EUROPEAN JOURNAL
FOR PHILOSOPHY OF RELIGION

THE THEME OF THE ISSUE

DIVINE MOTIVATION VS. HUMAN AUTONOMY?

GUEST EDITORS

GODEHARD BRÜNTRUP & LUDWIG JASKOLLA

Munich School of Philosophy

OTHER ARTICLES

DISCUSSIONS AND REPLIES

EDITORIAL

The special focus of the papers in this issue is 'Divine Motivation vs. Human Autonomy?' They engage with the perennial debate on autonomy and heteronomy in ethics. This dispute has been experiencing a renaissance due to the works of analytic philosophers of religion such as Robert M. Adams and Linda Zagzebski, who purport to ground morality in God, thus making it heteronomous to the human mind.

The central tenets of the two opposing positions are clearly illustrated in the first paper 'Human Autonomy and Theological Ethics' by Robert M. Adams. He distinguishes an individualist (Kantian) outlook from a social (Rawlsian) perspective, and argues that the social character of moral knowledge necessarily implies some version of heteronomy.

The *individualistic account* lies at the very heart of enlightenment moral philosophy, and has as such shaped the intellectual landscape of European continental thought. Kant argues that autonomy consists in a certain form of self-legislation. The Kantian usage itself in fact stems directly from the etymology of the word: 'autos' and 'nomos', which, when combined, translate to 'a law, which is prescribed by the individual itself'.

The idea of the autonomy of the moral self is paramount to the Kantian outlook. By means of pure practical reason – independent from others and grounded only in one's own reasoning – human beings legislate norms *for* themselves. This Kantian outlook achieves universality by claiming that pure practical reason will lead all human beings to legislate the same norms.

In this sense the Kantian account truly stresses the autonomy of human beings in legislating moral norms: pure practical reason is detached from constraining, external factors, and rather is grounded solely in the individual itself.

In many respects the *social account* can be described as the antithesis to the Kantian account. This might come as a surprise because John Rawls is seen by many as a moral philosopher of Kantian heritage. Rawls claims that in a hypothetical situation where nobody knows her own place in a future society, people would choose the maxims which would maximize the prospects of the least well-off in the society.

Yet, as Adams points out, the pure practical reasoning in the Rawlsian 'original position' is merely hypothetical, a heuristic construct. In reality, we are embedded in various social relationships and biographical constraints which dictate the content of our moral thought. Morality,

we might claim following Rawls and Adams, originates within a second-personal relationship – a relationship between 'me' and 'you'.

This social outlook thus seems to threaten Kantian autonomy: The essentially social nature of moral norms renders adherence to a moral authority, as opposed to pure, individual reasoning, inevitable. In a final analysis, this might lead to the idea that norms are not arrived at by practical reason, but are rather discovered as entities grounded in something which is greater than the determinations of the human mind: the Platonic idea of the Good or the Christian God.

Against this background of moral deliberation as inherently embedded in social practices and as something encountered as being intrinsically independent from us, we are able to reformulate the philosophical question which lies at the heart of this set of texts: Can there be a substantial sense in which human beings are autonomous, while still acknowledging the intrinsically social, and thus heteronomous, character of moral knowledge?

Most of the papers included in this special focus issue elaborate on the idea that there is indeed a substantial sense in which human beings remain autonomous, even if the 'original position' is merely an idealized construct. Autonomy then no longer consists in the pure and independent self-determination of an individual endowed with pure practical reason; it will rather be a second-order stance, which is exercised in the capacity of critically evaluating the moral norms experienced in relation with others.

In quite different ways, the various authors explore this middle-ground between autonomy and heteronomy. Linda Zagzebski's paper 'Divine Motivation Theory and Exemplarism', which concludes this special focus issue, develops this very idea from the assumptions of exemplarist virtue theory. Assimilating oneself to an exemplar of human virtue does not preclude human autonomy, but is indeed the very origin of moral judgment within social relations: 'We should investigate the judgment of exemplars, and that is revealed in narratives about them and other forms of observation of them.' (Zagzebski *this issue*)

Godehard Brüntrup & Ludwig Jaskolla
Munich School of Philosophy

HUMAN AUTONOMY AND THEOLOGICAL ETHICS

ROBERT M. ADAMS

Rutgers University

Abstract. It is argued here that we have good reason to aspire to be autonomous in certain ways that deserve a place in the theory of virtue, but not in some of the ways that have figured most prominently in theories of moral obligation. This grounds an argument that the sorts of autonomy to which we have reason to aspire need not be enemies of theological ethics. The focus is on the relation of autonomy to obligation in sections 1-4, and on the relation of autonomy to love in section 5.

Can theological ethics be reconciled with a proper valuation of human autonomy? To answer that question we must address a prior question: should we indeed aspire to be autonomous? If so, in what ways? And why?

I. A KANTIAN VIEW REJECTED

I begin by sketching a Kantian metaethical view that I reject. I will call it 'the Kantian view', believing that what I say about it is suggested by passages in the 'critical' works of the last decades of Kant's life. Whether it was exactly Kant's own view, it is a view inspired by Kant that undoubtedly has had, and continues to have, a great deal of influence. I think it worthwhile to think critically about that view, without getting involved in the question whether it was exactly Kant's own view or not.

According to the Kantian view, a good will, which we can have as well as God, is the sole determinant of moral obligation – and, some Kantians would add, the sole ground of any sort of real value. It is a main point of this view that in having a good will, a person is autonomous. One requirement for the relevant sort of autonomy is that the moral law

which one obeys in having a good will is a law that one legislates for oneself. Nevertheless it is also part of the Kantian view that all rational agents legislate the same moral law for themselves. That is because a second requirement for Kantian autonomy is that one is led by pure practical reason alone to legislate that law for oneself, and Kant supposes that pure practical reason leads all of us to the same moral law. The point of speaking of 'pure practical reason alone' in this context is that one is not to be led by faculties or factors external to one's own reason. If one were led by emotion, or desire for an extrinsic end, or by the opinions or desires or commands of other persons, to that extent, according to the Kantian view, one would not be autonomous but heteronomous.

That we legislate the moral law for ourselves, and are led to do so by our own pure practical reason, might also be called 'the Kantian story'. It is an odd story, representing as an exercise of my freedom – for Kant the most important exercise of my freedom – an act by which I obviously limit my freedom, subjecting it to a host of obligations. One might think that our individual autonomy is in tension, not only with theological ethics, but with any morality that places us under obligations.

That tension is not removed by the idea that our own practical reason directs us to accept the obligations. Even if we agree that it is always wisest (perhaps even by definition wisest) to follow the directions of practical reason, the wildest dreams of freedom may whisper in our ear that it would be nice if we could afford sometimes to do something really stupid. The tension between autonomy and obligation is still there. It may not be any harder to come to terms with it on the assumption that the demands of moral obligation come to us from other people and/or from God, if we value (and have reason to value) our social relations with other people and with God at least as highly as we value our own practical reason. That is the heart of my view of the relation of theology to moral obligation.

I do not believe the Kantian story. I will explain why I don't – first with reference to the idea of pure practical reason, and then, at somewhat greater length, to the idea of giving oneself the moral law.

(1) I am not aware of having a Kantian faculty of pure practical reason. This is not to say that I am not aware of choosing and acting for reasons, of reasoning about what would be good or bad, or right or wrong to do, and of choosing, at least sometimes, on the basis of what seem to me the best reasons. To that extent, I agree, I have a faculty of practical reason. But not a faculty of pure practical reason of the Kantian kind. That is,

I do not have a faculty of practical reason that can operate fruitfully without being importantly influenced by my own emotions and my own desires regarding possible consequences of my action. I also do not have a faculty of practical reason that can operate fruitfully without being importantly influenced by my awareness of the beliefs, emotions, and desires of other persons.

In my moral thinking I find that I must rely to a considerable extent on what my socialization as an analytic philosopher has led me to call 'moral intuitions'. These are beliefs, or in many cases inclinations or dispositions to believe, which are not reached by logical inference from other beliefs. They are not in that way products of reasoning, though they often serve as starting points for reasoning. They also are often targets of critical reasoning, for I do not suppose that I should embrace them uncritically. But even after they have survived critical scrutiny, they are not exactly products of reasoning. In their development and staying power they are sensitive to many features of my state of mind and my context, notably including my feelings and desires, and those that I perceive other people as having. I would trust them less rather than more, if I believed they were not sensitive to my own feelings and desires, and to those of other people.

(2) In addition to not being aware of having a Kantian faculty of pure practical reason, I have never experienced the moral law as a law that I give myself. In my first encounters with moral imperatives, undoubtedly, I experienced them as commands of my parents, and sometimes of other adults. This was not enough, by itself, to give me a conception of moral obligation, or moral rightness and wrongness. I did not have that until I saw a difference between something that was forbidden just because my parents did not want me to do it, and something that was wrong whether or not my parents actually commanded me not to do it. Not that it was up to me whether such a thing was right or wrong – rather it was wrong in such a way that it was not even up to my parents whether it was right or wrong. And I did not figure out for myself that there was such a difference; my parents made it clear to me that they recognized such a difference, and that they believed there were things that would be wrong for them to do – indeed, that would be wrong for any human being to do. In this way I encountered the moral law as a law that is binding on us in a way that we cannot achieve by legislating it for ourselves.

Obligation is distinct from any commitment we may have to fulfil the obligation. A commitment to obey the moral law I can (and even must)

give myself; the obligation to obey the moral law I cannot give myself. The moral law, and obligation, as such, meet me as a demand on me, a demand that comes to me from outside me. I may, and I in fact do, interpret the moral demand, as my parents did, as a demand from God; and to some extent I experience it as a demand from God. As such it may stand opposed to demands from fellow humans. And the plausibility of interpreting it as a demand from God has as one of its supports the more general fact that I experience the moral imperative as something I cannot give myself but must encounter as a demand of another person or persons.

My voluntary choices, to be sure, can and do contribute to shaping the particulars of my moral obligations; but it is not in general simply up to me what those particulars will be. I can give myself an obligation by making a promise or volunteering for a responsibility. In most such cases, however, there is some dependence of my obligation on the will of one or more other persons. If others do not acquiesce in my having the responsibility, my volunteering probably will not result in my acquiring an obligation. And in typical cases of obligations arising from promises, the person to whom I made the promise can voluntarily release me from the obligation, but also has a right to hold me to it.

II. THE SOCIAL NATURE OF OBLIGATION

That obligation meets me in demands of another person or persons reflects the social nature of obligation. Facts of obligation are normative facts, certainly; but they are at the same time social facts, facts about relations between persons. Obligations are obligations to someone – to someone else, someone other than oneself. Talk of 'obligations to oneself' I regard as metaphorical. It may in some contexts be an illuminating metaphor, but the supposed obligations to oneself are not binding in the same way as real obligations are.

If I have a real obligation, there is some other person or group of persons to whom I owe the fulfilment of that obligation. If the social fact of obligation is clear, they and those who sympathize with them require me to fulfil the obligation. They may or may not express the requirement as an explicit demand. In many cases, the requirement may be left implicit, as an expectation understood by all who have been initiated into the social context. The expectation may be manifested in such facts

as that if I fail to fulfil the obligation, without a good excuse, those to whom I owe it will have reason to resent the non-fulfilment, and may be angry at me.

The social character of the morality of duty, and of its grip on us, finds more acknowledgement than one might expect in the work of one of the most influential and important Kantian moralists of the twentieth century, John Rawls. One of his best-known ideas is that of an imagined 'original position', defined by certain conditions of knowledge and ignorance. Rawls argues that the correct principles of justice that would govern an ideal political society or state are those that any one of us, in the original position, would have most reason to choose to govern a society in which we would live our whole lives. This is obviously reminiscent of the Kantian story in which each of us, led by our pure practical reason, legislates for ourselves the moral law. But Rawls' view is also different from Kant's in instructive ways.

In both stories we are meant to identify with a single individual being led by reason to adopt, and agree to be bound by, a set of principles that should govern us all in our relations with each other. That is the striking point of similarity between the stories. A deeper but less obvious similarity is that for both Kant and Rawls the principles that the individual is to legislate are meant to enhance the freedom of the individual. One's freedom is limited by one's own obligations under the principles, but is more importantly enlarged by being protected by the obligations assigned to others under the principles. According to 'the universal principle of right' in Kant's *Rechtslehre*, 'Any action is right ... if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law.'[1] Similarly, Rawls' first principle of justice, to which he assigns absolute priority for the ordering of a just society, is that 'each person is to have an equal right to the most extensive scheme of equal basic liberties compatible with a similar scheme of liberties for others.'[2] The protection of free choice, thus emphasized by both authors, is an important affirmation of a form of individual autonomy.

---

[1] Immanuel Kant, *The Metaphysics of Morals*, translated by Mary J. Gregor, in *The Cambridge Edition of the Works of Immanuel Kant*, the volume on *Practical Philosophy* (Cambridge: Cambridge University Press, 1996), p. 387 (Ak 6:230).

[2] John Rawls, *A Theory of Justice*, revised edition (Cambridge, Massachusetts: Harvard University Press, 1999), p 53. This formulation is a little more precise than that in the 1971 first edition.

The differences between Kant and Rawls, however, are many and important. Most obviously, the Kantian story is supposed to be true – about something that really happens, perhaps in our experience in time, and certainly in the noumenal realm above time. Rawls' original position, on the other hand, is acknowledged to be an impossible fiction that could not be true of any human being – though it may be a useful exercise to imagine ourselves choosing moral or political principles under the constraints imposed in the fiction. Connected with this literary difference is a substantive difference regarding moral psychology. The Kantian story is a central piece of Kant's seriously asserted moral psychology. But when Rawls proposes to build a moral psychology to assess the likelihood that his ideally just political society would be stable, he has to leave the original position. He proposes 'three psychological laws', about human behaviour that he holds would actually tend to occur in certain social and political circumstances. They are generalizations, of considerable plausibility in my opinion as well as his, and empirically testable in principle, though he does not present actual empirical data in support of them.[3]

According to Rawls' first law, if 'family institutions are just, and ... the parents love the child and manifestly express their love by caring for his good, then the child ... comes to love them.' His second law hypothesizes a person who has acquired attachments and a realized capacity for fellow feeling as predicted by the first law, and is living under social arrangements that are 'publicly known by all to be just'. It predicts that 'this person develops ties of friendly feeling and trust toward others in the association, as they with evident intention comply with their duties and obligations, and live up to the ideals of their station'. Rawls' third law hypothesizes a person who has lived in the favourable conditions, and developed the personal attachments and capacity for fellow feeling that are envisaged in the first two laws. It predicts that 'this person acquires the corresponding sense of justice as he recognizes that he and those for whom he cares are ... beneficiaries of these arrangements'.[4]

---

[3] In footnotes Rawls (*A Theory of Justice*, revised edition, pp. 402-11, especially p. 403*n*6) cites a wide variety of studies, both armchair and empirical, by philosophers, social psychologists, and educational theorists (notably including Jean Piaget and Lawrence Kohlberg). But so far as I am aware, the empirical literature on the subject does not provide decisive confirmation of any such laws.

[4] John Rawls, *A Theory of Justice* (Cambridge, Massachusetts: Harvard University Press, 1971), pp 490-91. I quote this passage from the first edition.

These 'laws' do not describe an individual arriving at a moral stance by a wholly self-generated exercise of pure practical reason and voluntary choice. Rather, they describe a process of moral development, conditioned by social relationships, and involving affective as well as cognitive responses to them. The individual is described as acquiring not a theory, but a sense, of justice. His (or her) willingness to comply with demands of justice, and positively to support just arrangements, is not seen as unconditional commitment to obey a categorical imperative, regardless of consequences, and of whether other people obey it too. Indeed Rawls rejects 'the doctrine of the purely conscientious act ..., that the highest moral motive is the desire to do what is right and just simply because it is right and just'. He rejects as well the view that 'the desire to do what is right because doing this increases human happiness, or because it tends to promote equality, ... [is] less morally worthy'. He regards such doctrines as irrational.[5]

In Rawls' view, rather, the individual's willing support for just arrangements and their requirements is seen as socially motivated. It is motivated by 'love' and 'friendly feeling' for other people who love him or her, or have at least proved themselves reliably benevolent toward the individual. It is also seen as motivated by a perception that the just social and political arrangements have beneficial consequences for the individual 'and those for whom he [or she] cares'. As Rawls states, his 'account of the development of morality supposes that affection for particular persons plays an essential part in the acquisition of morality.' He says the question, 'how far these attitudes are [still] required for ... moral motivation [at a later stage] can be left open', but adds that he would find it 'surprising if these attachments were not to some degree [still] necessary'.[6]

Rawls lays primary emphasis on his belief that persons growing up in loving families and loyally friendly associations in a substantially fair and just society would tend to value very highly the principles of justice prevailing in that society, for the sake of their role in structuring fair and beneficial social relations. The firm will of such people to uphold those principles and comply with their demands would be a crucial support for the stability of the just society that Rawls envisages. I will add that such

---

[5] Rawls, *A Theory of Justice*, revised edition, p. 418.
[6] Rawls, *A Theory of Justice*, revised edition, p. 426.

positive valuation of ethical principles can also play a part in constituting a morally important form of individual autonomy.

But before moving on to that point we should take note of a less autonomous, and less warm and sunny, aspect of moral motivation – namely, the awareness of moral obligation as a demand or requirement laid on us by other persons. That there are certain sorts of behaviour that other people will not tolerate is something that we have known since early childhood – something that is involved in knowing, for instance, what a police officer is. And we still know that. None of us, probably, should be too sure that our behaviour would not be affected for the worse if we did not know it. It is a background aspect of our lived moral awareness, our awareness of facts of obligation as social facts.

