3, p.264–279.

Narens, L., 1974. Measurement without Archimedean Axioms. *Philosophy of Science*, 41, p.374–393.

Niederée, R., 1992. What do numbers measure? A new approach to fundamental measurement. *Mathematical Social Sciences*, 24, p.237–276.

Pfanzagl, J., 1968. *Theory of Measurement*. Würzburg-Wien: Physica-Verlag.

Segel, L. A., 1991. The infinite and the infinitesimal in models for natural phenomena. *Reviews of Modern Physics*, 63, p.225–238.

Shuchat, A., 1984. Matrix and Network Models in Archaeology. *Mathematics Magazine*, 57, p.3–14.

Steiner, M., 1995. The applicabilities of mathematics. *Philosophia Mathematica*, 3, pp.129–156.

----- 2005. Application and Applicability. In S.Shapiro, ed. *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford: Oxford University Press. Ch. 20.

Suppes, P., 1951. A set of independent Axioms for Extensive Quantities. *Portugaliae Matematica*, 10, pp.163–172.

----- 1969. *Studies in the methodology and foundations of science: selected papers from papers from 1951 to 1969*. Dordrecht: Reidel.

## THE ROLE OF AFFECTIVE STATES IN PHILOSOPHICAL INQUIRY

AMANDA TAYLOR

UNIVERSITY OF DURHAM

### ***Abstract***

Affects have often been characterised as a hindrance to the rational thinker. In this paper I reconsider the role of affects in philosophical inquiry in the light of recent work on the emotions which suggests that affects play a role in framing the ways in which we experience the world. I explore affects as motivators and curtailers of philosophical inquiry drawing on work by Hookway (2002, 2003). I suggest that although Hookway is correct in identifying the motivating role of affects, his account is too sparse and does not take account of the wider role which affects play in philosophical inquiry. Drawing on phenomenological psychiatry I argue that affects play a pre-reflective role which enables successful 'explicit' reasoning to begin. Building on this I use accounts by de Sousa on emotions as salience providers (1980, 1987, 2004) and William James on the role of temperament in philosophical inquiry (1992) to supplement and fill the gaps left by Hookway's account. Here I draw a distinction between sporadic and sustained affects, claiming that a full account of the role of affects in philosophy ought to take account of both. This paper ultimately provides an examination of how the recent work on emotions affects the way we may view philosophical inquiry.

### ***Introduction***

The recent popularity of philosophical examinations of 'the emotions' has thrown new light on the way we understand ourselves, others and, arguably, philosophical methodology. Such work deviates from a 'traditional' understanding of affects as hindering rational projects and inquiries. Instead it focuses on the cognitive role of emotions, suggesting that they play a constructive role in shaping our experience of

the world and the way we approach problems (Damasio 1995, 2000; de Sousa 1987; Solomon 1993, 2004, 2004(ed)). If these accounts are to be taken seriously, then there must be ramifications for our conceptions of 'how we do philosophy'. In this paper I examine the effect this work has on the way we view philosophical inquiry by examining the relationship between affect and reasoning. I draw on three areas: affects as motivating<sup>1</sup> and curtailing philosophical investigation (section 1); affects as salience bestowers (section 2); and affects as dissolving deadlocks which appeal to reason alone cannot achieve (section 3). In what follows I use to 'affect' to denote a cluster of phenomena which are felt states and which are prior to reasoning<sup>2</sup>.

Following from these accounts I suggest that affects serve to 'draw us into the world' forming a tacit background from which successful reasoning must begin. Considerations from psychopathology highlight the fundamental role which affects play in healthy reasoning ability. I also consider the motivational influences of affects in philosophical inquiry, following Hookway's (2002; 2003) discussion of the feelings of doubt and conviction. Although Hookway's account provides an important discussion of affects as motivating inquiry, it fails to take account of how affects influence the course the inquiry takes. In light of this the paper moves on to consider ways in which affective states mould our philosophical inquiry through rendering certain features of the philosophical landscape salient. This leads into a brief examination of affective dispositions as what I term 'marginalised influencers'. Affects here, though traditionally disregarded as playing an integral role in the rational thinker's project<sup>3</sup>, influence the inquirer's choice of explanation and methodology. Here I draw largely on William James's account of temperament and the role it plays in forming a kind of philosophical allegiance. Questions of

---

<sup>1</sup> I use 'motivation' here in the restricted sense, that is, motivation of philosophical reasoning and not in terms of motivating action.

<sup>2</sup> This term incorporates emotions, feelings, dispositions, predilections, temperament and moods.

<sup>3</sup> See for example Descartes (1911a, p.291) in which he writes that 'affections or passions ... are confused thoughts'.

favouring one view over another without a ‘rational’ motivating reason shall feature here.

It shall become apparent that affects can be largely described as *boundary producers*, setting the limits of reasoning, encapsulating the field of reference to which we can refer in setting the bound of our enquiry. Affects, therefore, evoke a defined field of relevance in which inquiries are pursued, setting the reasonable excitation of doubt and curtailing it when we meet with conviction. Further, they have influence over the direction which we steer through the course of our inquiry<sup>4</sup>.

### ***Doubt and Conviction***

Hookway (2002; 2003) highlights how philosophical inquiry is motivated and curtailed by affects, in particular those of doubt and conviction. Doubt, as a felt bodily affect, induces the philosopher to begin their inquiry, whilst feelings of conviction act as a marker which curtails it (2002, p.256-259). In not only identifying, but in feeling the object of doubt to be ‘out of place’ the philosopher is motivated to question and examine further. It must be noted that in order to doubt something, background feelings of *security* or *trust* must first be in place; in order for something to feel ‘out of place’ there must first be a contentment or feeling of security which is compromised by the doubting. This feeling of security is tacit, forming the way we primarily occupy a philosophical stance; it is only revealed when an object of doubt causes feelings of unease.

---

<sup>4</sup> I wish to make it clear that in this article I restrict myself to highlighting the neglected role which affects play. In doing this I by no means wish to imply that philosophical reasoning can be reduced to, or invalidated by influence of affect. Instead, I merely emphasise that in philosophy the role and influence of affect has been neglected or largely cast in negative terms. This paper highlights this neglect by giving insight into how affects may *influence* philosophical reasoning in virtue of affects orienting us ‘in the world’ (either as salience providers, as de Sousa writes or in the form of temperaments as James claims). In using writers with differing account of emotion I show how, despite their different conceptions of emotion, these writers identify how affect and reasoning cannot be as easily polarised as has been assumed by much philosophical literature. Further, despite their difference, they recognise affects’ often central role in reasoning. Thus, my paper moves away from thought which opposes rational thought and affective influence, highlighting the ways they combine in philosophical reasoning and also, in everyday life.

In highlighting the role which doubt plays in motivating inquiry, Hookway attempts to show that affective states play a genuine role in philosophy. Philosophical inquiry is therefore underpinned by feelings of doubt and comfortableness with a philosophical system (2002, p.256-257). Doubt, for Hookway, is a feeling aligned with *salience*, whilst feelings of comfort allow things to retreat into a state of acceptability. He states: ‘real doubt involves some question or problem becoming salient’ (2003, p.93). Section 2 offers an account of the role of affects in conjunction with salience; for now, I wish to emphasise with Hookway that doubt serves to render certain features of a system as salient and thus as candidates for investigation. This may be easily seen on appeal to certain figures in philosophy, most notably Descartes.

Since Descartes famously asserted his method of doubt (1911a, p.219; 1998, p.18), doubting has been a key component within philosophical methodology. However Descartes, who rallied against the dangerous influences of ‘the passions’ for the rational thinker (1911b, p.425-427), failed to distinguish doubt as an affect. This suggests a more instrumental, less affective way of construing doubt. Doubt, in Descartes’ sense, becomes less of a motivating bodily feeling, and more of a methodological mindset utilised to arrive at foundational truths. With this in mind we may distinguish between ‘felt’ and ‘abstract’ doubt: that is doubt as *feeling* and doubt as *mind-set* which does not involve bodily feelings.

In his treatment of doubt as affect, Hookway fails to distinguish the ways in which doubt may indeed be ‘felt’. In his account doubt, and its related feeling of conviction, are very much felt in a *bodily* sense (2002, p.253); something is felt to be wrong, to be suspicious, to be treated with caution and, for Hookway, this feeling motivates closer philosophical scrutiny (2002, p.255-256). However, as we have seen, doubt constitutes a spectrum of states which might range from a spontaneous, felt bodily sensation to a mindset occupied in undertaking philosophy aimed at ‘foundational truths’. Although felt doubt might be the strongest instance of

doubting I contend that doubt may also be a mind-set providing an approach and methodology to our philosophical inquiries; thus Hookway fails to call into question whether a philosopher *always physically feels* doubt when they claim to doubt the external world around them or a set of propositions. Following these lines of thought, it is my contention that one may doubt without being assailed by the *physical feeling* of doubt. The philosopher may be provoked to begin inquiry through felt doubt, and yet may steer his way through the inquiry with ‘abstract doubt’. In this way the practice of philosophy involves both the abstract and felt modes of doubt at different stages in the inquiry. Hookway, then, needs to provide a subtler account of doubt which takes account of the spectrum of types of doubt involved in philosophy, in which felt doubt as motivator would feature.

This philosophical methodological doubt may be contrasted sharply with a very different kind of extreme felt doubt which features in many mental illnesses, notably schizophrenia. In many cases of schizophrenia the patient undergoes a loss of affect which impacts directly upon their reasoning abilities<sup>5</sup>. The schizophrenic suffers from an affective disorder in which they lack the feeling of tacit security which normally reigns prior to doubt. This entails a loss of the framework of salience and security, procured by affects, which facilitates healthy reasoning. The feelings of security which orient us within the world give way and the sufferer occupies a constant and unrelenting felt doubt (Sass 2004, p.136).

Thus, whilst Hookway is correct to identify the important role which feelings of doubt and conviction play in respectively motivating and curtailing philosophical inquiry, the strength and breadth of the feelings associated with doubt form a wide

---

<sup>5</sup> For detailed accounts of the relationship between loss of affect and reduction of practical reasoning ability in schizophrenia see: Blankenburg 2002; Sass 1992a, 1992b, 1994a, 1994b, 2001, 2004; Stanghellini, 2004 and Sèchéhaye 1970.

spectrum; the strongest feelings and widest scope for doubt being symptomatic of certain psychotic illnesses. It might be convincingly argued therefore that a philosopher may doubt the veracity or soundness of a proposition without *feeling* the unrelenting doubt which marks the condition of schizophrenia<sup>6</sup>. However, alongside the spectrum of doubt which I have identified at play in philosophical inquiry, affective feelings play a different role in *preventing* the philosopher from feeling doubt in the same way as the schizophrenic<sup>7</sup>; in this case, the feeling of at-homeness or what I have called *security* underlies the philosophical exercising of doubt as an instrumental mind-set. It invokes a pre-reflective trust in the world and one's place in it, facilitating an ability to 'get on with things'. In this way, affects orient and draw us into the world, allowing feelings of 'homeliness' and preventing the radical sceptical doubt of disorders like schizophrenia.

From this, we may borrow from the work of certain phenomenologists, in conjunction with what psychopathology has to show to present how it is that affects, more precisely *moods* (Heidegger 1962, p.172-179) constitute an affective platform from which philosophical inquiry, or any type of successful reasoning must begin (Heidegger 1962, p.84; Husserl 1989, p.183)<sup>8</sup>. Here affects form a lens through which features of the world are filtered and presented in a certain light. Moods, for Heidegger, attune us to the world, structuring and disclosing it to us in a light corresponding to the particular mood. The effect of this may vary from producing 'a way in which the world appears' (Ratcliffe in press, p.1), or a general context within which reasoning is tinted; or more parochially, it may reveal certain features of the environment within which we operate as

<sup>6</sup> Doubt involved in schizophrenia is also of a greater breath. Whilst few objects appear as candidates for doubt to the philosopher, the schizophrenic doubts everything. Termed as 'loss of self evidence', it is due to the absence of an affective background feeling of *security*, or in phenomenological terms, the feeling of 'everyday Being-in-the-world' (Sass 1992, p.2-8, p.45; Stangellini 2004, p.112-115)

<sup>7</sup> See Sass 1992b

<sup>8</sup> For a non-phenomenological account see Damasio 1995, p.150.

salient, or conversely as something to be passed over due to its uncontentiveness or indifference to our projects. Numerous cases discussed by phenomenologist psychologists such as Sass (1992a, 1999, 2004), Stanghellini (2004, p.95-109), and Blankenburg (2002) show that loss of these affective filters correlates with a loss of practical reasoning. They argue further that loss of affect is responsible for the phenomenological changes the schizophrenic undergoes in experience; the schizophrenic cannot constrain the field of reference in inquiry. Unlike the philosopher, the schizophrenic does not know when to stop doubting.