This sterner sort of moral awareness is not showcased in Rawls' three laws, but it is by no means absent from his whole account of the 'moral sentiments' and their development. In relation to the stages represented by the first and second laws, in particular, he states that 'parental norms are experienced as constraints', and that at the second stage 'the common sense rules of morality' and 'the moral standards appropriate to the individual's role in ... various associations' are 'impressed on [the individual] by the approval and disapproval of those in authority, or by the other members of the group'.[7] Rawls clearly regards his third stage 'morality of principles' as more autonomous; but (as I read him) it is not exempted from his general account of guilt, which he connects with 'the concept of right', whereas shame is connected with 'the concept of good'. According to Rawls,

> one who feels guilty, recognizing his action as a transgression of the legitimate claims of others, expects them to resent his conduct and to penalize him in various ways. He also assumes that third parties will be indignant with him. Someone who feels guilty ... is apprehensive about the resentment and indignation of others ...[8]

In this way, we may say, awareness of the fact or possibility of guilt, which accompanies our awareness of moral obligation as its shadow, so to speak, expresses our awareness of being required, in various ways, by other persons, to fulfil our obligations.

---

[7] Rawls, *A Theory of Justice*, revised edition, pp. 407, 409.

[8] Rawls, *A Theory of Justice*, revised edition, pp. 423-4.

III. AUTONOMY AS INTEGRITY
AND AS INDEPENDENT-MINDEDNESS

I return now to my suggestion that we can see a morally important form of individual autonomy in acting from one's own positive valuation of ethical principles, even if one's coming to that positive valuation was in many ways socially conditioned. Here we begin to think about autonomy, not as an aspect of human nature, or a source or explanation of moral obligation, but as a virtue, an intrinsically excellent personal quality, and a source and explanation of moral behaviour.

I did not create myself out of nothing, and I could not have done that. I am a creature of God and a child of my parents. Neither did I create my morality out of nothing, nor could I have done that. My morality has been shaped in various ways by my own thinking, feeling, and acting; but it began with my initiation into an ethical culture that had been developed over many generations. And no doubt it is still much more like than unlike the morality of those with whom I have been most closely associated. Were that not so, our social conflicts would be grievous indeed. It would be ridiculous to aspire to a moral autonomy that would consist in living by a morality of which one is the sole creator, not even influenced by the morality of those other persons whom one respects most highly.

Other ways of living remain, however, that may reasonably be regarded as versions of autonomy and as worthy objects of aspiration. One of them I will call moral integration. It is an ability, and disposition, to govern one's life in accordance with one's own views and feelings, one's own attitudes, aims, and commitments, and especially with those that one values most highly and most persistently. It is an important excellence or strength of character; perhaps indeed it simply *is* strength of character – a trait without which one's moral virtue can at best be sadly limited. Just how virtuous it is to be autonomous in this way will depend, of course, on how virtuous the individual's predominant attitudes, aims, and commitments are. But in acting morally correctly in a way that reflects one's own deepest views and feelings, aims and attitudes, one is certainly more virtuous than one would be in following correct moral principles without much attachment to them, simply as a matter of habit, or social convenience, or safety.

And such deep and integrated commitment to moral views and principles is surely a more durable and reliable source of moral action

than one would be likely to have with a less autonomous style of self-government. That view is supported by common sense, and consistent with contemporary research in social psychology.[9] If a sense of justice is to be strong support of the sort of stability that Rawls argues that his ideally just political society would have, it must inspire motivation that amounts to autonomy of this sort.

Another virtuous version of autonomy that requires discussion here may be called moral independent-mindedness. It presupposes the consistent self-government and centred prioritization of motives characteristic of moral integration, but it involves something more than that: a developed ability and disposition to think for oneself about ethical questions and situations.

Our moral need for independent-mindedness is underlined by empirical research in social psychology (for example, Stanley Milgram's famous experiments on obedience to authority).[10] A general factor emerging in such research of social psychologists seems to me to have more fundamental significance for a theory of virtue.[11] It has been suggested that 'there is a single thread that runs through social psychology's discoveries of people acting in surprising and demoralizing ways: people's understandings of the world … are strongly influenced by what they take to be other people's perceptions…'[12] I would add that the empirical evidence suggests that we are similarly influenced by what we take to be other people's expectations. More generally, we are strongly motivated to be, and to seem to be, in tune with our social surroundings, and especially with those around us who seem to have some authority.

That is not a totally bad thing. Such a drive for attunement with those around us is equipment for our multi-dimensional cooperation with each other as social animals. Without it perhaps we would not have been able, for instance, to learn a language. It makes us vulnerable, however, to temptations of social conformity. They are a very important category of moral temptations: Much of the most appalling moral wrong done by human beings occurs in following a crowd to do evil.

---

[9] Cf. Robert Merrihew Adams, *A Theory of Virtue* (Oxford: Clarendon Press, 2006), chapters 8 and 9, especially pp. 130-38.

[10] Stanley Milgram, *Obedience to Authority* (London: Tavistock, 1974).

[11] Here, and in the remainder of my discussion of moral independent-mindedness, I draw at more than one point in my discussion in *A Theory of Virtue*, pp. 153-55.

[12] John Sabini and Maury Silver, 'Lack of Character? Situationism Critiqued', *Ethics*, 115 (2005), 535-562 (p. 559).

This is a category of temptations that does not receive its due in traditional lists of the virtues. To deal with temptations of fear, the list has included courage. To deal with temptations of pleasure and physical desire, it has included moderation. What is the virtue of dealing well with temptations of social conformity? We can hardly raise that question without realizing that we do not have a standard answer to it. I suggest that moral independent-mindedness is the virtue we are looking for.

At the heart of this virtue, the habit of thinking for oneself involves a certain critical stance, a readiness to question, particularly in moral matters, one's own opinions and aims, and those of everyone else. At the same time moral independent-mindedness involves an ability and disposition to act with some confidence on conclusions one reaches in one's thinking, even if they do not agree with those of other people. Contrasted with this virtue are vices of deficiency and excess: social conformism, as a deficiency, on the one hand; and an excess of moral self-confidence, or even a contempt for the views and aims of other people, on the other hand.

IV. AUTHORITY, INDEPENDENT-MINDEDNESS, AND GOD

Persons or institutions have authority insofar as they are authorized to act, at their discretion, on behalf of some person, or group or organization of persons, or to decide how those persons will or should act, and treat each other, or what their structures for cooperation, and their policies and practices in relation to other groups, should be. Social groups of large or medium size, and some smaller groups, seem to need to have leaders with some measure of authority in that sense if they are to cooperate effectively and avoid unnecessary conflict. Recognition of authority, and deference to authority, are therefore part of ordinary life in almost every human culture. As we come to understand how our societies work, most of us conclude that it is good to have some people authorized to make certain kinds of decisions on behalf of all of us. And if we start a new organization we will very likely give some individual or group authority to act on its behalf, and to coordinate our cooperative endeavours as an organization.

In recognizing, or bestowing, authority, or in being born and growing up in a society that bestows authority on its leaders, one is apt to acquire obligations to obey or comply with the directions of those in authority.

These may be obligations to those in authority or to other members of the society, or commonly to both. In everyday life we routinely, and often casually, accept obligations of this sort: legal obligations, obligations to comply with the instructions of supervisors at work, obligations to collaborate in accordance with decisions of the organizers of a conference in which we participate, or the coach of a team to which we belong, or the conductor of a band in which we play or a choir in which we sing.

Can we be independent-minded in our acceptance of these obligations? It may seem that we cannot. The middle of a concert or a football game hardly seems the right time for a critical stance, or a readiness to question the direction of the conductor or the commands of the coach. More to the point, the middle of a combat operation or a violent crime scene may well seem not to be the right place for a soldier or a police officer to question critically the command of a superior officer. Yet in some situations of those types, those in authority have certainly issued orders that morally ought not to have been obeyed. Clearly there can be a tension between our ordinary and usually justified, or even urgently needed, deference to established authority, and our need for independent-mindedness as a barrier to stop us from immoral obedience to immoral directives that may come from people in authority.

I have suggested that the virtue of independent-mindedness can be seen as a mean between vices of too much and too little deference to the views and aims of other people. The quantitative form of the expressions 'too much' and 'too little' could mislead us here, however. The question that the virtuously independent-minded person must always keep at least in the back of her mind is not how much deference, or obedience, is too much, and how much is too little, but in what cases deference is owed, and in what cases obedience would be wrong. But the independent-thinking person needs not only to seek right answers to the latter questions; she also needs to give unhesitating obedience to authoritative directives when that is right and urgent, while retaining a sensitivity that would bring the critical stance rapidly into play if an immoral directive were given. Such a sensitivity might well be an aspect of autonomy as moral integration. And it might help to have thought in advance, in a cool hour, about what sorts of command it would be morally wrong to obey.

Here, since our topic is autonomy and theological ethics, we may turn our attention to God's authority and God's commands. First I will sketch a line of argument I have developed more fully elsewhere,[13] in support

of the view that a grounding of moral obligation in God's authority and God's commands can support the social demand character of the obligation, and at the same time can support its objective validity more adequately than any merely human demands can. Then I will reflect at greater length on the question whether there is any place for the virtue of independent-mindedness in our relation to God's authority and God's commands.

The notion of divine commands plainly applies to God an analogy drawn from human institutions. A conjectural history of the conception of moral obligation might begin with social practices of commanding and obeying, and associated roles of authority. These practices will have involved a conception of obligation, though probably not a fully moral conception at first. For initially no distinction may have been drawn between what is required by human authorities and what is truly, objectively, or morally required.

At some point, however, it will have been realized that actual human social requirements are often not good enough to constitute an adequate basis for moral obligation. Experiences of conflicting social demands, and abuse of authority, inspire thoughts of a moral demand that may be superior to the demands of human authorities. Belief in superior personal powers or gods, with whom we may be seen as having a social relationship of some kind, suggests the obvious hypothesis as to the source of such a superior demand. Thus a divine command theory of the nature of moral obligation might arise as an idealized version of a social requirement account of the matter. I suspect that this conjectural history approximates the actual history of the conception of moral obligation in more than one society.

My argument continues with reflection on the question, what reasons should move us to comply with God's directives. I do not begin with, nor even include as a reason, the consideration that God is supremely powerful and can force us to obey, or can punish us for disobeying and reward us for obeying. That would be the wrong sort of reason to motivate an autonomous obedience. Like the reasons that Rawls proposes for supporting principles of justice, the reasons on which I wish to focus here are reasons for valuing social bonds – in this case a social bond or relationship with God.

---

[13] Robert Merrihew Adams, *Finite and Infinite Goods* (New York: Oxford University Press, 1999), chapter 11.

Developmentally, in Rawls' account, the first reason for valuing a social relationship is that it is a relationship to someone who is good to us. The child comes to love the parents because they 'love the child and manifestly express their love by caring for his good'. Similarly, if God is our creator, if God loves us, if God gives us all the good we enjoy, if God has made a covenant of faithful friendship with us, if God has become incarnate and even died for us to be our friend, to rescue us from our own perversity, to bring us to the greatest good – those are all among the reasons that theistic religion has found for valuing our relation to God. It is no accident that the Ten Commandments, in the Bible, begins with the reminder, 'I am the Lord your God, who brought you out of the land of Egypt, out of the house of bondage' (Exodus 20:2).

A second type of reason for valuing a social relationship is admiration for one's partner in the relationship. '[M]oral exemplars, that is, persons who are in various ways admired and who exhibit in high degree the ideal corresponding to their position' play a part in Rawls' account of our valuing just relationships.[14] This type of reason obviously plays a part in devotion to God. God is conceived as the most perfect, intrinsically the most excellent and admirable being. That is a central theme – I think the central theme – of theistic worship as such. I would push that theme further, into metaethics. Classic theistic philosophical traditions have ascribed to God the role assigned to the Form of the Good in Plato's Republic: the role of the Good itself, the exemplar and standard of excellence by which all other intrinsic excellence is objectively measured.[15]

A third type of reason for accepting requirements laid upon us in a social relationship is that we recognize the requirements themselves as good, either intrinsically in themselves or as reasonable in view of beneficial consequences that they have. Thus Rawls' 'third law' envisages a person coming to honour the requirements imposed by a just political order 'as he recognizes that he and those for whom he cares are … beneficiaries of these arrangements'. Similarly, in accepting God as the supreme moral authority it is crucial (and normally believed) that God's commands spring from a design and purpose that is good, and that the behaviour that God commands is not bad, but good, either intrinsically or in its effects.

---

[14] Rawls, *A Theory of Justice*, revised edition, p. 413.

[15] A view of this type is defended in Adams, *Finite and Infinite Goods*, especially chapters 1 and 2.

Against that background, we must raise the question whether there is a place for autonomy as independent-mindedness, and its critical stance, in relation to God's moral authority. I believe there is. Except insofar as theists are engaged with doubts about their faith, their independent-mindedness will not normally be manifested in questioning whether it is right to obey God, or entertaining the thought that perhaps they ought to disobey God. But that leaves plenty of room for critically questioning whether commands that are alleged to be God's really are from God, or whether they have been correctly interpreted and understood.

Independent-mindedness of this sort seems to me to be a requirement of theism. A temptation that is hardly separable from religion is that of idolatry – specifically, of failing to distinguish one's own religion from God. Idolatry is misidentification of God.[16] Avoiding it requires independent-mindedness. It also requires balancing independent-mindedness with openness (reflective rather than unreflective openness) to influence of criticism from others. Both of these, therefore, have a part to play in devotion to God.

V. AUTONOMY AND LOVE[17]

The divine command theory sketched above is not a complete theological metaethics. It gives an account of only one of our central moral concepts – that of moral obligation. The rationale I have given for it plainly presupposes other ethical concepts for which I do not think it would be plausible to give an account in terms of commands (God's or anyone else's). Specifically, it presupposes concepts of value – specifically, of what is beneficial, or good for a person, and of intrinsic excellence which is worthy of admiration (or of worship, in the case of supreme excellence). In trying to develop a complete theological ethics and metaethics, in fact, I would not begin with divine commands but with God as the supreme Good and the standard of all other intrinsic value. In such a theological ethics, obedience to divine commands has a place, but is not the most central human response that is sought. The central response is worship, or more broadly, love.

---

[16] Cf. Adams, *Finite and Infinite Goods*, pp. 209-13.

[17] In much of section 5 I am remapping ground covered in parts of my paper on 'Christian Liberty', in Thomas V. Morris (ed.), *Philosophy and the Christian Faith* (Notre Dame, IN: University of Notre Dame Press, 1988), pp. 151-71.

Jesus is famously quoted as identifying two commandments in the Hebrew Bible as the most important of all. 'The first is, "... you shall love the Lord your God with all your heart, and with all your soul, and with all your mind, and with all your strength." The second is this, "You shall love your neighbour as yourself."'[18] The content of these commands stands in a somewhat paradoxical relation to their imperative form. Love may indeed be demanded, but where love is at the centre of a life, the role of obedience to commands must be limited. For obedience and love compete with each other as motives. Where obedience is my motive, it focuses my attention on my own action and its conformity or nonconformity with a command. But where love is my motive, it focuses my attention on the object of my love. Perhaps the objects of those motives may coincide, if the object of love is a command. But in loving God or one's neighbour, the object is presumably much more than any command.

No doubt we have obligations, and no doubt they are needed. But obligation that is morally compelling has its home in personal relationships. And in those personal relationships there are bonds that are prior, motivationally, to the obligations. There are 'attachments' or 'affection for particular persons', as Rawls puts it, or 'love' as we might say in a theological ethics.

In the 'farewell discourse' in the Gospel according to John (15:12-15), Jesus is reported as saying to his disciples,

> This is my commandment, that you love each other as I have loved you.
> No one has greater love than this: to lay down one's life for one's friends.
> You are my friends, if you do what I command you.

Then he adds,

> I no longer call you slaves, for the slave does not know what his master
> is doing; but I have called you friends, for I have made known to you
> everything I have heard from my Father.

In those words being friends of Jesus – and no doubt, by implication, friends of God – is linked with obeying a version of the command of neighbour-love. And in almost the same breath being friends of Jesus is linked with not being slaves – not being slaves of Jesus, at any rate, and by implication, I think, not being slaves in relation to God. And in relation to God, being friends rather than slaves is linked with knowing, through Jesus, what God is doing. I find this concatenation of ideas very suggestive for our topic.

---

[18] Mark 12:28-31, quoting Deuteronomy 6:4-5 and Leviticus 19:18.

If we seek a word in the New Testament that corresponds with our word 'autonomy', we might think first of 'freedom', '*eleutheria*' in Greek. This word had less individualistic connotations in the social context of the New Testament than 'freedom' does in much of our discourse. It did imply exemption from some types of social constraint; but more fundamentally it signified a social position. The free person, the *eleutheros* in the literal sense, was precisely a person who was not a slave – who was not owned by a master, as a horse or a plough might be owned – someone who did not belong to another person in that way.

To want someone as a friend is also to want him or her to belong to you, but in a very different way. Friends want to have claims on each other, which will constrain their choices in some ways. But friends, as such, are less interested in controlling each other than in sharing in each other's being themselves, speaking their own minds and acting in their own integrity.

Friends share. They share things; they share experiences; they exchange thoughts and feelings; they share projects. A friendship itself is indeed a shared project. This brings us back to John 15:15. Why is it that God's friends are to know what God is doing? The point, I take it, is that whereas God's slaves, if God had slaves, would only need to know what God told them to do, God's friends are to share projects with God. And if they are to enter into God's projects and make them their own, they need to understand the projects, so as to recognize what will and won't serve to fulfil them.

In the text I have quoted, Jesus tells his disciples that if they do what he commands them, they are his friends. They enter into a common project with him. And the command he gives them is that they are to love each other. The common project is love. They are to enter into God's project by loving what God loves. And Jesus offers them understanding of the project. They are called, not to blind obedience, but to vision – called to understand what God is doing, and discern what they can do to participate in the project.

What Jesus asks from his friends in these sayings is something like what I called 'moral integration' in section 3. In this case it is an integration centred in love. God's love for people is to become their love for people too; their lives are to cohere with that love as a central and controlling value. Their action is to flow from their own loving, and from their own discerning how that love may be fulfilled. Yet that is not

to be separate from their entering into God's love, and understanding what God is doing.