### ***Salience***

From phenomenological writers we have begun to see how affective states orient us in the world. In this section the idea of *how* affective states may do this within philosophical inquiry will be pursued. Two areas will be examined forming a wide and narrow focus. The wide looks at the role of affective states in producing what de Sousa terms ‘patterns of salience’ which configure the way we broadly view the world (in line with the Heideggerian idea of moods above) and more intricately, of how we view problems or inquiries. This supplements and fills in the gaps of Hookway’s account (2002, 2003) which focus primarily on the methodological importance of feelings of doubt and certainty for philosophic inquiry. I go further than Hookway suggesting that these ‘feelings’ are key in not only beginning and ending the route of philosophical inquiry, but that they play a more active role in *framing* and *landscaping* the course our philosophical inquiry takes. As we have seen, for Hookway, doubt and conviction are instrumental in beginning and ending inquiry; however salience provided by affective states steers a way though, proffering certain routes more readily than others, and flavouring the journey our enquiry makes due to the more general sense in which affective states may be said to attune us to the world.

Ronald de Sousa states that '*[e]motions are determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies*' (1980, p.137)<sup>9</sup>. de Sousa is primarily concerned with the *rational value* of emotions, turning away from caricatures of emotions as divorced from reason and a hindrance to rationality. de Sousa works against this conception to show how affective states 'underlie' the process of reasoning (1980, p.141). Although his work is primarily concerned with how affective states affect cognition in general, I will draw on his work to illuminate and expose implicit assumptions of 'how we do philosophy'. I shall issue a cautionary warning before I properly begin: although I am speaking largely of 'positive' roles of affective states on philosophical methodology, this by no means equates to a wholesale unreflective promotion of affective states in philosophy. There are affective states which may hamper a philosophical project; for example, those leading to dogmatism (Hookway 2002, p.256; James 1956, p.82) – such as pride in one's own system or contempt for others and so on may serve to close off the openness to new ideas which challenge ones already held, which is (supposedly) integral to the spirit of philosophising. However, as already stated, it is my purpose to highlight the manner in which recent literature on emotions might be brought to bear on assumptions made on how philosophy is done.

In his examination of the rational role of affective states, de Sousa describes how it is that affective states are intrinsically linked to saliences (2004, p.66-67). For him they are salience bestowers.<sup>10</sup> This concept is important and is one that has hitherto been neglected by philosophy, which traditionally construes itself as working apart from the influence of affects. The question of what forms the field of our inquiries must be important to any philosopher, and is one that has not been adequately pursued. It impacts directly on methodological assumptions in philosophy which

---

<sup>9</sup> See also de Sousa 1987.

<sup>10</sup> See de Sousa 1980 for further information.

have previously not been questioned. It is, therefore, a question which will impact not only on theorising, but on the *structure* of philosophical theorising, that is how we are disposed to approach and work through problems identified by philosophy. The identification of philosophical problems hinges upon what the philosopher takes to be questionable either through doubt or through the particular way their experience is structured by affects as touched upon in the previous section. Thus the issue of what the philosopher takes for granted depends largely on what appears as salient to him, and more importantly, what does not appear as salient, or questionable.

In his work de Sousa argues that affects highlight certain features of experiences (1980, p.136); this might be because it is useful to our projects, or fearful or so on. The point here for de Sousa, is that affects define a field of reference within which we operate or in this case, within which we do philosophy. Affects limit what we have to take into consideration in our inquiries, by making contentious propositions salient (as with doubt), or by more generally evoking a field of what we may comfortably trust, or take for granted (1980, p.136). Involved in this field is an implicit consideration of the stance which we take in doing philosophy. The stance to which we are aligned and in which we work will undoubtedly flavour the nature of our enquiries and further our conclusions<sup>11</sup>. Such stances provide a further orientation in the way which we approach philosophical problems. However, for the purposes of this paper, I shall continue to focus on affective states.

Affective states as salience prescribers therefore have a direct impact on the ways in which we approach philosophy<sup>12</sup>. In highlighting certain elements of our systems or inquiries, affective states perform pre-theoretical roles, which according to the phenomenological psychiatrists mentioned earlier, must be in place before healthy

<sup>11</sup> For further reading on the centrality of stances in philosophical investigation see Van Fraassen (2002) and Ratcliffe (forthcoming).

<sup>12</sup> For arguments surrounding this see Damasio 1995, de Sousa 1980; 1987.

explicit reasoning begins: affective states prepare an arena in prescribing the bounds within which the inquiry may take place, something which, as de Sousa remarks, cannot be performed by reason alone. He writes:

Despite their reputation for motivating bad behaviour, emotions are essentially implicated in our capacity to live a coherent and reasonably well-regulated life. (2004, p.65)

There must be some standard in virtue of which the terms on which we inquire are set. As we may see from Blankenburg's example of the schizophrenic woman who could not constrain a field of reference in which to place her inquiry as to which dress she ought to wear that day (2002, p.309), reason standing apart from any other function cannot readily or immediately recommend a course of action (see also Damasio 1995, p.70). Instead she took into account every small, trivial and irrelevant detail, most of which would not be entertained by someone with healthy reasoning abilities. This, as Blankenburg identifies, is the problem of 'overinclusiveness' (2002, p.307-309). Thus, if we place de Sousa's arguments for affective states as salience providers, and lessons learned from descriptions of certain psychotic illnesses, we find that one 'function of emotions is to fill gaps left by ... "pure reason" in the determination of action and belief' (de Sousa 1980, p.136).

Following from de Sousa's considerations, and the ones we have previously gleaned from phenomenology it seems reasonable to suggest that affective states perform 'behind the scenes work' in philosophical inquiry, forming predilections and spotlighting features of the world which are beneficial, dangerous or somehow worthy of attention. I have so far spoken in broad terms of affective states, but now I shall move on to consider the role which specific affects provide in guiding and colouring philosophical inquiry.

### *Affects as Marginalised Influencers*

As already stated, Hookway (2002, 2003) examines and promotes the methodological roles the feelings of doubt, and following from this, conviction play in philosophical inquiry. He unearths the roles that these two feelings play in both beginning and ending philosophical inquiry. However, in looking only at doubt and conviction as the beginning and end of philosophical inquiry, his account misses out a great deal. Hookway focuses primarily on that which appears salient through doubt, and its resolution through conviction. Just as important as these considerations is an account of the role of more subtle affective states, in the guise of temperamental dispositions which favour one course of explanation over another, or to tend to look at a problem in a certain light. James famously characterises such philosophical allegiances and how they shape an individual's philosophical inquiry in conjunction with his work on feeling in *Pragmatism*<sup>13</sup>. 'Feeling' for James becomes that which forms a background to all our endeavours, producing an individuality of method. As Carroll Izard writes, for James:

[F]eelings are the roots of personality or individuality and the best means of explaining the events and behaviours of everyday life. (1990, p.631)

Although James's discussion is framed in terms of feelings, in alignment with his wider work, I believe that we can take feeling as a subset or type of affective state<sup>14</sup>. The important thing to note about James's work is the frank acknowledgement of the role of affective states as what I call marginalised influencers. That is, affective states dispose us to a certain philosophical outlook or methodology. In the first essay of *Pragmatism* James opens by boldly declaring:

<sup>13</sup> For example see his discussion of the 'tough headed empiricists' and the 'tender minded rationalists' (1992, p.23ff).

<sup>14</sup> See Ratcliffe 2005 for a reinterpretation of William James, emotion and feeling.

The history of philosophy is to a great extent that of a certain clash of human temperaments. Undignified as such a treatment may seem to some of my colleagues, I shall have to take account of this clash and explain a good many of the divergences of philosophers by it. (James 1992, p.22)

This bringing to the fore of the role of temperament may shed a great deal of light upon philosophical methodology. James goes so far as to claim that an appeal to temperament ‘explains a good many of the divergences of philosophers’ and this may be so, but I believe that room is available to appeal to a related concept which draws directly from considerations of the role of temperament; that is what I call affective predilection.

Affective predilection works in a different way to affects in conjunction with salience. Instead of highlighting certain things as relevant or problematic, affective predilection instead disposes us to unreflectively favour one methodology over another, or frames the types of question we are likely to ask. As de Sousa (1980, p.140) writes: ‘emotion can be clearly seen as a disposition to ask certain questions’. Affective states may in this case be seen as that which may break or avoid a deadlock which rational means alone could not. Here, we arrive again at a consideration of a situation which reason may not single-handedly dissolve. A Burden’s ass scenario can easily be imagined involving two philosophical models of exactly equal rational integrity pertaining to examine the same puzzle; the question arises of how you chose your philosophical allegiance when there seems no strictly *rational* reason or consideration to opt for one position over another. An answer which follows from the discussion so far is that our affective orientations predispose us to be more sympathetic to one system or methodology than another. This is a point to which de Sousa is sympathetic when he writes:

The same goes for choices of strategies in the light of existing desires: there are choices that no rational calculation can make, because they are between alternatives that on rational calculation turn out the same. (de Sousa 1980, p.136)<sup>15</sup>

In this way James's discussion emphasises the *sustained* role of affect in philosophy. In doing so it differs from Hookway's, whose exposé of the role of doubt uses a notion of sporadic doubt linked with salience. James thus fills in the gaps of Hookway's account, showing how affects are at work in inquiry in the background; steering the philosopher through the inquiry according to their tacit predilections.

Doubt, for Hookway, is a specific state which reveals certain propositions as candidates for investigation through their questionability, thus rendering them salient (2003, p.93). However, his account focuses primarily on what might be termed as *sporadic* affective states in terms of doubt and the corresponding resolution of doubt, and equally on conviction. These affective states are relatively short term. This forms a contrast with James's account which shows how affects play a *sustained* role in our enquiries. As was emphasised in the first section of this paper, a condition of a healthy reasoning and mind is that we do not constantly doubt, we feel a sense of familiarity in relation to our background environment, enabling anything untoward to become salient.

Hookway's account is lacking in that it only scratches the surface of the impact of affective states on reasoning, focusing only on the sporadic ones. James, on the other hand, in attempting to call attention to something which had gone unnoticed or unspoken of in philosophy, emphasises the manner in which long term or 'consistent' affective dispositions impact on the ways in which we go about doing philosophy.

---

<sup>15</sup> See also: James 1956, p.92: 'Why does [x] believe [y]? Simply because, like every human being of the slightest mental originality, he is peculiarly sensitive to evidence that bears in some one direction. It is utterly hopeless to try to exorcise such sensitiveness by calling it the disturbing subjective factor, and branding it as the root of all evil.'

In this way, his account reveals that which is usually tacit in our philosophical methodologies; that which is, according to traditional standards ‘unphilosophical’ according to the ideals of ‘detached and coolly rational’ inquiry, untainted by base influences of mood, temperament or other such ‘subjective’ things. In uniting Hookway’s account of ‘sporadic’ affects with James’s a more encompassing account of the role which affects play in philosophical inquiry might be arrived at; allowing us to rethink the way in which we have conceived of ‘how we do philosophy’.

### ***Conclusion***

In highlighting the ways in which affects may be said to impact on philosophical inquiry I have worked against a longstanding conception of affects as mere disruptors of successful rational thought. I have shown that affects play an active role in both motivating and moulding the course of our inquiries, and further, that an encompassing account should discuss both of these roles. Throughout this work I have attempted to show how assumptions which hold that philosophical inquiry is carried out apart from the influences of affective states ought to be reviewed. This paper ultimately offered a re-examination of how we have viewed the relationship between philosophy and affect in light of the recent work on emotion; and, in doing so, has called for attention to be given on how affective states inform and prepare us for philosophical investigation.

### ***References***

Blankenburg, W., 2002. First steps toward a psychopathology of ‘common sense’ (Trans. A. L. Mishara). *Philosophy, Psychiatry & Psychology* 8, p.303-315.

Damasio, A. R., 1995. *Descartes’ Error: Emotion, Reason and the Human Brain*. London: Picador.

----- 2000. *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. London: Vintage.

----- 1980. The rationality of emotions. In A. Rorty (ed.) *Explaining Emotions*. London: University of California Press, p.127-153.

----- 1987. *The Rationality of Emotions*. Cambridge: MIT Press.

----- 2004. Emotions: what I know, what I'd like to think I know, and what I'd like to think. In R. C. Solomon (ed.) *Thinking About Feeling*. Oxford: Oxford University Press, p.61-75.

Descartes, R., 1911a. The principles of philosophy. In his *The Philosophical Works of Descartes* (Trans. Haldane, E. S. and Ross, G. R. T.) Cambridge: Cambridge University Press, p.201-302.

----- 1911b. The passions of the soul. In his *The Philosophical Works of Descartes* (Trans. Haldane, E. S. and Ross, G. R. T.) Cambridge: Cambridge University Press, p.329-428.

----- 1998. *Meditations and Other Metaphysical Writings*. (Trans. Clarke, D. M.). London: Penguin.

Heidegger, M., 1962. *Being and Time* (Trans. Macquarrie, J. and Robinson, E.). Oxford: Blackwell.