I used to speak of 'theonomy' in this connection, quoting Paul Tillich's formulation, that 'Theonomy asserts that [God's] law is, at the same time, the innermost law of man himself, rooted in the divine ground which is man's own ground'.[19] I said that the ideal proposed in these sayings from John's Gospel 'is not one of heteronomous subjection to a law whose motives are alien to the human agent; it is an ideal of theonomous permeation of the human faculties by the Spirit of God, so that the human agent comes to love what God loves and to see ethical priorities as God sees them'. Much of that formulation still seems to me apt. What is envisaged in the sayings is loving what God loves (loving it for its own sake, as well as for God's sake) and seeing ethical priorities as God sees them. And those motives are not to be alien to the human agent. They are to be the human agent's own deepest motives. But I now think it was misleading to use the term 'theonomous' here. For the motive whose possibility I wish to affirm in this context is not a law – not even an innermost law in the agent – but simply a love.

For the same reason perhaps 'autonomy' also is not the best term to use in this context. It carries connotations – or even implications – of law (*nomos*), which are certainly intended in a Kantian context. But the focus in the Jesus sayings of John 15:12-15 is not on government of lives by laws or rules, but motivation of lives by love. Laws and rules certainly have their place in human life. Societies need them for social peace and cooperation. And personal rules and principles can help us individually to resist temptations and live more productively. They can play a part in the virtue of consistency and living by one's own deepest values that I call moral integration.

To speak of a moral integration centred in friendship with God and with other friends of God, however, is to speak of something more than principled self-government. It is to speak of a system of loves that imprint their own organization on one's life – an organization of motives and not primarily of rules or laws. It is like principled self-government in engaging and drawing out what is central to one's selfhood. Friendship requires engagement and expression of the selfhood of each of the friends. But that involves much more than rules or laws.

---

[19] Paul Tillich, *The Protestant Era*, abridged edition (Chicago: University of Chicago Press, 1957), pp. 56-57; and Adams, 'Christian Liberty', p. 157.

COMPULSION OR ATTRACTIVENESS – A FALSE DICHOTOMY?
A KANTIAN PERSPECTIVE ON THE SOURCES OF MORAL MOTIVATION

MARIA SCHWARTZ

University of Augsburg

Abstract. The essay questions the dichotomy between 'push' and 'pull' motivation to act morally, asking for the motivational power of Kant's categorical imperative instead, its functionality as well as its sources. With reference to Christine Korsgaard it can be shown that personal integrity together with the notion of an ideal common world form one single source of motivation, grounded in exercising our autonomy. In a last step this outline of a kantian ethics of automony is related to the notion of God, whose role is illustrated in Kant's *Religion Within* in a surprising way.

Since G. E. M. Anscombe's *Modern Moral Philosophy* (1958) many if not most scholars in ethical philosophy, Linda Zagzebski amongst them, open a sharp dichotomy between two types of ethics (cf. Zagzebski 2004: xi f.):

(1) An ethics of duty (deontological), associated with a commanding law, duty (the 'ought'), obligation, compulsion and force, concerned with morally right or wrong acts.

(2) An ethics of value or virtue (teleological), associated with attractiveness, thick moral concepts and being a good or virtuous person.

However, taking a closer look at the ethics of Immanuel Kant, I think this dichotomy is wrong. Specifically, there is no such conflict as compulsion versus attractiveness once we look at the sources of moral motivation from a Kantian perspective. My contribution in this article

aims at pointing out two main and closely connected sources of moral motivation, which underlie Kantian ethics:

(a) Personal integrity and self-conception as a free human being.

(b) The notion of other people in this world as free human beings.

Both of these sources are related to compulsion and attraction alike. They are grounded in valuing oneself and others as free human persons, living in a common world.

In the following I will ask 1) for possible categories of factors motivating our actions, drawing a very simple picture. My questions are a) Who or what is motivating? and b) *How* does motivation work? In part 2) I move on to take a closer look at the motivational power of the Categorical Imperative. In part 3) Conception of the Self and 4) Kingdom of Ends, I rely mainly on Christine Korsgaard's *Sources of Normativity* (1996) to discuss personal integrity and the notion of an ideal common world as motivating factors.

In the final part 5) I attempt to relate this outline of an ethics of autonomy to the notion of God, asking if there is any place left for religious approaches at all and if so, what they would be.

I. MOTIVATING FACTORS – THE (VERY) SIMPLE PICTURE

Let us consider a basic situation of moral challenge:

A person P is in desperate need of help. I am the only one present and able to help by doing X. So X is the moral deed in question.

Question 1: What or who is motivating me to do X?

Once we ask *what or who* possibly is motivating me to do X, there are several everyday-answers like:

– God
– People
– Me, myself (Reason? Emotion? Will?)
– Values
– Duty

However, the philosopher intervenes instantly, pointing to at least three entirely different categories of concepts all regarded as 'motivating factors' here. Assuming God is a person these are:

a) persons/people, b) human faculties and c) abstract concepts like values or duties.

Question 2: How are these factors motivating?

Once we ask in a second step *how* these factors should be considered 'motivating', the mentioned dichotomy between attractiveness and compulsion arises. Yet there are significant differences in how well that dichotomy fits to each category. Let us again consider some everyday-answers people might give when questioned about their motivation to do X. Applied to the first category, motivated by persons (from a Christian perspective, including God), a first, encouraging answer might be:

> I love God and his commands, so I do X. Or: I like the person in need, so I do X.

We can regard this motivation as a 'pull-motivation', working via attractiveness. It is attractive to do X, so I act accordingly.
Another, more sobering type of answer might be:

> God (or people) will punish me if I don't do X.

We can regard this motivation as a 'push-motivation', working via force or compulsion. It is not attractive to do X. However, something very bad will happen if I refuse to help, thus I feel compelled to do X.

In the whole area of relationships, dealing with persons, distinguishing between the two different kinds of motivation seems just fine. And it makes even more sense in a theological frame like 'Divine Command' versus 'Divine Motivation' theory.

Moving on to the other two categories the distinction becomes questionable, mainly because certain types of motivation (like attraction) are tied to certain concepts (like values). It appears that inner-personal faculties like reason and emotion or concepts like duty and value are *as* ambivalent as persons with regard to their motivational powers. Each concept or faculty is capable of motivation via compulsion or attractiveness. For example, to look at everyday-language once again, consider these sentences:

> 'I will feel remorse when I don't do X.' – 'I like to be a helpful, virtuous person.' – 'I love to follow my duty.' – 'I feel compelled to realize a value, to carry it into effect.' – 'I *have* to change the world, no matter what!'

As the last two examples indicate, strong values might attract somebody in a way that makes it extremely difficult to see any difference between a strong attraction and a compulsion at all. Pull easily becomes push and vice versa.

I am questioning if the distinction between 'push' and 'pull', between 'compulsion' and 'attractiveness' is helpful at all. As a side remark, Zagzebski's theory is much more complicated and tries to bind these concepts together. She would most likely regard this picture as over-simplified (where I would agree). It might also look different once there is no clear dichotomy between reason and emotions in the first place[1]. However, this simple picture will be useful for the following parts where I am going to outline a theory of autonomy, especially the concept of duty ('Pflicht'), addressing the kind of motivating power in Immanuel Kant's ethics (part 2) and its sources (part 3 and 4).

II. THE MOTIVATIONAL POWER
OF THE CATEGORICAL IMPERATIVE

How to describe the motivating power of the Categorical Imperative? Is it based on attractiveness or force? The answer I am trying to give is: Neither, because Kant has a wholly different account of motivating forces, the 'Triebfedern'. The moral law is regarded as the sole, objective ground of motivation. To become effective, however, it has to become the subjective 'Bestimmungsgrund'. How does this happen?
A very interesting remark in Kant's *Groundwork* tells us:

> ... [the human being] is subject *only to laws given by himself but still universal* ... if one thought of him only as subject to a law (whatever it may be), this law had to carry with it some interest by way of attraction or constraint, since it did not as a law arise from *his* will. (G 4:432-433)

Therefore, attraction or constraint ('Reiz oder Zwang') is *not* how the law given by one's own will works. Why is it not motivating in the way of a push-motivation, by force? Kant is pointing out here that we are used to relating the term 'law' instantly to force, since when there are laws, there are also institutions which care for these laws to be enforced. We speak of 'armed forces' or the 'force' of the circumstances which leaves us 'no choice'. 'Force' could also describe the power to make people do something they do not want to do, be it by means of pure violence or by manipulating them otherwise. The latter might be the crucial point.

There is not the slightest force of that kind in the demands of the Categorical Imperative (CI). To be obligated means to know you *should*

---

[1] Zagzebski discusses cognitive and affective aspects of emotion (cf. Zagzebski 2004: 68).

do something, just because you know it is good to do – especially when there is nobody around to *force* you to do it.

There are other quite interesting remarks in Kant's later works, mainly the *Metaphysics of Morals* and the *Religion within,* where he assumes any incentive ('Triebfeder') has to be willingly included into one's maxim:

> ... freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim* ... (Rel. 6:23-24)

Even the Categorical Imperative itself has to be willingly endorsed, be adopted as one's highest maxim, in order to be put into effect. At least this is what the *Religion within* tells us. Thus, to speak of attractiveness or compulsion is equally wrong, because the underlying principle of adopting moral maxims is freedom. Starting from the adoption of maxims and down to the specific actions we perform, it is always our very own freedom that enables us to decide. It is our free choice to do X, even and specifically *if* there are strong external or internal forces standing against a moral decision.

In both the *Metaphysics of Morals* and the *Religion within* alike a decision is needed for whatever good motivating factor, even the moral law itself. Otherwise it will have no effect on one's maxims, and consequently on one's actions, at all. The rational endorsement of any factor whatsoever, apparently seems to neither have nor need any further source, because at that point it would no longer be a free decision.

In our initial example you might expect to get punished for helping the person in need by an unjust regime or P might be your worst enemy. Yet you are the only one who is able to help, knowing it would be a good thing to do. This 'knowing I should do X' amounts to the moral obligation standing against other inner and outer forces alike.

Assume in our example that you would just walk away and nobody will notice your refusal to help. Nothing will happen to you if you just decide differently. You might escape punishment, which is normally considered a good thing. Or you might experience some inner satisfaction of paying your enemy back, successfully suppressing any forthcoming feeling of remorse.

Thus, the rigid 'commands' of the Categorical Imperative are probably the only type of commands in this world which are *free* from force. To be obligated by the CI means to be motivated (in a way I will

discuss shortly), but still always be free to do otherwise. From an outside perspective, its demands seem to be powerless, un-founded, so to say. Gerd Haeffner, a teacher of mine and Phenomenologist, once spoke of the harmlessness and innocence ('Unschuldigkeit') of the demands of the Categorical Imperative. I never heard a more apt expression to describe its motivating power.

When it comes to explain acting on duty ('aus Pflicht'), it is often stated this means doing your duty just because you know it is duty. To do a good deed just *because* it is a good deed seems like an empty command, however, it fits to the notions of choice and freedom.

A brief side-remark regarding possible objections: Kant does speak of 'Nötigung' frequently (for example in G 4:413, G 4:434, MM 6:437), which could be translated as 'necessitation' or even 'force'. It is also difficult to see how the intellectual feeling of respect ('Achtung') fits into the motivational picture, which would be a separate topic. I think Kant is constantly struggling to hold on to the initial idea in the *Groundwork* that the CI is adopted due to rational, free decision alone, and nothing else.

Still, the standard interpretation that acting morally in Kantian ethics amounts to following duty for duty's sake might be a preliminary, outside view. Contemporary Kant scholars like Christine Korsgaard are asking further questions. We are surely not moral for *no* reason, just to be moral. The moral saint will rescue her enemy, even if she might endanger her own life doing so. In answering the normative question of *why* she would act in this way, there must be more to say (cf. Korsgaard 1996: 13, 17-18). Why should moral demands bother us at all if they were only suggesting 'doing duty because it is duty'? Is there really no further explanation? This question leads us to part 3), Korsgaard's motivational account.

III. PRACTICAL IDENTITY – OUR CONCEPTION OF THE SELF

Christine Korsgaard offered a very interesting answer to this question with regard to personal identity, combining the Kantian with a Platonic approach (cf. Korsgaard 1996). To understand her thesis in a nutshell, again considering our initial example: I cannot 'live with myself' any longer if I do not do X because my deepest identity consists in acting according to the CI. A human being's identity first and foremost consists in acting as a free person, in making her or his own free choices. The actual place where freedom fulfils itself is the moral realm.

Kant assumes there is a hierarchy of different classes of imperatives and accordingly of different maxims[2]. What Korsgaard additionally assumes is that we are acting in the light of a hierarchy of different practical identities or conceptions of ourselves[3], illustrated by the following graphic:
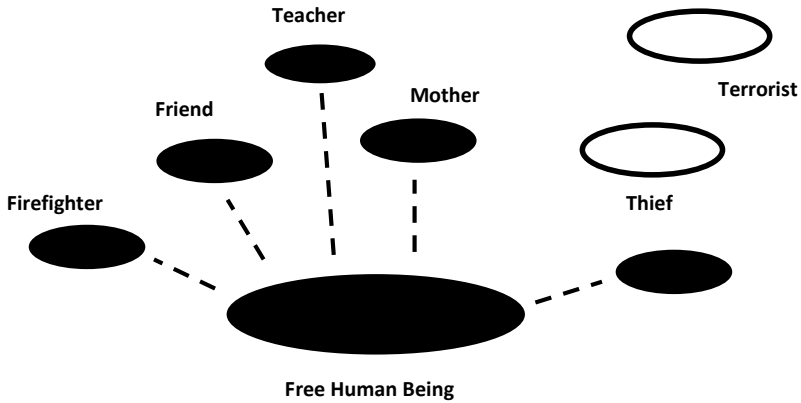
Identity as a free human being (in my graphic at the very bottom) is connected to the CI. In this light I choose all my different contingent social identities, which could also be regarded, in the Kantian terminology, as whole sets of maxims that we are trying to live up to.

Contingent identities are such as the identity of being a good friend, mother, teacher, or firefighter. All of those come with different duties. If a teacher cannot get up in time to hold his classes he will probably lose this identity sooner or later or at least not be regarded as a good teacher. The dotted lines in the graphic indicate all those identities which could be chosen in accordance with the Categorical Imperative. There are also some which can't be chosen without damaging my fundamental identity, like the one of a terrorist or thief. Both identities include using other people as a mere means to one's ends, impairing or even eliminating the freedom of others, so they must not be chosen. How it is even possible to choose an identity like that, willingly, which amounts

---

[2] An imperative suggests something I *should* do. As soon as I answer positively, deciding that this is something I really *want* to do, it becomes my maxim (cf. Schwartz 2006: 44-76).

[3] Cf. for the following Korsgaard 1996: 120-123.

to the abandonment of a fundamental part of what makes us human, remains an unsolved mystery in the *Religion within*, where Kant speaks of the radical evil (for example Rel. 6:32, 6:35, 6:43)[4]. Once chosen, the practical identities are a powerful source of motivation or, as Korsgaard states, the source of obligation:

> It is the conceptions of ourselves that are most important to us that give rise to unconditional obligations. For to violate them is to lose your integrity and so your identity, and to no longer be who you are. ... An obligation always takes the form of a reaction against a threat of a loss of identity. (Korsgaard 1996: 102)

Back to our initial question about the kind of motivation: If not the Categorical Imperative itself, are practical identities perhaps providing a push- or pull-motivation, based on attractiveness or force? Judging from the latter quote, Korsgaard rather seems to think it is a push-motivation, a negative one, grounded in the fear of damaging one's integrity.

However, I would like to add that if you are living up to an identity, for example, as a good friend, by choosing not to betray your friend you are strengthening your identity. On closer inspection it might be very attractive to not only preserve, but live up to the identities we value. The feeling of respect ('Achtung') also points in this direction. After doing X one could say: 'This might have been hard, but I'm proud I did X. It was the right thing to do.' It does not seem to be only the fear of losing one's identity, which might motivate some of our actions, but also the positive feeling of forming a character and living up to one's true self.

This ambivalence especially applies to the most fundamental identity, the one of being a free human being. An insight of spiritual literature (for example in the works of St. Ignatius of Loyola) is that the more you practice your freedom, the more free you are becoming. This bottom source of motivation, one's human identity, can neither be categorized as compelling nor attracting. It is about freedom, lying above or in between.

And yet this approach faces some common criticism. If my self-conception is the underlying source of moral motivation, wouldn't it just be my own selfishness motivating me, striving for preserving my

---

[4] The inclinations ('Neigungen'), can't be regarded as root of evil choices; these are only responsible for *akrasia*. As Kant writes, inclinations are opponents of principles in general ('Grundsätze überhaupt'), whether they are good or bad (cf. Rel. 6:59, Footnote). Once a person suffers from *akrasia,* she can't even rob a bank successfully. The free decision of the will is necessary condition for good and evil actions alike.

identity? I think that is only one side of the coin. One further step is left to answer this objection which I will outline in the following part 4.

IV. THE KINGDOM OF ENDS – BEING PART OF A COMMON WORLD

Taking a closer look at *what* the Categorical Imperative actually commands adds a whole new dimension to the picture. The CI is not concerned with specific advises like 'Do this' or 'Don't do that'. It is a formal principle. Thus, to understand the motivational force of the CI, the formulas and their application need to be discussed. To consider one of its most known formulas in the *Groundwork*:

> '... act only in accordance with that maxim through which you can at the same time will that it become a universal law.' (G 4:421)

What does it mean to 'will that it become a universal law'? How are we to judge what should become a universal law and what not? To even try to do so, I have to take the view of other people into account, treating them as ends in themselves. The universal law is not exactly a random law I am giving *myself*, stuck in some sort of solipsism. To focus on autonomy ('Selbst-Gesetzgebung'), like we did so far, would be only one – and an incomplete – description, omitting the vital point. Once we take a closer look at the formula, it opens a new horizon towards other people: The horizon of all of us forming a common ideal world, a Kingdom of Ends.