Hookway, C., 2002. Emotions and epistemic evaluations. In P. Carruthers, S. Stich and M. Siegal (eds.) *The Cognitive Basis of Science*. Cambridge: Cambridge University Press, p.251 – 262.

----- 2003. Affective states and epistemic immediacy. *Metaphilosophy* 34, p.78-96.

Husserl, E., 1989. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy: Second Book*. (Trans. Rojcewicz, R. and Schuwer, A.). Dordrecht: Kluwer.

Izard, C. E., 1990. The substrates and functions of emotion feelings. *Personality and Social Psychology Bulletin* 16, p.626-635.

James, W., 1956. *The Will to Believe and Other Essays in Popular Philosophy*. New York: Dover Publications.

----- 1992. *Pragmatism*. London: Routledge.

Ratcliffe, M., In press. Existential feeling and psychopathology. *Philosophy, Psychiatry & Psychology*.

----- Forthcoming. Stance, Phenomenology and Feeling. *Synthese*.

----- 2005. William James on emotion and intentionality. *International Journal of Philosophical Studies* 13, p.179-202.

Sass, L. A., 1992a. *Madness and Modernism: Insanity in the Light of Modern Art, Literature and Thought*. New York: Basic Books.

----- 1992b. Heidegger, schizophrenia and the ontological difference. *Philosophical Phenomenology* 5, p.109-132.

----- 1994a. *The Paradoxes of Delusion: Wittgenstein, Schreber and the Schizophrenic Mind*. London: Cornell University Press.

----- 1994b., 'Civilised Madness: Self-Consciousness and the Modern Mind'. *History of the Human Sciences* 7, p.83-120.

----- 1999. Schizophrenia, self-consciousness and the modern mind. In Gallagher, S., Shear, J., (eds.) *Models of the Self*. Exeter: Imprint Academic, p.319-343.

----- 2001. Self and world in schizophrenia: three classic approaches. *Philosophy, Phenomenology, Psychology* 8, p.251-270.

----- 2004. Affectivity in schizophrenia: a phenomenological view. In Zahavi, D. (ed.) *Hidden Resources: Classical Perspectives on Subjectivity*. Exeter: Imprint Academic, p.127-147.

Sechehaye, M., 1970. *Autobiography of a Schizophrenic Girl*. New York: Signet.

Stanghellini, G., 2004. *Disembodied Spirits and Deanimated Bodies: The Psychopathology of Common Sense*. Oxford: Oxford University Press.

Solomon, R. C., 1993. *The Passions: Emotions and the Meaning of Life*. Cambridge: Hackett Publishing Co.

----- 2004. Emotions, thoughts, and feelings: emotions as engagements with the world. In R.C. Solomon (ed.) *Thinking About Feeling*. Oxford: Oxford University Press, p.76-88.

Van Fraassen, B., 2002. *The Empirical Stance*. London: Yale University Press.

## **A PROBLEM WITH DEFINING TESTIMONY: INTENTION AND MANIFESTATION:**

MARK NICHOLAS WALES

UNIVERSITY OF ST ANDREWS

### ***Abstract***

Within current epistemological work in the field of testimony it is generally considered usual to accept cases of non-verbal testimony within definitions of testimony (testimony through manifestation); it is also considered usual to accept cases in which the speaker does not intend to testify (non-intentional testimony). In this paper I show that while considered individually these two cases are unproblematic, this is not the case when both are considered as part of a definition of testimony. I suggest a case in which a person testifying non-verbally and non-intentionally is uncomfortably close to a case which we must count as perception; this is undesirable, as we clearly want to separate cases of testimony from cases of perception. I then look at possible solutions to this problem rejecting a number of solutions; I settle on a solution based on Jennifer Lackey's work on testimony. I conclude that we need not worry about the conjunction of non-intentional testimony with manifestational testimony as long as we are careful when we are trying to define testimony.

### ***Introduction***

Epistemological work on the subject of testimony has primarily focussed on how (and if) we can be justified in gaining knowledge from testimony. However, a recent paper by Jennifer Lackey, "The Nature of Testimony" (Lackey 2006), focuses on how we should define testimony. In her paper Lackey argues that certain definitions of testimony cannot work, regardless of issues of justification. In a similar vein, I hope to show a further problem that will occur if we are not careful with how we define testimony.

In this paper I will look at two particular aspects of testimony: a person's meaning to testify (intention) and testifying non-verbally (manifestation). In §1 and §2 I will look at how intention and manifestation are handled in current views on testimony. In §3 I will show that if certain cases of non-intentional manifestation are allowed as cases of testimony we will run into problems; namely, that these cases of testimony seem to be nothing more than perception. In §4 and §5 I will further analyse the ideas of intention and manifestation to see if therein lies a clue towards a possible solution. In §6 I will briefly look at the possibility of defining testimony in a highly restrictive way that allows us to avoid the problem in §3; I will, however, reject this formulation. Finally, in §7 I will look at two solutions that can be found in existing definitions of testimony. I will conclude in §8 that we can avoid the problem from §4, but only if we are careful when defining testimony.

I will be focussing on what Coady calls 'natural' testimony; this is in contrast with 'formal' testimony, which is testimony in a legal setting (Coady 1992, p. 26). Natural testimony is the testimony of everyday life: when someone tells me their name, when a lecturer explains a theory, or when the news-reader on TV tells me the day's events. Testimony has, by certain philosophers, also been thought to include such cases as road-signs, posthumous writings, and nodding.

There are many different views on what testimony is (from hereon in I will use 'testimony' to refer only to natural testimony unless specifically mentioned): these range from the technical definitions offered by Coady (1992), Lackey (2006), and Graham (1997), to the more general concepts used by Fricker (1995), Sosa (1992), and Price (1969). I will now look at what these views have to say about two particular forms of testimony: non-intentional testimony and testimony through manifestation.

*Intention: Current Views*

Non-intentional testimony, when S testifies that p without meaning to do so, is touched upon by many of the views mentioned in the introduction. Let us see why it is an issue. Imagine that Gerry's bike has been stolen and one of the following two situations occurs:

- 1) Peter tells Gerry that his bike has been stolen
- 2) Gerry overhears Peter muttering to himself about how he just stole Gerry's bike.

Now, (1) is unequivocally a case of testimony: Peter is telling Gerry a piece of useful information and Gerry is paying full attention. But what about (2)? Intuition would, I think, suggest that this is not a case of testimony. Firstly, it was not meant to be heard by anyone, and secondly it was not aimed at anyone, least of all Gerry. The one thing that seems to be different between these two cases is that in (2) Peter has no intention that what he has said should be heard. On this subject Lackey says: "testimony is often thought of as an intentional activity on the part of the speaker... [that] a speaker cannot testify unless there is some intention on the part of the speaker to convey information" (Lackey 2006, p. 187). But there are times when non-intentional testimony does seem to be possible. Consider the following two situations:

- 3) For most of his life Albert kept a diary of everything he did, including a time when he cheated on his wife. The diary was kept in a safe and was never meant to be read by anyone other than Albert. But Albert died a number of months ago. His wife has decided to go through his belongings to check for anything important before she gets rid of them. She hires a lock-smith to open the safe and reads Albert's diary. She is horrified to find out that Albert cheated on her. Albert had no intention of his wife ever reading his diary, and yet this still seems to be Albert testifying (through the diary) that he cheated on his wife.

4) Consider the case of Hannah who by accident sends an e-mail not to the intended party, her best friend Jenny, but to Katie, who is the unfortunate subject of the e-mail. Now in this case it would intuitively seem that Hannah has testified her feelings about Katie to Katie. But there was no intention to do this: it is the opposite of what was intended. Again we seem to have a case of non-intentional testimony.

What do the main views on testimony have to say about such cases? Well, Fricker's view of testimony as "tellings generally" (Fricker 1995, p. 397) would seem to allow for this: it would include all communications of information, whether they were intentional or not. Sosa's view is quite general also, whereby testifying is simply stating "one's belief that p" (Sosa 1991, p.219). Sosa specifically mentions that his definition is "a broad sense of testimony that counts posthumous publications as an example" (Sosa 1991, p.219). Lackey's view is also open to the idea of non-intentional testimony: "...testimony does not depend on the intentions of the speaker but, rather, on the needs of the hearer" (Lackey 2006, p. 187). The stricter definitions of testimony offered to us by Coady (Coady 1992, p. 42) and Graham (Graham 1997, p.227) do not seem as open to the idea of non-intentional testimony: they both heavily rely on the speaker meaning to convey some piece of evidence that p. However, Coady does touch upon the matter: he says that journals and the like are "not really very far from what we have been treating as paradigm testifying since the only difference lies in the original speaker's or writer's ignorance that he will be testifying to us" (Coady, 1992: p. 50). So non-intentional testimony is widely accepted within the main views on testimony.

### ***Manifestation: Current Views***

Manifestation, when S testifies that p non-verbally, is another element of testimony that must be considered. Let us consider another two examples:

5) Shaughan walks into the sitting room and asks James if there are any cookies left. James, whose mouth is full of cookies and is unable to speak, nods his head and points towards a cupboard in the kitchen.

6) Pete is giving auditions for his choir. He wants to know which voice part each individual is; he asks each person to sing a song to him instead of telling him what voice part they are. Stewart, a bass, sings to Pete in a very low voice and Pete takes from this that Stewart is a bass. Kat, a soprano, sings to Pete in a very high voice and Pete takes from this that Kat is a soprano.

Again, both of these seem to be cases of testimony. In (5) James is telling Shaughan that “Yes, there are some cookies left, and they’re in the cupboard over there”. If this had been spoken it would certainly count as testimony. In (6) Stewart seems to be testifying “I’m a bass” and Kat seems to be testifying “I’m a soprano”. Although in this case the testifying is not quite as direct as in (5), we can still class it as answering a question with information, and so it again seems to count as testimony.

The first of these examples is based on one given by Lackey (Lackey 2006, p. 186), so it is not surprising that Lackey supports the view that you can testify through manifestation. However, it seems that Lackey thinks she is unique among the views in holding this position: “...notice that, in one form or another, all of the proposals [mentioned in the previous sections] flesh out the notion of testimony in terms of statements” (Lackey 2006, p. 185). But I do not think that this is the case.

Consider Fricker’s “tellings generally” comment. ‘Telling’ is not necessarily verbal, as in “The smile on her face told him everything”; and I think it would be quite normal for someone to say that, in (5), James told Shaughan where the cookies were and that, in (6) Stewart told Pete that he was a bass. It is also interesting to note that Coady talks about ‘institutional testimony’, which can include “road signs, maps, the measurement markings on rulers...” (Coady 1992, p. 51). Although most

of these are just forms of written testimony, road signs are an interesting case. Many road-signs cannot be fully understood just from what is written on them: a sign pointing towards Aberdeen would be of no use unless the person reading it also gets the directional information that is offered. But this directional information comes not from written testimony (the reading of 'Aberdeen') but from manifestational testimony (that the sign is pointing in some particular direction). So although it is not explicit in Coady's account of testimony I can see no reason why he would deny that testimony is possible through manifestation. I also think that cases (5) and (6) are convincing enough to sway most philosophers into thinking that you can testify through manifestation.

So far I have shown that both non-intentional testimony and testimony through manifestation should be considered under a plausible account of testimony. However, I hope to show in the next section that if we accept both at the same time we will run into problems.

### ***Intention And Manifestation: A Problem***

Consider the following example. Russell is both a talented musician and one of the world's leading experts on cars.

7) He is walking along the street and he overhears someone singing in their sitting room. He can tell that the singer is a soprano, and so he comes to rightly believe that the woman who lives at No. 52 is a soprano.

8) He continues to walk along the street when he hears a car approaching from behind. Before the car comes into view he can hear its engine. He can tell from this that the car that is about to pass is a 1953 Ford Prefect, and so he comes to rightly believe that the car that is about to pass is a 1953 Ford Prefect.

In (7) we seem to have a legitimate case of non-intentional manifestational

testimony: the woman at No. 52 is doing nothing differently from the singers in (6), other than that she has no intention to testify. But, as we saw with cases (3) and (4), having no intention to testify should not be a significant reason to think that (7) is not a case of testimony.

And yet (7) seems almost identical to (8). But we would surely not want to say that the car has testified; (8), in fact, seems to be nothing more than perception: Russell hears the car and combines it with his knowledge of cars to conclude that it is a 1953 Ford Prefect. This seems to be no different from any case of perception where we combine something we perceive with some background knowledge to form a new belief: for example, seeing a yellow door and coming to believe that it is yellow. So the argument (which from now on I will call the ‘non-intentional manifestation problem’ – or N-IMP) is as follows:

(7) is a legitimate case of testimony: it is simply cases (3) and (6) combined which were both valid individually.

If (7) is a legitimate case of testimony, so is (8).

Therefore, (8) is a legitimate case of testimony.

But (8) is a case of perception.

Therefore, (7), and any other case of non-intentional manifestational testimony, is a case of perception.