The reason why many people regard the Categorical Imperative as implausible, not matching our everyday experience, is that this horizon is never fully conscious. Though we are acting in its light all the time, we are not doing so consciously. Nobody asks himself 'can I universalize the maxim underlying my action?', deriving it step by step from the CI as if in a mathematical operation. All of this is happening way faster, parallel to deliberation or rather, forming the background of our decision making.

If we want to apply the Categorical Imperative consciously, like we do in ethical theory, it requires nothing less than imagining a whole ideal world in which we all want to live. Becoming a law means that everybody is allowed to act like me, in a similar type of situation. It further means that I will grant all my reasons justifying my action to everybody who acts the same way, while I am affected this time. The notion of universal law requires taking other human beings seriously, which leads us instantly to the second formula, to treat them as an end, not as mere means:

'*So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.*' (G 4:429)

And once we combine both of these formulas, we arrive at the formula of the Kingdom of Ends. Like the following graphic shows, there must be a number of persons to whom the law applies (here pictured as dots):
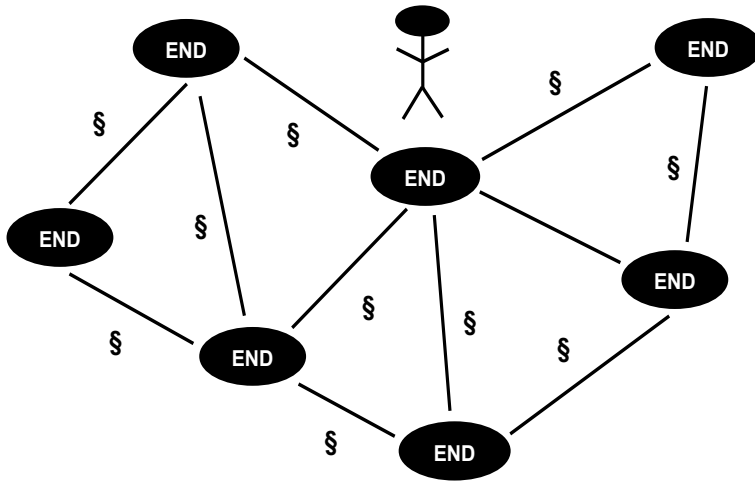
The structure provided by the universal law-formula necessarily has to be 'filled' by people which are ends in themselves. Laws regulate the relations between people. Once I am doing something which affects another person, it always should be something which I also agree to be done to me, justified with same relevant reasons.

A side remark on the importance of reasons given in the maxims: For example, suicide is not regarded as a moral decision in the *Groundwork*. The reason for this verdict, however, is not the obvious one that I am getting rid of my humanity altogether (since then suicide would be prohibited under all conditions). It is rather the reason the agent himself provides for his decision: The suicidal person in the example is neither sick nor about to be executed. He simply doesn't expect the rest of his life to be more pleasant than unpleasant (cf. G 4:422). In the *Metaphysics of Morals* Kant questions if suicide shouldn't be allowed in other cases, for other reasons, like committing suicide in order to rescue his country

or to save others from a contagious illness (cf. MM 6:423f.). So the universalization of maxims includes a universalization of reasons. Not mere acts are assessed, but whole rich descriptions of actions, including circumstances and reasons, like they are already found in the examples of the *Groundwork*.

Applying the CI as a formal principle requires 'filling' it, considering all the values, emotions, the specific factual needs and interests of other people. Kant himself is not very much bothered with this empirical base, though. He is only unfolding the basic structure and theory.

What does the notion of the Kingdom of Ends add to our understanding of motivational sources? After pointing to one's own integrity in part 3, we now possibly notice an even deeper motivating factor, which stems from our humanity *as* being part of a common world full of relationships and other people. What does it mean that these people are not allowed to be used as a mere means to our personal aims? They have to agree to our actions, *in case* these are constraining their legitimate freedom.

Like we saw in part 2, however, there is not the slightest force making us take the view of other people into account which are affected by our actions. We are free to ignore the common world idea, anytime. Most people committing crimes know perfectly well the world wouldn't function anymore if everybody acted like them. And most even rely on others *not* doing the same, making exceptions for themselves, as Kant writes in G 4:424. The common characteristic of immoral actions is that nobody could rationally wish to live in a world where such courses of action become a law.

The advantage of the criteria of universalization is that it does not stay as opaque as an ethics of moral examples. Reading Aristotle can leave us puzzled and helpless when it comes to the question of what exactly is a moral thing to do. I simply do not know what the moral expert (a *phronimos*) would do if he were me. Like John Hare pointed out (cf. Hare 2005) most of us probably even more lack the empathy to be able to judge what God would do. To even identify moral examples like St. Francis of Assisi or Gandhi, which are suggested by Zagzebski (cf. Zagzebski 2004: 46), we need to have a notion on what exactly is good or bad in the first place. It requires a lot of imagination and empathy to see what a morally wise person would do in my situation. Answering the question in which kind of world we would like to live in, on the other hand, is much easier. Most likely this will be a world of peace, of respect and caring, no cheating, no stealing, no killing and so on.

To summarize: First, like outlined in this part, I do not think human autonomy in a Kantian framework has to do with solipsism or selfishness at all. The motivational power of the CI stems from a fundamental identity, which is inseparably connected to the world of all human beings around me.

Secondly, as shown in part 2, since all moral decisions are free decisions, both approaches, be it a push- or pull-motivation, are actually not applicable to the Kantian framework. Considering our example of doing X, it does not matter if it is attractive to help somebody – for example my best friend or my child who is in danger – or if strong forces of whatever kind are holding me back. It would be moral to just help, not because there is a strange value to be realized through that deed, but because there is another human being in need of help, which is valuable as a person.[5] Nobody wants to live in a world where people make it a law to ignore each other. If I would constantly demand others to treat me as an end in itself while at the same time excluding myself from this demand, treating everybody as a pure means to my ends, I would be acting immorally. This is also not how God created us or meant us to be. Finally this leads us to a few brief remarks on the relation of an ethics of autonomy and the religious dimension.

V. RELATING AN ETHICS OF AUTONOMY TO THE NOTION OF GOD

Is there any place for God in this picture of human autonomy, and if so, what would it be? A first answer assumed by many is: There is none, God is left outside. It is like Plato vs. Protagoras: No longer God, but humans are the measure of all things, deciding independently what is right and wrong.

A second, theological answer has been given by Aquinas and others.[6] As the creator of persons which are a source of value and valuable themselves, God is also the creator of all values entirely. He constitutes freedom in a deeper way, enabling us to make moral decisions and do what we recognize as morally good. Even more, he also provided a fundamental moral criteria or moral insight to be found in every human's soul, religious or not. This answer might be compatible with

---

[5] It is possible to extend the scope of obligation to all living beings. Cf. Korsgaard 2004.

[6] See the contribution of C. Schröer in this volume for an elaborated account of this approach.

a Kantian account (in this case the criteria is regarded the CI), however, it is not what Kant would say.

Once we turn to Kantian ethics, the *Religion within* suggests a third and different answer. There is a very surprising remark in the Preface, where Kant writes:

> This idea [of a moral, most holy and omnipotent being] is not … an empty one; for it meets our natural need, which would otherwise be a hindrance to moral resolve, to think for all our doings and nondoings taken as a whole some sort of ultimate end which reason can justify. What is most important here, however, is that this idea rises out of morality and is not its foundation; that it is an end which to make one's own already presupposes ethical principles. (Rel. 6:5)

What exactly is 'a hindrance to moral resolve'? This is perhaps the most decisive part in Kant's ethical works, where it is not entirely clear if he is not close to introducing another motivational factor apart from the moral law on a different level.

God is introduced here as the provider of the highest good, a dimension of hope. If the common world we strive towards is only wishful thinking, an empty fantasy which will never become real, this could be an obstacle to moral decisions. God guarantees that sooner or later everybody will reach happiness (*eudaimonia*), not unconditionally, but according to her or his character, while he guarantees as well that we finally become those virtuous persons we are striving at. Yet this is not a different motivational ground for morality, as Kant instantly adds. If there were no hope for heaven or eternity, we are still supposed to act morally and be able to. For the Categorical Imperative to make sense, it must be assured the Kingdom of Ends is realized, sooner or later, as a community of rational beings in which nobody is using other persons as a mere means to his or her ends any longer. Since humans seem incapable to ever achieve this goal fully, God is the only guarantor that our moral efforts are not in vain. However, this dimension of hope is not regarded as a motivating factor for adopting ethical principles in the first place.

As I have tried to argue, the several sources of motivation outlined initially are actually only one: Valuing myself as a free human being, living in a world of other beings just like me. Referring to the title of my contribution, I do not think moral conflicts are taking place between compulsion and attractiveness, at all. They all come down to the question which attraction to follow, which compulsion to bow to, and why.

BIBLIOGRAPHY

Hare, John. 2005. 'Review of Linda Zagzebski, *Divine Motivation Theory*', *Notre Dame Philosophical Reviews*, 2.13

Kant, Immanuel. 1997. *Groundwork of the Metaphysics of Morals*, ed. by Mary Gregor (Cambridge: Cambridge University Press) [G]

Kant, Immanuel. 1998. *Religion within the Boundaries of Mere Reason And Other Writings*, Allen Wood & George di Giovanni (eds) (Cambridge: Cambridge University Press) [Rel.]

Korsgaard, Christine. 1996. *Sources of Normativity* (Cambridge: Cambridge University Press)

Korsgaard, Christine. 2004. 'Fellow Creatures: Kantian Ethics and our Duties to Animals', *Tanner Lectures on Human Values* 24: 77-110

Schwartz, Maria. 2006. *Der Begriff der Maxime bei Kant. Eine Untersuchung des Maximenbegriffs in Kants praktischer Philosophie* (Berlin: LIT)

Zagzebski, Linda. 2004. *Divine Motivation Theory* (Cambridge: Cambridge University Press)

FREEDOM: ON THE CONCEPT AND THE REALITY OF HUMAN AUTONOMY

MATTHIAS LUTZ-BACHMANN

Goethe-Universität, Frankfurt am Main

I. FREEDOM AS A BASIC CONCEPT IN PRACTICAL PHILOSOPHY

1.1. Introductory remarks

In the first paragraph of my paper I will discuss some aspects of the concept of freedom as a basic concept in practical philosophy. With the term 'practical philosophy' I refer to the part of philosophy which is dealing with the human as an 'acting being' in a proper sense. Additionally it is important to introduce the reflections in practical philosophy from the perspective of 'us' as the actors, that means from our experiences as acting beings. While we are acting we are somehow interacting and sometimes even co-operating with other actors like us. This is the way how Aristotle at first distinguished methodologically practical from theoretical philosophy.[1] In my introduction into 'Ethics'[2] I explained why this point of departure is necessary for the practical disciplines in philosophy like ethics or political philosophy. But in difference to Aristotle we have to refer, from the very beginnings of practical philosophy, to a 'intersubjective reality' of us as inter-actors, that means as beings who are necessarily and always mediated in our actions which other reasonable actors. This holds true even if a special human interaction is a 'limited one' in following a pure 'strategic' calculation of

---

[1] Cf. Aristotle, *Nichomachian Ethics*, Book 1, 1094 a 1 ff. and Book 6, 1138 b 20 ff.

[2] Matthias Lutz-Bachmann, *Grundkurs Philosophie Band 7: Ethik* (Stuttgart: Reclam Verlag, 2013), pp. 13-26.

others without the entire elements which have been analyzed in modern discourse-ethics as the unavoidable 'consensus'-oriented structures of human inter-action.[3]

In analyzing human action in that way, practical philosophy is dealing with the condition of human acting which we might call 'free action', introduced from the so-called first and second persons-perspective. This is true because we experience as actors our own acting and vice-versa the acting of all the other actors as somehow 'self-determined' or 'self-defined' if we are not totally obedient to someone else. That experience seems to differ from some of the ongoing debates in theoretical philosophy today, especially in the philosophy of mind. That is why I will shortly refer to this discussion in asking whether or not the framework of theoretical philosophy does allow us to assert that there is something like the reality of 'human freedom' in the outer physical world which cannot be denied in the name of scientific knowledge. I will point shortly to the so-called 'compatibilist' position which accepts the 'theoretical' possibility of something like 'human freedom' as human 'self-determination'.

This debate is not an invention of modern philosophy. We can find an evident analogy to this current debate in the philosophy of Immanuel Kant who accepted in his first 'Critique' (the 'Critique of Pure (Theoretical) Reason'), the idea of human freedom as a necessary concept of reason even if according to the premises of the 'Physics' of his days something like 'freedom' was not understood as an empirically secured 'appearance' in the outer physical world.[4] Nevertheless argued Kant in his second 'Critique', the 'Critique of Practical Reason' that from the prospective of practical argumentation – that means as I have already explained from the prospective of 'us' as inter-actors who are necessarily raising questions like whether or not a certain way of acting is *morally* right or wrong – we cannot avoid to presuppose reasonably the practical 'reality' of human freedom.[5] But this debate is even much older than Kant's discussion with David Hume and other philosophers following the principles of Newton's 'Physics' in the 18th century, it has its background in the late antiquity. I mean Augustine's controversy with Cicero and

---

[3] Cf. Jürgen Habermas, *Moral Consciousness and Communicative Action* (Cambridge, MA: MIT Press, 1990). For the status of consenus-orientation of discourse-ethics see e.g. William Rehg, *Insight and Solidarity. The Discourse ethics of Jürgen Habermas* (Berkeley and Los Angeles: University of California Press, 1994), Chap. III, pp. 56-83.

[4] Cf. Kant, *Critique of Pure Reason*, Akademie-Ausgabe III, 362 ff.

[5] Cf. Kant, *Critique of Practical Reason*, Akademie-Ausgabe V, 103 ff.

the Stoic tradition in 'De civitate Dei',[6] where we find something like a first philosophical debate on the theoretical possibility of human freedom (here: 'freedom of the will') under the conditions of a strict necessity and causality in the outer physical world: a necessity which is drawn from either an inescapable 'fate' or an idea of cosmological determinism which is the case in Cicero and the tradition of Stoic philosophy or from the premise of God's 'foreknowledge' and 'omnipotence' in the case of both some of the pagan ancient philosophers as well as the Christian thinkers. It is interesting to see that even Augustine's discourse is not looking for a pure theoretical solution for this problem even if it is true that he is pointing at the theoretical self-contradictions of the positions of Cicero and the Stoics.[7] For the philosophical solution of the problem how human freedom can be stated under the epistemological condition of a presupposed strict necessity in the physical world Augustine is aiming at a practical outcome meditated through the insight that a theoretical negation of the free will is leading the speculative thinking into unavoidable self-contradictions. (Here it is not the right place to refer to that very complex discussion in the history of thought which laid the ground for almost endless disputes in medieval schools of theology and philosophy).

1.2. Dimensions of human freedom as a 'practical concept'

Concerning the concept of 'action' in practical philosophy I suggest a first distinction between, on the one hand, actions that are determined externally or even compelled from the outside, that is, actions that are in the basic sense not free and, on the other hand, actions whose source could be traced to us. This is not to say that every action that is not completely compelled from the outside is free in an unlimited sense. That applies probably to only a few actions because actions occur in predetermined situations, that is, in circumstances which we cannot determine or change and as such can only react to in an appropriate manner. This also applies to our co-operation partner whose intentions and strategies we factor into our own actions. This way, we experience

---

[6] Cf. Augustine, *De civitate Dei*, V, 8-11, ed. B. Dombart/A.Kalb CCSL 47 (Turnhout: Brepols Publ., 1955).

[7] Cf. Eleonore Stump, 'Augustine and the Free Will', in *The Cambridge Companion to Augustine*, ed. by E. Stump & N. Kretzmann (Cambridge: Cambridge University Press, 2001), pp. 124-147.

ourselves as agents which are always defined and conditioned by certain constraints which we did not chose and which we aren't able to change in a short time. On the other hand, however, when we are not completely constrained or compelled, we know that outside the sphere of external influence, we act in accordance with our own chosen ends and means and we can thereby follow our intentions, plans, motivation and convictions. In this connection, we can speak of our freedom at the threshold of our practical experience.

The interpretation of this experience has led practical philosophy to the distinction between two forms or aspects of freedom: the 'negative freedom' and the 'positive freedom'. 'Negative freedom' manifests itself therein that an action is carried out to a certain extent without external influences or even physical interference. To this end, we define negative freedom as the absence of constraints to actions emanating from other actors. 'Positive freedom' manifests itself therein that, in what an agent does, his conviction, for example the rightness of reasons, follows his self-chosen intentions and plans. The concept of self-legislation or 'auto-nomy' which is central to ethics, although not first influenced by Kant, presents the only possible way by which the idea of a positive and self-determining freedom could be articulated. That means that the semantics of the concept of freedom is somehow larger than the concept of autonomy. Like I will show in the second part of my paper we can define 'autonomy', at least with regard to Kant, as a specific state of *reality* of freedom. With this way of reading practical freedom, we touch on further aspects of the concept of freedom which, in relation to Kant, could be differentiated by means of the concept of freedom of action and freedom of the will.

Through the concept of the freedom of the will in the sense of a negative freedom of an actor from the will of others or freedom from internal compulsion ('hetero-nomy'), we can distinguish the concept of '*arbitrary freedom*' (in Kant: 'Willkürfreiheit') which limits itself to the bounded scope of possible choice between many courses of action or predetermined goals of action. While the concept of 'freedom of action' refers to the capacity of the agent to initiate activities which in the real sense are external activities, the concept of the 'freedom of the will' presupposes the thought of an external being-in-action and a conceptually differentiated capacity of the subject of the will towards internal self-determination by choice and by modification of his/her action plans, intentions, maxims and preferences.