Clearly this is problematic. I take it that even if we are reductionists about testimony we would not want a case of testimony to be case of perception simpliciter; we would want it to be somehow reducible to perception. Take a standard instance of testimony:

9) I tell my friend that wildcats live in trees.

Although my friend has heard something, namely my uttering that wildcats live in trees, and formed a belief based on this, that wildcats do indeed live in trees, it

seems very different from (8). The information that my friend gains in (9) seems to come from hearing and understanding what has been said by someone else, whereas in (8) it seems to come from hearing alone.

I plan to use the remainder of this essay to try and find a way to avoid the N-IMP argument. There are a number of ways that this can be done:

Deny cases of non-intentional testimony. [Deny cases (3) and (4)]

Deny cases of manifestational testimony. [Deny cases (5) and (6)]

Deny the specific combination of non-intentional and manifestational testimony.

Deny that (8) is analogous to (7). [Denying premise II of the N-IMP argument]

Deny that (7) is a case of testimony. [Denying premise I of the N-IMP argument]

I will cover these in order: §4 will focus on non-intentional testimony; then in §5 I will look at manifestational testimony in more detail; having shown in §4 and §5 that we cannot reject either view completely, in §6 I will see if we can reject the combination of the two views; finally in §7 I will look at two solutions that can be found in the definitions of testimony given by Coady and Lackey, that both use strategy IV above. Finally I will conclude that strategy V, denying that (7) is a case of testimony, which I think seems intuitively most plausible, is the best way to avoid the problem.

### ***Intention: Further Analysis***

So, can we completely reject non-intentional testimony? If we could it would stop the N-IMP argument before it could even get started. If we consider cases of formal testimony, it is not unheard of to use journals and diaries as testimonial evidence, even if the person who wrote it had no desire for anyone else to ever read it. If formal testimony, the most narrow and rigidly defined form of testimony, accepts non-intentional testimony, then surely we must accept such cases in natural testimony? But perhaps it is not the case that everything formal testimony allows should be allowed under natural testimony too, so we will need another reason.

Let us look at (3) and (4) again:

3) For most of his life Albert kept a diary of everything he did, including a time when he cheated on his wife. The diary was kept in a safe and was never meant to be read by anyone other than Albert. But Albert died a number of months ago. His wife has decided to go through his belongings to check for anything important before she gets rid of them. She hires a lock-smith to open the safe and reads Albert's diary. She is horrified to find out that Albert cheated on her. Albert had no intention of his wife ever reading his diary, and yet this still seems to be Albert testifying (through the diary) that he cheated on his wife.

Can we really deny that these are cases of testimony? I do not think that we can without denying our basic intuitions of what testimony should be. In these cases someone, X, is gaining useful information as a result of something that has been said by someone else, Y. The fact that Y had no intention that this information should get to X seems unimportant in these situations. This suggests that the person who is being testified to is not necessarily determined by the person testifying.

So, with this in mind, perhaps the problem lies just in how we define "non-intentional". One suggestion might be that the person testifying in non-intentional cases is in fact intending to testify: they intend to testify to themselves.

Let us consider Albert in (3) again. Although he does not intend to testify to his wife, Albert is perhaps writing down his thoughts so that sometime in the future he can read them back to himself to see what he used to think and feel: his intention is to testify to himself later in life. So Albert does not intend to testify to his wife, but he does intend to testify to himself. Thus, Albert's testimony is not non-intentional at all, and therefore we can avoid allowing non-intentional testimony into our definition of testimony. This would mean that the N-IMP argument cannot be formulated in the first place.

But I do not think that this suggestion will work in all (or even most) cases. While some people may write diaries to read back to themselves later in life it is likely that many people write diaries for different reasons: perhaps just for therapeutic purposes; getting their thoughts out to see what they really think. In cases such as this it does not intuitively seem like the content of what is being written is aimed at anyone – even the author. Because this possibility exists, the idea that all cases of seemingly non-intentional testimony are really just instances where someone has written something down for their future self seems unlikely.

***Manifestation: Further Analysis***

Perhaps, then, the problem is with manifestation. Maybe we should just reject the idea of testimony through manifestation completely: just reject the idea that we can testify non-verbally. What would be lost if we were to do this? Firstly, we would lose the cases like those mentioned in (5) and (6), and I think this would be a loss to an account of testimony: someone nodding to say “yes” or pointing to say where something is seems to be desirable in a definition of testimony. But more importantly, we might lose things like sign-language if we were to reject manifestational testimony. This definitely seems like a bad result: it is not unusual for sign-language to be used in cases of formal testimony. And, if we understand sign-language, we can have very informative conversations with someone using it: but sign-language is certainly non-verbal.

Manifestation can also have an essential part to play in understanding some testimony. For example body language can be very important when trying to work out what someone means. And directional road-signs, as mentioned in §2, cannot be fully understood without some manifestational quality. I do not, therefore, think we can reject all cases of manifestational testimony.

Maybe, then, there is something about the particular cases in (7) or (8) that make them somehow different from the other cases of manifestation. One suggestion might be that normally someone is responding to a question through manifestation, but that this is not the case in (7) or (8). In (5) James is answering Shaughan's question and in (6) the singers are responding to Pete's request. If this were the case we could reject cases (7) and (8), as neither are responses to questions, and so N-IMP would fail as an argument.

But notice that this suggestion would have to be limited to manifestational testimony. If we were to restrict all instances of testimony to cases where a question is being answered, we would not be able to count lectures as testimony, many everyday conversations, or cases such as (3) where someone is reading a journal. But these are often the situations that are most likely to give us knowledge. This means that the 'answering a question' suggestion must apply only to manifestational testimony. But this seems like an ad hoc requirement, so I will not consider it any further.

A final attempt might be to point out that cases (5) and (6) are very different. Singing to testify what voice part you are is not really the same as nodding and pointing. If we count (5) as a case of testimony but not (6) then we could avoid the N-IMP argument simply by saying that the woman singing in (7) is not testifying through manifestation. While I do think that the answer to the N-IMP argument lies in denying (7) as well as (8), I do not think that we can do this by rejecting (6). Perhaps I am wrong, but it seems to me that (6) should be counted as a case of testimony: the singer is answering a question in some way that gives the information that was needed – we would need a detailed account of why this does not count as a case of testimony given that it fits these conditions.

***Intention And Manifestation: Rejecting The Synthesis***

So we do not want to reject all cases of non-intentional testimony and neither do we wish to reject all cases of manifestational testimony: there do seem to be cases when each is desirable. Maybe, then, we can just reject the synthesis of the two by adding something like the following to our definition of testimony :

T': It is only a case of testimony when it is either non-intentional or manifestational, but not when both are combined.