In this history of philosophical ethics, we find an in-depth conception of a positive freedom in the person of Socrates. His idea of freedom peaks therein that he does not offer any resistance to the compulsion of external laws and, by that means, against the compulsion of the state by resisting the unjust verdict through external action to absolve himself, for instance by fleeing, but rather that he demonstrated in a morally practical sense the paradox of his freedom by voluntary accepting the externally imposed unjust death sentence. In this presentation of Plato we can identify the reality of positive freedom even in the face of an extremely limited negative freedom. Socrates documents in his action a reality of the freedom of the will which reaches far beyond arbitrariness. He presents himself to us as an acting agent who determines himself as a free moral being. The drama of this proof is encapsulated in the fact that he can only be a free moral being by liquidating his physical-natural being through the suicide which was ordered by the authorities of the state.

This example also shows that Socrates' freedom was so extremely constrained by the social-political situation to the extent that we cannot talk of finding in his action a situation of negative freedom, that is, an action carried out without external compulsion even when the external compulsion was not organized in a way that left Socrates no other choice. And in the sense of recent political philosophy, we can say that the freedom of the will demonstrated by Socrates does not contain within it freedom in the sense of external independence from the arbitrariness of others.[8]

The ethics of Aristotle presupposes a concept of practical freedom in the analysis of human actions even if one has to admit that unlike modern ethics Aristotle's ethics is not construed on the basis of a concept of freedom nor on autonomy. Freedom is presented by him implicitly as something like freedom of action in the sense of freedom of choice or freedom from arbitrariness. According to him we are free to develop the virtues of our character which might help us to become morally good human beings and living a good life. In this way, Aristotle does not orient himself towards the concept of freedom as 'autonomy', but rather towards freedom as 'autarchy' or *self-sufficiency* which means towards the quest

---

[8] The idea of freedom as 'non-domination' is important for the theory of Philipp Pettit. Cf. e.g. Philipp Pettit, *Republicanism. A Theory of Freedom and Government* (Oxford: Oxford University Press, 1997).

for *self-preservation* in the natural striving of the human actor towards his/her aims like the fulfilment of one's needs and desires. In the tradition of Stoic philosophy the question of human freedom is contradictorily addressed. Here we can identify on the one hand a speculative theory of cosmological determinism which sees everything that happens within the cosmic order as predetermined by eternal laws (in Greek: 'logoi', in Latin: 'leges') which constitutes for good or bad conditions our fate which we cannot avoid. On the other hand Stoic philosophers propagate in their ethics a principle of virtue ethics that stands in opposition to the blind course of fate according to which only those who preserve their internal autarchy in faithfulness to necessarily formulated moral responsibility act morally correctly (cf. Seneca, Marc Aurelius). It is instructing that for this ethical view, a balance of indeterminacy (or of negative freedom) must be presupposed because without it, we would not have any possibility of choice.

Mediated through the theory of freedom of the will by Augustine (he speaks of a true 'liberum arbitrium')[9] through the discovery of the *internally human will* and later by the invention of the *intention* of each and all actions in Peter Abelard's writings[10] as well as through the theory of morality and the authority of *individual conscience* in Thomas Aquinas (he speaks of a so-called double aspects of conscience, the 'synderesis' designated practical reason und the separate or single conscientious judgement of the 'con-scientia'),[11] it has been – among others like Francesco di Vitoria[12] – Immanuel Kant who is grounding *a theory of human freedom* which has at its centre the concept of *autonomy* of a free acting subject which is morally good only if his/her will is determined through the *self-legislation* of the practical reason.[13]

A close reading of Kant reveals however that this practical theory of the autonomy of the will is confronted in Kant's theoretical philosophy with the argument that in the physical world of appearances, there is

---

[9] Augustine, *De libero arbitrio*, ed. W. M. Green, CCSL 29 (Turnhout: Brepols Publ., 1970).

[10] Cf. Peter Abelard, 'Scito te ipsum', see: Peter Abelard's *Ethics*, ed. and transl. by D. E. Luscombe (Oxford: Clarendon Press, 1971).

[11] Cf. Thomas Aquinas, S.th. I, q.79, a.12 ff.; I-II, q. 104, a. 1.

[12] Cf. Anselm Spindler, *Die Theorie des natürlichen Gesetzes bei Francisco de Vitoria, Politische Philosophie und Rechtstheorie des Mittelalters und der Neuzeit II, 6* (Stuttgart: Frommann-Holzboog Verlag, 2015).

[13] Kant, *Groundwork for the Metaphysics of Morals*, Akademie-Ausgabe IV, 393.

no place for that philosophical 'idea' of freedom – in accordance with the Newtonian physics. This situation made it imperative for Kant to limit, on the one hand, the talk about freedom of human beings to the arena of the knowledge of practical reason and, on the other hand, to circumscribe it in the sphere of a world of physicalistic appearance thought of as thoroughly predetermined, to the limited proposition that freedom refers to the capacity of 'beginning a condition by oneself'.[14]

This finding indicates that theoretical philosophy systematically encounters a problem regarding the concept of freedom in the context of law and experiential concept of modern science which cannot be probably solved by adopting the empirical paradigm of research in the natural sciences. It must be treated more comprehensively, as is possible here – for instance in view of the methodological premises of descriptively oriented theoretical philosophy. Hereby I mean a point of view which is avoiding to accept and to integrate the first and second person's perspective into its epistemological approach.

1.3. The 'theoretical approval' of human freedom

I would have to limit myself to the following in a rather *meta-ethical* perspective: the theory of the freedom of the will in theoretical philosophy in contemporary time, like for instance in the philosophy of mind, is one of the most controversial questions. Here, the disagreement oscillates, when simplified, between compatibilism which proposes that the endorsement of a practical freedom of human beings can be reconciled with the endorsement of a universal causality – which is definitely controversial in physics – or the predetermination of every occurrence in the world by the law of nature; and incompatibilism which disagrees with exactly this proposition. Some incompatibilists defend the view which proposes that to save the theoretical possibility of freedom, the theory of the universal causal or natural determinism of every occurrence must be denied (this enables chance as a space for freedom). Other incompatibilists however arrive at exactly the opposite conclusion by maintaining that the theoretical endorsement of freedom in the face of natural causality is impossible (that is the position of the so-called impossibilism). The problem of the incompatibilist position consists in the fact that to negate a universalist causal determination on

---

[14] Kant, *Critique of Pure Reason*, ibid. III, 363.

a general level doesn't support the concept of freedom but of chance or coincide in the name of an unclear metaphysical indeterminism.

These debates in theoretical philosophy today suffer from some problematic misunderstandings and some reductionisms which cannot be comprehensively discussed here. In this context, I would like to point at four more general philosophical problems: First we are confronted in these debates with a problem of – at least as defended by a few protagonists – quasi worldview *naturalism* which arises from unexamined philosophical as well as unproved scientific axioms. This is obviously unsatisfactory in the light of general reasonable arguments and not only from a philosophical point of view. Second we can point at the methodological problem which consists in the fact that some protagonists claim the request to answer philosophical questions abruptly by means of the procedure and results of *empirical natural sciences.* It is obvious that such epistemological transfer from the neurosciences to philosophical debates will never contain appropriate arguments for a reflexive and comprehensive philosophical theory of knowledge. Third it strikes one that internal philosophical questions of *ontology* are being solved through contributions of a *philosophy of mind* which does not seem plausible for diverse reasons since it indicates an inner-philosophical reductionism. In these debates, there is very serious need for clarifications, especially but not only, from methodological perspective. Fourth it should be finally pointed out that a sophisticated concept of freedom does not have to be contaminated with the concept of '*chance*' or '*coincide*' because the two of them, the concept of 'freedom' in the above mentioned meaning as a basic condition of human acting on the one hand and 'chance' as a 'fact' or a 'phenomenon' in a causally closed world on the other hand relate to totally different topics which should not be confused.

In order to refer to the three positions I mentioned before, the 'compatibilist', the 'incompatiblist' and the 'impossiblist' position, I can assume that at least a number of compatibilists and some of the incompatibilists – as long as they do not on their path appropriate the natural science axiom of a universal natural law determinism of the world of appearance so as to deny like the impossibilists, basically and in theoretical perspective the possibility of human freedom of action – are in agreement that a weak reading of freedom of action could be thought about and defended, even when it is done by means of very different arguments. This concept of freedom operates under the label of 'freedom as self-regulation' ('Freiheit als Selbstbestimmung') or alternatively

'freedom as self-determination' ('Freiheit als Selbstdetermination') in contemporary debates. We must in this context still nonetheless differentiate this interpretation of freedom from Kant's practical concept of freedom as autonomy or freedom of the will.

It appears – the in the real sense inconclusive controversy in theoretical philosophy notwithstanding – that against the background of these debates, the concept of freedom as self-determination could be taken to be capable of finding consensus, at least in a theoretically weak reading. This takes into cognizance the two aspects of freedom which we have mentioned with regard to human action, namely, the positive and negative aspects. This is the case because the concept of freedom as self-determination, or rather, self-regulation, defines freedom on the one hand by means of the absence of external influence or compulsion which is observable from a third person perspective. On the other hand, it qualifies the action as an occurrence which is factually positively determined by the actor as the causal originator in accordance with his action plan ('Handlungsplan'), his expressed reasons as well as his articulated intentions. In that sense the fact of an action could be understood as something which is caused by the actor and which could be traced back to him as the causal agent purely from the theoretical third person observer's perspective. For a practical perspective of philosophy this would not yet be a sufficient understanding of freedom, namely from the actor's perspective and his/her partners of cooperation. It is here where the more comprehensive perspective of intersubjectivity comes into consideration reflecting the proper meaning of human action as an intersubjective practice.

The compatibilist position which in theoretical perspective defends an ontology of occurrences of universal causal determination could be deduced from the understanding that the theoretical position of indeterminism does not bring any advantage for the endorsement of the concept of freedom as self-determination or self-regulation. Indeterminism substitutes the theoretically presumed closed determinism with the presumption that instead of a universal causality, chance is what is operational. From this presumption there is no direct way leading to a theoretical rehabilitation of a concept of free action in any sophisticated sense. That is why the proponents of compatibilism accuse incompatibilists that it is first of all through the theoretical position of a causally closed, pre-determined world occurrences, that philosophically the possibility could be imagined that in exactly this physicalistic world

an actor determines quasi him/herself in his/her actions. For proponents of compatibilism, actions are results of practical willing of preferences and goals of actions of the agents whose intention and preferences are to be understood theoretically as causally active determiners of the action even when, besides the preferences and intention of the acting subject, other factors which can never be ascertained from an observer perspective, condition the external action. But this makes not reversible, on the one hand, the ontological understanding of the action totally determined causally, partly by the acting agent himself and partly due to other factors. On the other hand, freedom should be understood in this framework as a way of self-determination, as a sort of causal self-influence. In doing this, it must be conceded in keeping with the conviction of compatibilism that only an incomplete knowledge of all causally active antecedent conditions of actions can be ascertained. Freedom in the sense of self-determination of actions is totally a part of the cause of actions in this understanding of causal determinism of actions.

For the status and content of the concept of freedom, it is important to note that the compatibilists as proponents of a physicalistic determinism as well as the proponents of indeterminism agree in a weak concept of freedom as self-determination, although the debate in theoretical philosophy about the possibility and reality of freedom does not arrive at any satisfactory conclusion. It is above all important to note that the concept of self-determination is given here a reading which, in the main, must be differentiated from Kant's concept of practical reason and his freedom of the will.

While Kant takes the non-empirical background, that is, in his language about the intelligible character of human beings morally acting persons – a perspective which, in accordance with our understanding of ethics, could be comprehended only from the perspective of the participating observer of an acting subject concerned about the moral rightness of his actions – the position defended by the compatibilists regarding freedom as self-determination which we have described here focuses on the extension of the theoretical observer's third person perspective. Within that perspective it should be verifiable externally that an actor is the cause of his/her own actions. This obtains in accordance with the methodical prerequisite of such a description when first the actor/agent in question does not act under external influence or compulsion (this should redeem the negative aspect of freedom)

and second when his action plans, wishes or intentions, like mediating instruments or partial causes, become the causal determination of his actions (this should redeem the positive aspect of freedom).

This version of freedom makes it possible to think of a constructive relationship of action execution, the acting agent and his action plan, even within a theory of a causally closed world of pure physical occurrences. But then, one can object that the concept of freedom as self-determination remains methodically closed for the ethical question regarding the morally right central access to the perspective of the actor/ agent in the first or second person. One can thus regard this version of the concept of freedom as a theoretically 'weak' theory of freedom in the context of a physicalistic oriented ontology for which – as much as I can tell – there is no compelling philosophical argument. This ontology is neither proven nor likely or probable. But nevertheless, from a moral point of view we can conclude that this ontology (even if unlikely) does not oppose the practical idea of freedom at all nor one can claim the emptiness of a more substantive concept of freedom which goes beyond the elements of causal agency and self-determination.

II. THE REALITY OF FREEDOM
AND ITS PHILOSOPHICAL IMPLICATIONS

2.1. Freedom and autonomy: Reflections about an often neglected difference

In the first part of my paper the focus of my interest was directed at the different perspectives of theoretical and practical philosophy concerning the question how to understand basically the *concept* of human freedom and its *possibility*. In reference to the ongoing discussions within philosophy we could state that a minimal concept of negative and positive freedom, understood as a missing of outer compulsion and as a kind of self-determination, is in a certain way agreed by almost all disputing parties (with the exemption of the so-called 'impossibilists'). I argued in addition to this statement that the semantic meaning of the more or less agreed concept of freedom is different concerning the different perspectives of theoretical and practical approaches. In the following second section of my paper I will focus now on the philosophical implications of the understanding of 'freedom' in practical philosophy because it is obvious that only from the perspective of the first and second person's view we may expect an appropriate or at least a more

comprehensive account of what is human freedom about: 'Human freedom' interpreted from the experiences of 'us' as beings who cannot avoid to interact permanently with others like us while living our life.

In the centre of Immanuel Kant's practical philosophy we find the philosophical statement of the practical *reality* of human freedom. That means not only that we cannot start to discuss problems of morals or politics if we do not presuppose a state of ontological reality of both: negative freedom *as well as* positive freedom. But additionally we cannot even conceive and describe 'human acting' as such if we do not understand ourselves as the 'free authors' of our own acting. That does include that the questions concerning 'intentionality', the conceptual difference between 'goals' and 'means', and the idea of a 'plan' behind single human acting – these and many other concepts we are accustomed to apply to human behaviour, language and action presuppose the implicit affirmation of human freedom. Here in Kant's philosophy we learn about the difference between the *'concept of freedom'* and the *'concept of autonomy'*. This difference is often ignored with misleading consequences not only with reference to a correct interpretation of Kant's texts but also to an appropriate understanding of the status of the reality of freedom. Concerning Kant the *'concept of freedom'* is a 'necessary postulate' of reason in theoretical philosophy or, in Kant's words, it is an 'idea of reason' which the faculty of human understanding is forced to accept as a reasonable postulate,[15] while practical philosophy is starting with the basic experience of us as practical acting beings.[16] That means that we as actors already know that we are acting free if there are no outer restrictions for our actions, if there is negative freedom. Kant's practical philosophy is starting from the experience of the reality of freedom. And without this there would be no proper practical philosophy even thinkable. On the other hand does the *'concept of autonomy'* explain how to execute human freedom in a morally correct way. Kant's answer to that question is referring to the 'reality of pure practical reason' ('Faktum der Vernunft')[17] in us and the necessity to focus all our practical interests on the normative force of 'pure practical reasoning', that means to let our practical will be determined by nothing else than our practical reason as its deciding factor: human practical reason giving the human by his/

---

[15] Kant, *Critique of Pure Reason*, ibid. III, 366 ff.
[16] Kant, *Critique of Practical Reason*, ibid. V, 19 ff.
[17] Ibid. V, 31.

her proper authority alone ('*autonomously*') the rules for her/his moral acting ('the norms' or 'laws', in Greek: '*nomoi*'). Only insofar as this constellation is *realized* Kant is talking about '*human autonomy*'.[18]

We can therefore conclude that the concept of autonomy is part of Kant's answer concerning the specific problem of morality, while the reality of freedom is already presupposed as essential for human practice at all. It has its general function to explain the specific character of human acting in difference from other occurrences in the world. We therefore can learn that the meaning of the term 'autonomy' does not simply coincide with the concept of 'freedom' in Kant, but the concept of 'autonomy' indicates the two specific dimensions: first the *reality of freedom* (in the above defined double meaning of negative freedom as 'non-domination' and the positive understanding of freedom as 'self-determination') and second *the reality of the effective moral orientation* of a human actor. Therefore the concept of 'autonomy' in Kant is not only pointing at the reality of moral freedom in us but does articulate that nothing else than 'reason' in its practical application, that means human reason concerning our specific *moral* acting and its principals. Reason is thought to be the only source of insight for morally correct acting. As a measure for moral orientation or a proof whether or not 'we' as human individuals are defined by pure reason alone in our moral orientation Kant offers the regulation by the so-called 'categorical imperative' and its rule of generalization concerning the maxims of our acting. Within the philosophical debates this argument of Kant was often criticized as a pure *formalist* criterion for ethically correct behaviour and acting. But a more comprehensive reading of Kant's texts allows to acknowledge that we can identify in the Kant systematic 'Metaphysics of Morals' are much more versions of the categorical imperative. Even in the earliest version of the 'categorical imperative' in the 'Groundwork for the Metaphysics of Morals' we are confronted with the Kant's postulate that we should never act in a way that we treat a human as a pure instrument. This postulate is thought to be a version of Kant's 'categorical imperative' and it seems to me that this postulate is one of the prominent and important formulas of Kant's 'categorical imperative'. The prohibition of a pure instrumental treatment of other humans does content not only a pure formal but a *material* dimension of Kant's concept of morality insofar as Kant's argument is referring to the concept of human dignity and

---

[18] Ibid., V. 33.

the statement that the human is an 'end in itself'. This formula of the 'categorical imperative' is obviously stating a universal, but substantive and *material*, although *negative* criterion for morality. In my above mentioned introduction into philosophical ethics I argued that it is this aspect which does allow to refer Kant's ethics today to a concept of basic and universal, 'negative' or 'prohibiting' human rights.[19]

2.2. 'Real freedom' and its conditions concerning human practice

On this background of a first clearance of the basic concepts in Kant's contribution to the questions concerning human freedom and moral autonomy I would like to refer now to the question how to develop a deeper understanding of the reality of human freedom in general. In this final part of my paper my considerations are obviously inspired by Herman Krings. In his essay 'Freiheit. Ein Versuch Gott zu denken'[20] we can find the suggestion first to widen the Kantian analysis of freedom into an intersubjective concept and second to take into consideration that the way how we realize the intersubjective structure of our freedom together with all the others is always in various ways qualified as a relative, contingent or even limited form of freedom. The insight in the 'relative' or 'contingent' way how we realize freedom in our practice does lead Hermann Krings to the postulate of a 'not relative' or 'not contingent' manner how freedom might be realized. This idea of a not contingent way to realize freedom enables us to reflect about the possibility of an 'infinite subject' of freedom by which the finite versions of freedom we are representing in our human practice are ultimately made possible.