This would avoid the N-IMP argument, by saying that (7) is not a case of testimony, so neither is (8), because (7) combines the two ideas, which is not allowed. But, much like the 'responding to a question' suggestion in §5, this seems to be a totally ad hoc solution which has only one purpose: to avoid the N-IMP argument.

***Intention And Manifestation: A Solution***

So we cannot reject non-intentional testimony or manifestational testimony, and we cannot justifiably say that it can be one or the other but not both. This leaves two options: we can deny that (7) and (8) are analogous, or we can deny that (7) is a case of testimony. I will now look at two solutions that can be found in the literature on testimony that may help us to deny that (7) is analogous to (8).

The first possible solution is hidden away in a comment made by Coady: he thinks that in cases of 'institutional testimony' (the use of road signs or maps) it is obvious that what makes these cases different from others is that the signs and maps were made by other people (Coady 1992, p. 51). So perhaps what makes cases of testimony different from Ford Prefects and yellow doors is that the information that they offer is provided by other people (albeit indirectly). So, let us add a new proposition to our definition of testimony:

T+: And it is only a case of testimony when the information that is provided comes from other rational agents.<sup>1</sup>

However, this new definition is not sufficient for our purposes. For example, consider Bob who is very angry: it is clear to anyone who sees Bob that he is angry. Yet, even though Bob is angry and anyone looking at him can see that he is angry, this does not seem like a case of testimony, unlike one of our standard cases such as (6). So even when the information that is being provided comes from another person (or other rational agent) we do not always face a case of testimony; this means T+ will not do as an addition to our definition of testimony.

The second possible solution can be found in Lackey's paper "The Nature of Testimony" (Lackey 2006). In this paper Lackey offers her definition of testimony:

"S testifies that p by making an act of communication a if and only if (in part) in virtue of a's communicable content, (1) S reasonably intends to convey the information that p, OR (2) a is reasonably taken as conveying the information that p." (Lackey 2006, p.193)

The important feature of Lackey's definition of testimony is that she splits testimony into two kinds: speaker testimony and hearer testimony. However, for our purposes, it is another part of her definition that is important; it is her focus on the idea of an "act of communication":

"In my proposal, I shall focus on the notions of an act of communication, a, conveying the information that p. ...I am construing the concept of an act of communication broadly so that it does not require that the speaker intend to communicate to others; instead, it requires merely that the speaker intend to express communicable content." [Emphasis Lackey's] (Lackey 2006, p.187)

---

<sup>1</sup> This would be preferable to 'people', which would rule out possible exceptions such as intelligent aliens, etc. Another option here might be 'beings capable of holding beliefs'.

Lackey makes an important distinction between the speaker's intent to communicate to others and the speaker's intent to express communicable content. Consider (3) again: Albert's diary seems to testify about his cheating even though he had no intention that anyone should read it, hence it seems to be a case of non-intentional testimony. But if we use Lackey's distinction we can say that there was no intention to testify to anyone else, but there was intention to express communicable content (the words in English); the fact that later on someone reads this and takes it as testimony is due to the reader's interpretation of the material as testimony, and not the intentions of Albert.

This distinction might allow us to avoid the N-IMP argument. In (7) the soprano intends to make an act of communication, but does not intend for it to be taken as testimony. This can still count as testimony in Lackey's account. However, the car has no intentions at all, so it cannot even intend to express communicable content. Hence, case (7) is a case of testimony, but (8) is not. N-IMP seems to be solved.

However, to me this solution is still lacking something. Intuitively I feel that the best solution is to deny that (7) is a case of testimony. Even though (7) was, I think, legitimately constructed from cases that we definitely do want to count as testimony, I think there is something very unusual about it being accepted. I do not want to deny that (6) is a case of testimony, as, to me, it seems to be legitimate. But (7) comes quite naturally from it when taken with cases like (3) or (4). So how might we avoid this?

A first attempt might be to say that the woman singing in (7) is not even intending to make an act of communication. It is quite possible that she might have spontaneously burst into song with no intention of communicating anything. This is unlike the singers in (6) who are singing with the deliberate intention of showing what voice part they are. But I do not think this will work; after all it is only a possibility that

she has no intentions to communicate. It is also quite possible that she does. Maybe she thinks that she is communicating “Look at me, I’m such a brilliant soprano”. In this case we could, using the Lackey fix, still take this to be a case of testimony.

Or could we? There is one other important part of Lackey’s distinction which has remained unnoticed . Her definition says: “[...] OR (2) a is reasonably taken as conveying the information that p.” (Lackey, 2006: p.193). The important part here is the use of ‘reasonably’. On this Lackey says, “Roughly, reasonably taking a as conveying the information that p requires that a normal hearer in similar circumstances would take a as conveying the information that p..” (Lackey 2006, p.190).

Now, going back to (7) it is now clear why we should not count it as testimony. Russell is both a talented musician and a leading expert on cars. He can hardly be called a ‘normal hearer’. So (7) no longer counts as a case of testimony and we can finally legitimately avoid the N-IMP argument. By using some aspects of Lackey’s definition of testimony we can now say that non-intentional manifestational testimony cannot be reduced to perception.

### ***Conclusion***

To conclude then, if we do not define testimony carefully we run the risk of having cases where testimony is not just reducible to perception but actually is perception. This is not desirable to even the reductionist on testimony. We can, however, avoid this problem if we are careful in our definition of “non-intentional”: if we use Lackey’s distinction between the speaker’s intent to communicate with others and the speaker’s intent to express communicable content, as well as her idea of a ‘normal hearer’ no problem arises. Although we should, I think, focus on the more important issue of justification in testimony, we must always be aware that our definition of testimony can be important and should not simply be ignored.

***References***

Coady, C. A. J., 1992, *Testimony*, Clarendon Press Oxford, pp. 3 - 75

Fricker, E., 1995, "Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony", *Mind* 104, pp. 393 - 411

Graham, P. J., 1997, "What is Testimony", *The Philosophical Quarterly* 47, pp.227 - 232

Lackey, J., 2006, "The Nature of Testimony", *Pacific Philosophical Quarterly* 87, pp. 177 - 197

Price, H. H., 1969, *Belief*, George Allen & Unwin Ltd: Humanities Press, pp. 112 - 129

Sosa, E., 1991, *Knowledge in Perspective*, Cambridge University Press, pp. 215 - 222

***Acknowledgements***

Thanks to Martijn Blaauw (Free University of Amsterdam), Vasco Castela (University of Manchester), and to the Praxis referees for advice on this paper.