As I have already shown in my previous argumentation the concept of practical freedom is not only an element of an ethical reflection but expresses more comprehensively 'the condition of being human as such'.[21] In the approach of Herman Krings, the practical concept of freedom takes the center stage in a philosophical anthropology which conceives the whole of humanity as rational beings capable of morality. From this

---

[19] Ibid., p.87.

[20] Hermann Krings, 'Freiheit. Ein Versuch Gott zu denken', in *Krings, System und Freiheit* (München/Freiburg: Alber Verlag, 1980), pp. 161-184. The translation of all texts of Hermann Krings quoted in this article are mine.

[21] Cf. Hermann Krings, Art. Gott, in *Handbuch philosophischer Grundbegriffe* (München: Kösel Verlag, 1973), Band III, pp. 614-641 (p. 634).

background, Herman Krings develops his extended analysis of freedom as 'autonomy' or 'self-determination'. Understood in this fundamental sense, the proposition that human beings autonomously determine themselves is, at least from a logical point of view, in a *contrary* opposition to external 'determination' (heteronomy) and compulsion and at the same time in a *contradictory* opposition to 'non-determination'.[22] Put differently: the autonomy of human beings which Krings conceives as the necessary condition for human beings to be in an unlimited sense humans, contains within it fundamental observations on two counts. *Firstly*, it implies negatively, that the self-determining person is free from external compulsion, be it of natural or social form. One can speak here of a concept of '*negative freedom*' in Krings' work. *Secondly*, it implies positively that the person has the capacity to give his free will a proper content or an individual entry which makes it a specific kind of that will. That represents the positive will and that will is coherent with the concept of '*positive freedom*'.[23]

It is the concept of positive freedom that helps Krings to distinguish the concept of freedom of the human will clearly from the concept of 'liberum arbitrium' of the tradition which does express just an *arbitrary freedom* which is indifferent with regard to all further agency content of the will. This differentiation is important with regard to the progression of the further analysis by Krings. In his argumentation the expression of autonomy as a self-determination of the human will means thus that 'the will gives itself its content'. In order to avoid misunderstandings Hermann Krings explains that the expression 'the concept of "self" is referring to a pure reflective determination, not a metaphysical claim'.[24] With this explanation, Krings draws attention on his own part to the epistemological status of his expositions. For him, they do not present any empirical oriented sociological or psychological 'description' of freedom in the outer world of physics or society neither do they present any 'ontological theory' of human beings and their 'proper nature'. They are rather to be understood as a *structural analysis of human freedom* in the sense of conceptual analytically proceeding reconstruction of the implications which are necessarily connected to the expression that human beings are autonomous acting beings or in my language 'free

---

[22] Ibid., p. 635.
[23] Cf. Hermann Krings, *Freiheit*, ibid., pp. 171 ff.
[24] Ibid., p. 173.

agents'. Krings himself speaks of a non-metaphysical and 'transcendental-reductive analysis of the structure of freedom'.[25]

The crucial and central point of the subsequent analysis of the structure of practical freedom as autonomy is its double or ambiguous character of the reality of freedom including with necessity both dimension: the negative and positive dimension of freedom. Otherwise freedom is according to Krings not conceived as 'real' in the comprehensive or full practical meaning of the word. This implies that according to Hermann Krings analysis the reality of freedom of an agent is conceptually defined, on the one hand, by the absence of heteronomous or external determination of the subject of freedom. On the other hand, real freedom is qualified through the further requirement of a real self-determination namely that the will gives itself its proper content. Otherwise we couldn't speak of human 'autonomy' in an unconditioned way. Exactly because the human will should not be misunderstood as arbitrariness, the question of this level of the analysis is focusing an understanding that the will does not choose, say, a specific object of desire, pursuit or arbitrariness. Hermann Krings is mainly concerned with ensuring that the will, without losing its autonomy in the situation of its practical application, 'opens' itself up 'initially' and 'decides' primarily for its own materiality or determinacy: 'The will must determine itself to itself as willing something specific.'[26]

This therefore qualifies the positive moment of freedom as autonomy or self-determination, that is, that the human will is to be thought in the reality of human acting not only 'independent from' every form of external determination, but 'the will' must also, beyond that, 'give itself a specific content'. The concept 'proposes that willing originally affirms itself as *willing something specific*'. Here we can see that a basic relationality of the human will and a dimension of self-reflexivity of freedom are identified as necessary elements of the reality of human freedom. And it becomes explicit but also an element of self-affirmation and positive recognition of the will. In the tradition of Kant and Schelling Hermann Krings describes that structure how the human will becomes real like an act of a 'transcendental self-determination'. It fortifies the empirical freedom of human beings in the contingent context of action *against* losing itself in limited conditions of external actions. In this sense,

---

[25] Hermann Krings, *Gott*, ibid., p. 639.
[26] Hermann Krings, *Freiheit*, ibid., p. 173.

Herman Krings is able to propose that 'transcendental freedom is *grounding* empirical freedom'.[27] In other words: Krings states that a comprehensive philosophical theory of the *reality* of human freedom (that means: a theory which is referring to the analysis of concrete actions and interactions of free human actors) is aware of and pointing at the double structure of negative and positive freedom and its conditions or prerequisites which are expressed conceptually by the basically relational and self-reflexive structure of human autonomy.

It is important to recognize that these conceptual elements of human autonomy are not ontologically 'just given' or are 'there' by nature or by the social order of human life, instead they have to be *realized* by ourselves *practically* in our concrete actions. Only if human autonomy is preserved in the practice of human acting, human freedom is real and is more than a possible thought. Therefore the important question has to be answered which proper content of the freedom can be identified that would help to avoid the always existing possibility (or 'danger') that the condition of 'human autonomy' is practically giving up itself in becoming an 'arbitrary will'. To put it in this way: By which of our decisions our 'free will' is 'opening itself' towards its 'specific content' and in doing so not giving up its identity as an autonomous or self-determining capacity of us? What content is of such state that it does not, like random object of knowledge or desire, contingently reduce human freedom and – in the best case – make it an arbitrary freedom? That is the decisive question of Hermann Krings. Kant's answer to this question on which he also reflected is well known as I mentioned above: For Kant, the so-called 'pure' and practical reason 'a priori' is the only candidate which the free will of human beings can chose without losing its autonomous status. Former interpreters of Kant argued that Kant's moral theory is able to avoid the problem to minimize human autonomy by reducing the content of moral decisions to a pure formal categorical decision making. But we have to recognize that this statement is not complete and is insofar incorrect as Kant is claiming too a material criterion for his ethics. This expressed in the imperative of the 'Groundwork' to avoid everything in human acting which might treat other humans as instrument. But nevertheless according to Kant each of us has to become obedient without exception to his or her personal practical reason and

---

[27] Ibid., p. 174.

it is this orientation alone by which autonomy of the will is reached and morality is made possible.

Compared with Kant, Hermann Krings puts the same question differently when he asks: Which 'content' is appropriate to the 'form' of the unconditioned and insofar 'transcendental' freedom? The will is not indifferently opposed to its content; but which content of the empirical autonomy in the sphere of concrete human acting – without limiting or conditioning its autonomy and in doing so minimizing or even destroying human autonomy? Hermann Krings answer to this challenge consists in the discovery of the freedom of the others: 'The satisfying content of freedom cannot, as long as it lives up to its form and dignity, be any other thing but freedom. This freedom is not, unlike we find in Kant, a monological personal freedom but rather the "freedom of the other".'[28] With this argument Hermann Krings does not only go beyond Kant's alleged 'empty formalism' but also determines an 'intersubjective content' of freedom. In another essay he expresses this thought as follows: 'It belongs to the condition of freedom, as long as it is understood as freedom oriented towards freedom, that it is originally affirmative and open to the freedom of the other. Only in such a resolve can it realize its necessary self-opening and self-determination implicitly and explicitly.'[29]

This intersubjective grounding of the necessary conditions for human autonomy and its logically necessary implications led Herman Krings also to a discussion of the discourse-theory, especially the outline of universal pragmatism by Jürgen Habermas and the concept of transcendental philosophical motive incorporating linguistic pragmatism of Karl-Otto Apel at the end of the 1970s. In that debate, Krings focused on the claims of freedom to necessary application or transcendental non-circumvention. In this time, Habermas connected this idea to the 'ideal communicative situation'[30] and Apel to 'a priori of the communicative community'.[31] Krings, in his analysis, agreed with the argument of Habermas which proposes that the idea of the ideal communicative community which is

---

[28] Ibid., p. 174.

[29] Hermann Krings, *Gott*, ibid., p. 636.

[30] Cf. Jürgen Habermas, 'Wahrheitstheorien', in Habermas, *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns* (Frankfurt: Suhrkamp Verlag, 1986), pp. 127-183.

[31] Cf. Karl-Otto Apel, 'Das Apriori der Kommunikationsgemeinschaft und die Grundlagen der Ethik', in *Apel, Transformation der Philosophie II* (Frankfurt: Suhrkamp Verlag, 1973), pp. 358-435.

contra-factually embedded in factual discourses should neither be misunderstood as an 'empirical phenomenon' nor as a 'mere construct'. But Krings objected to the assumption that effectually operative communicative process in Habermas is already enough to ground the demands of universal validity claim of, for instance, basic moral norms like we can identify with regard to the unconditioned normative claims of the human rights. Against Habermas, he argued that this analysis does not go beyond the insights bound to the factually occurring discourse. This holds true insofar as even within the later developed position of Habermas in his version of discourse-ethics there are finally no other criteria given for moral reasoning and the justification of moral norms than the 'interests' of the participating individuals.[32] 'Against the idea', Krings wrote, 'that there is no objection to the thought that the validity claim in relation to the factum of the communicative speech act to be identified and reconstructed; it could be objected that with it validity as such should be justified.'[33] In his argument Herman Krings was pointing at the manifest problem of an unsatisfactory circularity of argumentation in Habermas and Apel. This consists in the fact that the authors of discourse theory tie the conviction of the necessity of moral norms to the success of a communicative practice in which these norms already apply. This claim of linguistic pragmatics leads to error when it tries 'to a certain extent to be itself in its own meta-system'.[34]

Even if not all comments of Hermann Krings concerning the position of Habermas and Apel are convincing, the main argument of Krings is according to my understanding nonetheless still interesting. The reason for this consists in the fact that he clearly gives here to understanding and justifies with his so-called 'transcendental analysis' of human freedom and its necessary conditions for its implementation in our morally correct practice also the demand for necessary validity of morality. This is so because the affirmation of one's and other's freedom receives in Hermann Kring's analysis the necessary practical validity claim which Kant claims for pure practical rational insight and his understanding of what he called 'the good will'. According to Krings, it is first from this, clearly in opposition to Kant, extended transcendental insight,

---

[32] See in my introduction into ethics, ibid., pp. 94-113.
[33] Cf. Hermann Krings, 'Reale Freiheit. Praktische Freiheit. Transzendentale Freiheit', in Krings, *System und Freiheit*, ibid., pp. 40-68 (p. 64).
[34] Ibid., p. 65.

that, in reality, proceeding from freedom, it should be possible to ask the fundamental moral question regarding moral difference (sittliche Differenz), that is, the question regarding good and bad. For Krings, this understanding is tied to his analysis of the concept of transcendental freedom: 'The transcendental philosophical meaning of the concept of necessity as justificatory instance for the moral character of actions points to the structural unity of real freedom, practical freedom and transcendental freedom. It serves mainly the function of making conceivable the profound structure of the problem of freedom.'[35]

2.3. The inner ambivalences of 'real freedom' and its positive relation to an 'idea of God'

The invention of the necessary intersubjective structure of real freedom is an important contribution to the philosophical debate on freedom. It contains a correction of Kant intuitions who defined moral autonomy in terms of an inner-subjective reflection with the consequence that political autonomy was seen as a matter of legal obligation of agents and their outer interaction, that means as a matter not of their inner will and conscience but of coercion by a legitimate law which Kant defined by the expression of the common will of all in a given political or societal community. In opposition to Kant Hermann Krings is arguing in favour of a concept of *intersubjective freedom* claiming the necessity to recognize the other agents as equally free and in their autonomy unconditioned other beings. The affirmation of the real freedom of others with whom we cannot avoid to interact is not any longer, like in Kant's Philosophy, a matter organized by the system of a legal order of coercion and focusing on the outer behaviour but a necessary postulate without which no single person is able to preserve his or her own autonomy and free will in the reality of its concrete personal acting.

With this argument Krings is able on the one hand to overcome the peculiar difference between the moral philosophy and political philosophy in Kant which is a final consequence of his epistemological distinction among the realm of the 'appearances' and the so-called 'thing in itself' as well as his talk of the human as a 'civic of two (separated) kingdoms'. On the other hand Krings has to address the problem that the logical necessity for 'us' humans to realize our freedom by acts of

---

[35] Ibid., p. 67.

an unconditioned recognition of the freedom of the others is never realized in practice without severe limitations. Here we can learn about an inner tension between the claim for an unconditioned preservation of freedom by an unlimited form of recognition of other freedom and the structurally always limited mode how this claim is realized each time given the practical circumstances and conditions of our practical lives. Not to speak about those environments or situations in which the freedom of others is not only structurally limited, but even intentionally harmed or the humans as the subjects of their freedom are themselves attacked, injured or destroyed by others. All these experiences of realities and facts are the reason for Krings to admit that his idea of 'real freedom' or 'human autonomy' is never fully realized in human practical life; but this insight doesn't change the unconditioned demand and postulate to acknowledge the freedom of others without restriction and vice versa. Insofar as we know the history of mankind and the practice of our given societies and political systems we can therefore conclude that the postulate of real freedom is something which entails a never ending program for political reform and societal change and the demand for each of us to orient our moral intuitions accordingly.

But from here it becomes clear that the content of real freedom as part of human practice is something we should talk about in the comparative way of 'more' or 'less'. This was taken into account by Herman Krings when, regarding the material aspect of the freedom of the will, he underscored: Freedom as the original determination of content has so much *relationality* and dignity like every terminal content to which it opens itself up. Freedom is therefore realized in human history and practice in different ranks, even if understood unconditioned in its form as the pure figure of autonomous self-determination and self-opening in reality. This qualifies the freedom of human willing and action always as a '*finite freedom*' which I want to refer to as freedom which remains in a certain way fundamentally *incomplete* in the practice of its self-realization. It receives on the one hand normative unconditionality and additionally logical necessity in keeping with the concept of transcendental logical analysis of its components and on the other hand factual finitude and subsisting imperfection in keeping with the degree of its realization. The two perspectives constitute a deep ambivalence or inner ambiguity of human autonomy. The reason for this consists in the fact that real freedom of human beings is realized under contingent conditions with the consequence that no structure of

realization of freedom can ever resolve or fully realize the normative idea of an unlimited necessity affiliated to the transcendental analysis of its unconditionality. This leads the philosopher Hermann Krings to the following reflection: 'But in whichever degree it [freedom – LB] choses to realize itself, no (human) realization ever exhausts the form of self-opening which is unconditioned or limited by nothing. From this unconditionality of the form, it follows that no content, not even the adequate content [=the other freedom], can ultimately set a limit to freedom which could to a certain extent disrupt the act of self-opening and resolve.'[36]

From this double structure of human freedom as '*unconditioned freedom*', in accordance with its concept and its normative necessity on the one hand, and as '*contingent freedom*', in accordance with the realization of its content on the other hand, Krings derives the idea that finite freedom is always *internally oriented* towards the shape of an unlimited realization of unconditioned freedom if it does not want to lose itself and give up the normative claim to unconditioned autonomy. Krings is stating that human freedom is in all stages of its realization 'already directed' towards the structure of realized *absolute freedom* in the mode of a dynamic practice of 'anticipation'. One could speak here of an argument which makes both the internal tension of freedom as an unconditioned postulate and its always conditioned way of realization explicit and in doing so is pointing at a dynamic process of the '*transcendence from within*'[37] which means nothing else than the practical human search for real freedom in our daily practice.

The finite freedom of human beings realizes itself in accordance to its structure, as Krings attempted to show, in affirming other's freedom and this affirmation takes on an unconditioned claim which cannot be invalidated by anyone. But this freedom does not realize itself other than 'in anticipation of unconditioned freedom' because, for other human beings as subjects or bearers of freedom, it also applies, as was determined for the realization of freedom, that they also can only realize their freedom in an imperfect, contingent and never exhaustive way. To this end Krings'

---

[36] Krings, *Freiheit*, ibid., p. 175.

[37] Cf. the Habermasian use of this term in his articles: 'Einen unbedingten Sinn zu retten ohne Gott, ist eitel', firstly published in: Matthias Lutz-Bachmann/Gunzelin Schmid Noerr (eds.), *Kritischer Materialismus* (München: Hanser Verlag, 1991), pp. 125-142 and 'Exkurs: Transzendenz von Innen, Transzendenz ins Diesseits', in Habermas, *Texte und Kontexte* (Frankfurt: Suhrkamp Verlag, 1991), pp. 127-156.

expressed '*anticipation* of unconditioned freedom'[38] aims at the *idea of a bearer of freedom* who realizes by himself/herself the normative idea of freedom as autonomy in an unlimited and insofar 'absolute' way. Should other people in their freedom therefore, as philosophically demanded, be affirmed in an unconditioned way, that would make it according to Krings in no other way possible and realizable besides in the form of the anticipation an 'absolute realization of freedom': This is named by Krings in the tradition of philosophical theology with the concept of 'God'. It is this kind of an unconditioned and unlimited way to realize freedom by which the limited and conditioned way how humans realize their freedom is only made possible. Otherwise the idea of autonomy would be an empty postulate and human practice would fail in its attempt to realize intersubjective freedom under the factual conditions.

Hermann Krings makes haste to add the following explanation in order to prevent possible or, in fact, very related misunderstanding. He does not wish to imply that the concept of an absolute freedom or God should be thought of and affirmed here as anticipation reaching out for an amplification of the concept of freedom neither is it an emphatic idealization or a lavish use of concept because, by absolute freedom, he does not mean an *object* which could be thought of as objective and as such conceivable in the mode of 'via eminentiae'. Instead, the concept of an absolute freedom refers for Hermann Krings to real relation or actual relation of necessarily required realization of the freedom of human will to a necessary anticipation of the finite freedom of the human beings to another subject of freedom, which, not being itself outside the dynamics of freedom and its realization, cannot be thought of outside of this relation.

The concept of absolute freedom results – like the entire discussion – this way for Krings strictly from the conceptual-reductive method of his transcendental-logical analysis: from the 'return to the formal unconditioned character of finite freedom'. For Krings, the idea of an absolute freedom – thought of as that content which alone has the capacity to fulfil and make possible the unconditioned claim of self-opening of finite freedom to other's freedom – necessarily results from the tension of the unconditioned form of self-opening of freedom as well as the required affirmation of other's freedom on the one hand and the always contingent and conditioned way to realize freedom for all bearers of finite freedom on the other hand. When alone such an idea

---

[38] Krings, *Freiheit*, p. 176.

of absolute freedom of the required unconditioned openness of all freedom to other freedoms is adequate, according to Hermann Krings, whether we know it explicitly or not, every other factual openness and affirmation of the other's finite and imperfect freedoms in the mode of anticipation of perfect freedom follows likewise. The presupposition of this begins the conceptual development of Krings' closing thought in alignment to Augustine: 'Augustine's cor inquietum does not lead to the thought of absolute rest but rather to the concept of that which satisfies cor inquietum and brings it to requiescere.'

The notion of God expressed through the concept of an absolute freedom makes it possible, according to Krings, 'to think of God' *without* making recourse to his ontological state as a first 'being' or his 'attributes' in an approach within theoretical philosophy. In analogy to Anselm of Canterbury's introduction of the concept of God in the mode of a logical-relational progression as a 'quo maius cogitari non potest' in his 'Proslogion',[39] Hermann Krings points out that his concept of God in addition should be thought of within a practical relation in which freedom becomes real alone. This is called by Hermann Krings a practical '*relation* quo maius cogitari non potest'.[40] And he adds the remark: 'Important is that especially not a concept of unparalleled being should be thought about, but rather simply the *fulfilment* of the original meaning of freedom.' In this sense the idea of an absolute freedom is '*as much as freedom can be a necessary thought*'.[41]

In other words: The philosophical concept of God is thought of by Hermann Krings as the necessary idea of the *absolute realization of freedom* without which the finite history of our human freedom and its practice cannot be thought of – even though it is imperative and an unconditioned postulate of our practical reasoning – as the necessary requirement of an unconditioned mutual recognition of external freedom by human beings. Without the philosophical affirmation of the idea of God as the absolute realization of freedom the finite actors like us could not continue to preserve the imperatives of a relational concept of freedom as autonomy since one could not see that anyone could fulfil its demands. This negative insight would lead to the statement that the practical concept of freedom must necessarily fail.

---

[39] Cf. Anselm of Canterbury, Proslogion, ed. F. S. Schmitt (Stuttgart: Frommann-Holzboog Verlag, 1984), pp. 84-89.

[40] Krings, *Freiheit*, ibid., p. 177.

[41] Ibid.

METAPHYSICAL SUPERNATURALISM AND MORALLY WORTHY ACTIONS

BRUNO NIEDERBACHER SJ

University of Innsbruck

Abstract. This article is an attempt to solve the question whether there is a version of metaphysical supernaturalism that grants both: first, that moral facts depend in a metaphysical strong way on God, and second, that agnostics and atheists are nevertheless able to perform morally worthy actions. The solution that is developed in this paper builds on a distinction between the proximate and the remote goodmakers of actions. It is argued that the proximate goodmakers of actions can be cognized also by the non-believer and that such knowledge or justified belief of the proximate goodmakers might be sufficient to perform morally worthy actions.

I. INTRODUCTION

Metaphysical Supernaturalism in Metaethics is a position according to which moral properties depend metaphysically on supernatural properties and moral facts depend metaphysically on supernatural facts. Proponents of such accounts of moral properties often insist that their claims are sheerly metaphysical and pertain only to the nature of moral values. They claim that these metaphysical claims do not imply that agnostics or atheists would be unable to perform morally worthy actions or be unable to have justified true beliefs about which actions are right or wrong. William Alston writes:

> What makes this table a meter in length is not its conformity to a Platonic essence but its conformity to a certain existing individual [footnote omitted]. Similarly, on the present view, what ultimately makes an act

of love a good thing is not its conformity to some general principle but its conformity to, or approximation to, God, Who is both the ultimate source of the existence of things and the supreme standard by reference to which they are to be assessed. [...] the view does not have the alleged epistemological implications. [...] The particularist is free to recognize that God has so constructed us and our environment that we are led to form sound value judgments under various circumstances without tracing them back to the ultimate standard. Analogously, we are so constructed and so situated as to be able to form true and useful opinions about water without getting so far as to discern its ultimate chemical or physical constitution, without knowing what makes it water. (Alston 1990: 320-322)[1]

However, this thesis might be challenged. One might contend that, in order to act in a morally worthy way, one ought to act for the right reasons. Let us call this the Right Reasons View. Acting for the right reasons implies having some grasp of these reasons. Moreover, having some grasp of these reasons implies not only having justified true moral beliefs such as *I ought to keep my promise*, but also having justified true beliefs about, or even an understanding of, *why* one ought to keep one's promise; that is, one must have true justified beliefs about, or even understanding of, what makes the action morally right or wrong. In this sense Alison Hills writes that, in order to perform a morally worthy action, 'you need to act for the reasons that make your action right' (Hills 2009: 117). Thus, the Right Reasons View is a normative claim according to which an agent's action *a* is morally worthy if and only if

   (i) it is right to perform *a*;
  (ii) the agent has the justified belief that it is right to perform *a*;
 (iii) the agent has the justified true belief that, or even
         an understanding of why, *a* is made right by feature *F*;
  (iv) the agent acts because of (ii) and (iii).

Now, if it is a theistic fact that makes actions right or good, or if a theistic fact is an essential part of what makes actions right or good, it would follow that people who do not believe in this fact could not act in a morally worthy way. Thus, the challenger would say: If Metaphysical Supernaturalism is true, then a person can perform morally worthy

---

[1] See also Adams 1999: 355: 'The crucial point is that my theory [...] does not require us to know anything about God, as such, before we can have knowledge, or adequately grounded belief, in ethics.'

actions only if that person has some specific theistic beliefs. Metaphysical Supernaturalism plus the Right Reasons View would thus have strong epistemological implications.

Of course, the thesis that an agent must have a grasp of what we might call the 'goodmaker' of her action in order for that action to be morally worthy is contentious, at least when stated in this general way. However, in what follows I will – for the sake of argument – assume the truth of this thesis. And I will examine whether – given the Right Reasons View – there is a kind of Metaethical Supernaturalism which grants both: (i) that moral values depend in a metaphysically strong way on God, and (ii) that agnostics and atheists are nevertheless able to perform morally worthy actions.

II. PRESUPPOSITIONS AND DISTINCTIONS

In what follows I will assume some version of Metaethical Cognitivism plus Moral Realism. I will presuppose that some moral utterances are moral assertions, that moral assertions are expressions of moral beliefs, that moral beliefs – or better: their propositional content – are truth-bearers, and that true moral propositions are made true by moral facts. Moreover, it might be helpful to introduce a few distinctions:

(1) I will distinguish first between (i) the 'goodmaker' of an action, (ii) the 'truthmaker' of a moral proposition, and (iii) the 'justifiers' of a moral belief:

(i) The goodmaker of an action is the (non-moral) property or the set of (non-moral) properties of the action that make it morally good. For example, what makes Fred's action of keeping his promise to Emma morally obligatory? One's answer to this question will, of course, depend on the normative theory one favours. A Utilitarian would say: It is the property of maximizing pleasure on earth that makes the action morally obligatory. A Kantian would perhaps say: It is the property of regarding the other as an end and not merely as a means that makes the action morally obligatory. An Aristotelian Perfectionist might say: It is the property of contributing to human flourishing that makes the action morally obligatory. And a Divine Command Theorist might say: It is the property of being commanded by the loving God that makes the action morally obligatory.

(ii) The goodmaker of an action has to be distinguished from the truthmaker of a moral proposition. What makes true the moral proposition that Fred's action of keeping his promise is morally obligatory? Proponents of Metaethical Naturalism think that Fred's belief is made true by a natural fact, say by the fact that keeping his promise maximizes pleasure, or contributes to human flourishing, etc. However, given that I favour a version of Non-Naturalism, I claim that Fred's belief that he ought to keep his promise is made true by the moral fact that he ought to keep his promise. The fact that his promise-keeping maximizes pleasure, or contributes to human flourishing, etc., is a different fact.

(iii) The goodmaker of an action has to be distinguished, at least in principle, from the grounds or reasons that justify moral beliefs. What justifies Fred's moral belief that keeping his promise is morally obligatory? A foundationalist might say: Fred's moral belief that he ought to keep this particular promise to Emma is justified by, among other beliefs, his belief that one ought *prima facie* to keep one's promises, a belief that is self-evident for Fred. A coherentist might answer: Fred's moral belief is justified by cohering with his moral belief-system. And a virtue reliabilist could say: Fred's moral belief is justified by being generated by the reliable cognitive moral faculty traditionally called 'phronesis'.

The distinction between the goodmaker of an action and the epistemic ground or reason for a moral belief is sometimes overlooked. One reason for this is that the goodmaker of the action might feature in any beliefs which might feature in the justification of the target moral belief. For example, Fred's moral belief that he ought to keep his promise might be justified by his beliefs that

(a) one ought to do whatever contributes to human flourishing;

(b) keeping my promise contributes to human flourishing.

As one can see, the goodmaker of the action features in the beliefs that justify Fred's target belief. Nevertheless, the goodmaker of the action is not identical with the belief about the goodmaker.

The second reason why the distinction between the goodmaker of an action and the justifier of a moral belief tends to be overlooked stems from the fact that some normative views include the condition that the agent must have a justified true belief about what makes the action good among the necessary conditions for morally good actions. The Right Reasons View is such a view, saying that an action is morally good only

if the agent has justified true beliefs about, or even understanding of, the goodmaker and acts because of them. Thus, on the Right Reasons View, the belief or understanding about the goodmaker of the action is part of the goodmaker of the action.

(2) The Right Reasons View distinguishes between doing something right and doing something right for the right reasons.[2] Thus if Fred keeps his promise but does it for the wrong reasons, he would be doing something right. But he would fail to do it for the right reasons and thus his action would not be morally worthy.

(3) We must distinguish between the moral value of action-types and action-particulars. In what follows I will consider only action-particulars, such as Fred's action of keeping the promise he has given to Emma at a particular time under particular circumstances. Furthermore, I will restrict the talk to human action. I will refer to such a particular human action with the symbol 'a'.

(4) 'God' is here used in accord with the traditional monotheistic religions, which conceive of God as a singular, eternal, immaterial, transcendent, personal, loving, almighty, omniscient, and perfectly good entity who created and sustains the universe and reveals himself to human beings. A believer in such a God might be called 'theist'.

III. KINDS OF METAPHYSICAL DEPENDENCE

Metaphysical Supernaturalists could claim:

> The action a has moral status M only if a bears the metaphysical relationship R with God (cf. Quinn 2006: 68).

The expression 'moral status M' can stand for a deontological status like *being obligatory, allowed, forbidden*; for an axiological status like *being morally good, morally bad*; or for a thick normative status like *being just, unjust, contemptible, pitiful, rude, kind, brutal*, etc.

The expression 'bears the metaphysical relationship R with God' can stand for *being commanded by God, being commandable by God, being in accord with/against God's command, being in accord with/against God's will, being what God/Christ would be motivated to do in like circumstances,*

---

[2] See: Hills 2009: 113: 'There is a well-known distinction between doing the right action and acting well or performing morally worthy actions. Your action is morally worthy only if it is a right action performed for the right reasons [...].'

or *being Godlike*, etc. I will not ask here which specification would be better, but will ask at a general level how the metaphysical relationship between what is stated on the left side and what is stated on the right side can be thought of. For the sake of simplicity I will abbreviate the expression on the left side with 'M-expression' and what it refers to with 'M-property', and the expression on the right hand side with 'SN-expression' and what it refers to with 'SN-property'. Three claims about their relation can readily be stated:

(1) The M-expression has the meaning given by the SN-expression.

(2) The M-property is identical with the SN-property.

(3) The M-property is grounded in the SN-property.

In what follows I will explore these claims. I will assess which of them is able to grant both (i) that moral values depend in a metaphysically strong way on God, and (ii) that agnostics and atheists are nevertheless able to perform morally worthy actions. For the sake of simplicity I will pick out in my examples the expressions 'being morally obligatory' and 'being in accord with God's command'.

3.1. The M-expression has the meaning given by the SN-expression.

Let us take the following two sentences:

(1) 'Fred's action of keeping his promise is morally obligatory.'

(2) 'Fred's action of keeping his promise is in accord with God's command.'

According to the thesis under consideration, sentence (1) means the proposition expressed in sentence (2), and, if the sentences (1) and (2) are true, they state the same fact, namely the fact that Fred's action of keeping his promise is in accord with God's command. This is a metaphysically reductionist position which starts with the semantic claim that 'morally obligatory' is defined as being in accord with God's command. Thus, the concept of God is thought to be part of the concept of moral obligatoriness. Let us call this position Analytical Supernaturalism, which has the following implications: First, the proponent of this view would have to think that someone who lacks the concept of God would not properly understand what 'morally obligatory' means. Second, the Analytical Supernaturalist would have to think that an atheist – who has the concept of God but believes that there is no God – would have to believe that the concept of being morally obligatory is empty, that there

are no actions that are morally obligatory (cf. Audi 2007: 122). And the Analytical Supernaturalist would have to think that an agnostic – who has the concept of God but refrains from believing that there is a God – could have no justified true belief that an action is morally obligatory. Thus, the Analytical Supernaturalist would have to think that atheists and agnostics can have no justified true belief about an action's being morally obligatory. Analytical Supernaturalism would thus have strong epistemological implications. And since atheists and agnostics could have no justified true beliefs about an action's being morally obligatory, they could not act because of such beliefs, and, *a fortiori*, they could not act for the right reasons. Thus they could not perform morally worthy actions.

It has been suggested to me that there is a version of a theory that grants both Analytical Supernaturalism as well as the possibility of atheists and agnostics performing morally worthy actions.[3] Atheists and agnostics could have a justified belief of the following sort:

The action *a* would be morally obligatory if, counterfactually, there were a God.

Atheists and agnostics could act because of such a belief, and could thus perform morally worthy actions. This is a possibility, but a restricted one. It is restricted to supernaturalist theories that spell out the SN-property in terms of possibility instead of actuality; that is, in terms of *being commandable by God* or *being what God would will* or *being what would be in conformity with God*, etc. Moreover, I doubt whether actual atheists and agnostics act for such reasons. But I nonetheless do not doubt that many of them have justified true beliefs about which actions are morally obligatory, and that many of them perform morally worthy actions. Therefore, I think that this solution is not compelling.

3.2. The M-property is identical with the SN-property.

Let us take our two sentences again:

(1) 'Fred's action of keeping his promise is morally obligatory.'

(2) 'Fred's action of keeping his promise is in accord with God's command.'

According to the identity thesis, sentences (1) and (2) differ in meaning but state, if true, the same fact. Let us call this position Non-Analytical

---

[3] I thank Katherine Dormandy for this suggestion.

Supernaturalism. There are two ways to understand the identity claim.
The strong version understands it as a Constitutive Identity Claim.
Proponents of this version often use analogies with scientific discoveries
to make their point, for example: Water = H_2O;[4] an example which is
nearer to our purposes because it is not about natural kinds but rather
about properties: being hot = having molecular kinetic energy. The idea
is then the following: As the nature of water *is* H_2O, and as the nature of
being hot *is* having molecular kinetic energy, in the same way the nature
of being morally obligatory *is* being in accord with God's command.

The second, weaker, version understands the identity claim in terms
of mere co-extensiveness. We can call it the Co-extensiveness Claim;
it says that the M-expression and the SN-expression are co-extensive.
However, we do not discover something about the nature of the property
of being morally obligatory by understanding it in terms of being in
accord with God's command. Think of proper names. Cicero is the same
person as Tully. 'Cicero' and 'Tully' are co-extensive. However, we do not
learn anything about the person's nature or character when we discover
that Cicero is also called Tully (cf. Audi 2007: 122).

According to the Constitutive Identity Claim a person can, of course,
justifiably believe that an action is morally obligatory without believing
that it is in accord with God's command. After all, a person can know or
justifiably believe perfectly well that some particular piece of iron is hot
even if this person has no beliefs whatsoever about molecular kinetic
energy. In a similar way, proponents of the Constitutive Identity Claim
could say: A person can justifiably believe that an action is morally
obligatory without having theistic beliefs. An atheist or agnostic can
justifiably believe that his promise-keeping action is morally obligatory.
Somebody favouring the Constitutive Identity Claim could conclude
that atheists and agnostics can perform morally worthy actions because
they can act on the basis of the belief that the action is morally obligatory.

However, there are two drawbacks to this account: First, it is obvious
that on the Constitutive Identity Claim agnostics and atheists would not
have true beliefs about the nature of the moral properties. They would
not know what the true nature of being morally obligatory is and thus

---

[4] Adams 1979/1999: 415: 'My new divine command theory of the nature of ethical
wrongness, then, is that ethical wrongness *is* (i.e. is identical with) the property of being
contrary to the commands of a loving God.' In Adams 1999: 29, he proposes the identity-
thesis also for axiological properties according to which *being excellent* is identical with
being Godlike.

would not know what 'makes' the action morally obligatory. Therefore, they could not act for the right reasons, if acting for the right reasons implies having some justified true beliefs about what makes an action good or bad, and the problem would remain.

Second, deeper problems lurk beneath the Constitutive Identity Claim. I think Derek Parfit is right when he says that such an important scientific discovery as

(i) Having molecular kinetic energy *is* being hot

states something about the relation among *several* different properties. This becomes clearer when we consider that 'being hot' expresses a complex concept, namely:

the property that can make objects have certain other properties, by turning solids into liquids or liquids into gases, causing us to have certain sensations, etc. (Cf. Parfit 2011: vol. II, p. 335)

Thus, claim (i) must be restated as: Having molecular kinetic energy is the property that can make objects have these other, different properties. If we apply this to our context we would have to say: Like 'being hot', the term 'being obligatory' expresses a complex concept, namely:

the property that makes an action have certain other properties, such as being a reason to act, having moral worth, being subject to praise, etc.

Thus, the claim

Being in accord with God's command *is* being morally obligatory

states a relation between *different* properties. The claim has to be restated as: Being in accord with God's command is the property that makes an action have these other, different properties.

If this reasoning is sound, the Constitutive Identity Claim would be in trouble. The Metaphysical Supernaturalist would have to find a different interpretation of the phrase 'being in accord with God's commands *makes* an action morally obligatory', an interpretation that does not amount to property identity but rather to a sort of constitution without identity.

How does the Co-extensiveness Claim fare in this regard? Recall that it is the claim that whatever is morally obligatory is in accord with God's command, and that whatever is in accord with God's command is morally obligatory. One reading of this claim has been proposed by Robert Audi. He says:

Suppose initially that we take the property of being obligatory to be the same property as that of being divinely commanded. Instead of stopping there, however, we might take '*both*' properties (i.e., the property expressed by the theological phrase 'divinely commanded' and the property expressed by the non-theological phrase 'being obligatory') to be (necessarily) consequential, in a strong sense, on non-moral, 'natural' properties belonging to the type of obligatory act in question [footnote omitted] [...] These natural (roughly 'descriptive') properties are the same ones central for understanding moral concepts and moral properties outside theological contexts. This is as it should be on the plausible assumption that properties F and G (as expressed by different terms, such as 'commandedness' and 'obligatoriness') are identical only if anything possessing them has them *in virtue of* the same property or set of properties [footnote omitted]. In rough terms, they are identically grounded. (Audi 2007: 123)

On this reading, agnostics and atheists can perform morally worthy actions because the goodmaker of the action is a natural property or a natural fact, and agnostics and atheists can have justified true beliefs about such facts. The price of this solution is, however, that morality – or at least a major part of it – does not metaphysically depend on God. Thus, the Metaphysical Supernaturalist will probably refrain from embracing this view and will instead look for a metaphysical reading of co-extensiveness which grants that the property of being obligatory depends metaphysically on God. A first step in his argument could be the negative claim that necessary co-extensiveness does not imply property identity. Let us take the following two terms: 'being an equilateral triangle in Euclidean space' and 'being an equiangular triangle in Euclidean space'. Both terms are necessarily co-extensive, that is, every triangle which has the property of being equilateral additionally has the property of being equiangular.[5] But why suppose that the two properties

---

[5] Jackson, (forthcoming), writes: 'But surely both properties are a certain shape, and we don't have two shapes. That is, what is true is something like: being an equilateral triangle in euclidean space = shape S, and being an equiangular triangle in euclidean space = shape S. But then the transitivity of identity delivers the conclusion that being an equilateral triangle in euclidean space = being an equiangular triangle in Euclidean space. What we have aren't two properties but two different ways of representing the same property.' I think that Jackson's argument shows only that the two triangles are identical but it does not show that the two properties are identical. They are properties of the same kind of triangle but not the same properties.

are identical? They are properties of the same kind of triangle but they are not the same properties. A second example: The terms 'being the only even prime number' and 'being the positive square root of 4' are necessarily co-extensive. Necessarily, they refer to properties that are had only by the number 2, but they do not refer to the same property.[6] Thus, there is good reason to suppose that the necessary co-extensiveness of the terms 'being obligatory' and 'being in accord with God's command' does not amount to property identity. The second step in the Metaphysical Supernaturalist's argument could then be this: He takes up Audi's idea of grounding, but in a different way. He claims that M-properties are grounded in the SN-properties; or to state it more precisely: He claims that the fact that the action *a* has an M-property is grounded in the fact that *a* has the corresponding SN-property. This is a kind of constitution without identity.

3.3. The M-property is grounded in the SN-property.

Let us take our two sentences again:

(1) 'Fred's action in keeping his promise is morally obligatory.'

(2) 'Fred's action in keeping his promise is in accord with God's command.'

According to the claim under consideration – let us call it the Grounding Claim – the sentences (1) and (2) differ in meaning, and if true, they state different facts. The fact that Fred's action is in accord with God's command makes it the case that Fred's action is morally obligatory. The fact that Fred ought to keep his promise is grounded in the fact that Fred's keeping the promise is in accord with God's command. There are two ways to understand this claim. The first way is in terms of causality. The fact that is stated by (2) causally explains the fact that is stated by (1). By commanding, God causes the action to be morally obligatory. The second way to understand the Grounding Claim is in non-causal terms.[7] This is what recent literature generally understands with the word 'grounding': a non-causal metaphysical relation between facts.[8]

---

[6] This example is given by Parfit 2011, Vol. II: 297.

[7] See Wierenga 1983: 389, who holds that the formulation 'makes it the case that' does not express a causal relation but rather some other asymmetric relation of dependence. Quinn 2006: 70, leaves both options open because he does not know of a conclusive reason for preferring the one version to the other.

[8] For detailed analyses of the concept of grounding see: Correia and Schnieder 2012.

In our talks we find many statements that express such relations, e.g., 'What he did is punishable because it is against the law' or 'Complexes exist because simples exist'. The 'because' in these statements refers to a non-causal relation. It is not the case that the law causes the action to be punishable, nor is the case that the existence of simples causes the existence of complexes. In the same way, Metaphysical Supernaturalists might understand the statement: 'Fred ought to keep his promise because doing so is in accordance with God's command.'

Proponents of both ways of cashing out the statement might say that their version lacks strong epistemological implications. They might claim that a person can justifiably believe that an action is morally good without having beliefs about what makes it morally good. But of course this will fail to satisfy proponents of the Right Reasons View. Although one can justifiably believe the one without the other, one cannot act in a morally worthy way without having justified true beliefs about the goodmaker of that action. Let us assume that Fred keeps his promise to Emma because he believes that

  (i) it is morally obligatory to keep my promise to Emma; and
  (ii) it is morally obligatory because keeping the promise contributes to human flourishing.

Let us further assume that the Metaphysical Supernaturalist endorses the Right Reasons View. How could he evaluate Fred's action? He could say that Fred is doing the right action for a wrong reason. Fred should have believed that

  (i) it is morally obligatory to keep my promise to Emma; and
  (ii*) it is morally obligatory because keeping it is in accord with God's command.

Thus he would have to admit that Fred does not perform a morally worthy action. But our Metaphysical Supernaturalist might take a different route and not share the Right Reasons View entirely. He could accept degrees of moral worth and thus say: Fred performs a morally worthy action, but he does not perform a morally perfect action. Following the scholastic tradition, one could also say: In such a case, Fred's action is good *secundum quid*, but not good *simpliciter*, that is, his action is good only under one consideration but not all things considered.

There are more plausible ways in which the Metaphysical Supernaturalist can embrace the Grounding Claim. One possibility is that the Supernaturalist could accept that there are two kinds of goodmakers of actions: proximate ones and remote ones. Fred's belief (ii) would refer

to a proximate goodmaker of the action. This does not exclude that there is a remote goodmaker that makes the action finally good. Thus, it could still be the case that the remote or last goodmaker of the action is some theistic fact. But Fred does not have to have beliefs concerning this remote goodmaker in order to act in a morally worthy way.

At this point we must ask how the proximate and the remote goodmakers relate. I can think of two ways of spelling this relation out. First, the Metaphysical Supernaturalist accepts that moral properties are grounded in natural properties, such as being a means, being a constituent of human flourishing, etc. Thus he accepts that what makes an action morally good or bad are the action's natural properties. He accepts that Fred does the right thing for the right reasons when he acts from the beliefs (i) and (ii). The supernatural claim comes in as an explanation of the grounding relation. This claim offers an answer to the question: Why is the fact that the action a has certain particular M-properties grounded in the fact that the action a has the N-properties ('N' stands for 'natural') which it does? The Supernaturalist's answer would be: Because this relation is established by God, for example by wanting humans to flourish. On this view, Fred could have a justified true belief about what is morally obligatory, as well as a justified true belief about what makes the action morally obligatory. If Fred acted for these reasons, he would perform a morally worthy action. The only thing Fred would lack would be true beliefs concerning the explanation of why certain natural facts make actions good or bad.

However, there is a drawback to this account. If the relation between the two kinds of fact is understood as 'established by God', it seems that this account leads directly to the first horn of the Euthyphro dilemma: morality is arbitrary. It seems that it is entirely up to God's free will how he combines facts in the relation of grounding. Thus, although this account is coherent, it suffers from implausibility.

But the phrase 'being established by God' admits of another interpretation, and this interpretation evades the objection that morality becomes arbitrary. The idea is the following: God's wanting or God's commanding is an expression of God's essence. And God is essentially good. Therefore, God's wanting or commanding is not arbitrary. And so we can say: The fact that an action a is obligatory is grounded in the fact that a has a certain natural property, for example the property of contributing to human flourishing. The fact that a has the property of contributing to human flourishing is a's proximate goodmaker.

For a person to perform morally worthy actions, it suffices that she believe, know, or understand this. She need not also have the additional theistic belief that the grounding relation between the two facts is established by an act of the essentially good God.

IV. CONCLUSION

In this article I examined whether there is a version of Metaphysical Supernaturalism which grants both (i) that moral values depend in a metaphysically strong way on God, and (ii) that agnostics and atheists are nevertheless able to perform morally worthy actions in the sense of the Right Reasons View. My result is that it is hard to find a version of Metaphysical Supernaturalism that grants both. The most promising candidate was the last version of the Grounding Claim. But even on this claim we must distinguish between a more or less detailed belief about, or a more or less complete understanding of, the goodmakers of the action. The Metaphysical Supernaturalist would have to say that theists' beliefs about, or understanding of the goodmakers of, actions would be more detailed or complete than atheists' or agnostics'. But this need not imply that atheists or agnostics could not perform morally worthy actions. For moral worth might require no more than knowledge or understanding of the proximate goodmakers of actions.[9]

BIBLIOGRAPHY

Adams, Robert. 1979/1999. 'Divine Command Metaethics Modified Again', *Philosophy of Religion: The Big Questions*, Eleonore Stump & Michael Murray (eds) (Oxford: Blackwell Publishers), pp. 412-416

Adams, Robert M. 1999. *Finite and Infinite Goods: A Framework for Ethics* (New York, Oxford: Oxford University Press)

Alston, William. 1990. 'Some Suggestions for Divine Command Theorists', *Christian Theism and the Problems of Philosophy*, ed. by Michael Beaty (Notre Dame: University of Notre Dame Press), pp. 303-326

Audi, Robert. 2007. 'Divine Command Morality and the Autonomy of Ethics', *Faith and Philosophy*, 24.2: 121-143

Correia, Fabrice & Benjamin Schnieder (eds). 2012. *Metaphysical Grounding: Understanding the Structure of Reality* (Cambridge: Cambridge University Press)

Hills, Alison. 2009. 'Moral Testimony and Moral Epistemology', *Ethics*, 120.1: 94-127

Jackson, Frank. Forthcoming. 'In Defence of Reductionism in Ethics', *Does Anything Really Matter? Responses to Parfit*, ed. by Peter Singer (Oxford: Oxford University Press)

Parfit, Derek. 2011. *On What Matters* (Oxford: Oxford University Press)

Quinn, Philip. 2006. 'Theological Voluntarism', *The Oxford Handbook of Ethical Theory*, ed. by David Copp (Oxford: Oxford University Press), pp. 63-90

Wierenga, Edward. 1983. 'A Defensible Divine Command Theory', *Nous*, 17.3: 387-407

SECONDARY-CAUSE AUTONOMY
AND DIVINE MOTIVATION

CHRISTIAN SCHRÖER

University of Augsburg

Abstract. An act-theoretical view on the profile of responsibility discourse shows in what sense not only all kinds of technical, pragmatic and moral reason, but also all kinds of religious motivation cannot justify a human action sufficiently without acknowledgment to three basic principles of human autonomy as supreme limiting conditions that are human dignity, sense, and justifiability. According to Thomas Aquinas human beings ultimately owe their moral autonomy to a divine creator. So this autonomy can be considered as an expression of secondary-cause autonomy and as the voice of God in the enlightened conscience.

I. TOPIC AND THESIS

What roles does the divine reason play in the motivation and commitment of human action? As noted Thomas Aquinas, one can ask this question in two senses, as he explains in the sum of theology:

> Even if the eternal law may be unknown to us, insofar as it is in the divine mind, as it is known to us still in some way, either by natural reason, which as its own image derived from him or by an additional revelation. (Summa Theologiae I-II q.19, a.4 ad 3)

In this paper I confine myself entirely to a discussion of natural reason, which, according to Thomas, corresponds to 'what human reason is able to understand about God'. According to Thomas, this concerns both theoretical and practical reason.

My thesis is that every moral motivation is essentially based on the use of practical reason. But this use is not possible without a basic acknowledgement of human autonomy. Human beings ultimately owe this moral autonomy to a divine creator. So the moral autonomy of human beings is an expression of secondary-cause autonomy. This then implies that the natural practical reason of the creature *is* the voice of God having an effect in the enlightened conscience. All other forms of religious motivation, just as a specific human motivation, can only be justified if they are compatible with the basic principles of this secondary-cause autonomy. Otherwise the instructions, which God gives by natural reason and revelation would not agree. (Cf. Aquinas, Expositio super librum Boethii De trinitate q.2, a.3c)

In the first step, I will introduce the systematic profile of a morally motivated action on the basis of responsibility discourse and deepen the topic by means of classical doctrines. I will then call on the basic forms of technical, pragmatic and moral motivation where they occur as part of this profile. In the last step I will show in what sense *all* of the seriously acceptable forms of moral motivation find their supreme normative condition in the three basic principles of secondary-cause autonomy.

II. THE BASIC PROFILE OF HUMAN ACTION

2.1. The Responsibility Discourse

The common theme of Socratic ethics is the responsibility discourse. Socrates calls on his fellow citizens to give answer for what they do and don´t do: *logon didonai*, which Cicero and the rest of Latin reception recite as *rationem reddere*. Plato emphasizes the universal character of this discourse several times through the picture of the judgment of the dead, which corresponds to the image of the Last Judgment in the Jewish, Christian, and Muslim tradition.

In order to hold a person responsible, six questions are necessary and only together sufficient. A responsibility discourse is usually triggered by an event seen as unpleasant or harmful and furthermore suspected to be attributable to human action or inaction. For instance, suppose that a train has derailed. In a first attempt to understand the situation, we will ask three questions, which relate to the matter of facts:

(1) We ask about the event: what exactly happened?

(2) We rate the event using two kinds of valuation, which usually arouse our interest in events: on the one hand judging it with regard to the difference of pleasant and unpleasant, and on the other hand with regard to the difference of useful and detrimental.

(3) We wonder about the cause: To what is it due that this event occurred? Was it a physical event only, just as other natural events? Or was it an action (such as sabotage / terrorism) or inaction (such as lack of maintenance)? If it was due to an action or inaction, then the question of what lead to the event is directed at a person. Only then it is expected that this person will be held accountable. The first three questions therefore concern three moments that are empirically recordable: event, rating, person.

Now if we actually can track down a certain person, we have to ask three more questions. For not everything people do can be attributed as an imputable act. John Austin describes the obstructive beginning of a day: I wake up, stumble over the carpet and spill my coffee. Here, all three verbs express an activity. Nevertheless, one would not readily speak of an attributable *doing*. You do not wake up, stumble and spill coffee *voluntary*. So we must ask:

(4) Did the actor really want to cause his action or inaction?

The fourth question thus concerns the voluntary mode of the doing (*hêkoûsion, voluntarium*).

Moreover, usually not every person is held responsible for every event they caused through their action or inaction. We do not blame the inventor of the railway for every train that derails. Rather, we focus on particular decisions of the actor, because with the decision to either take action or not, the actor usually intends to induce a very specific event. And this point also defines whether the actor will judge his action to have failed or been successful afterwards. Thus we can only speak of a morally attributable act if the actor knowingly induced or accepted an event as an inevitable event. The fifth question must therefore be:

(5) Did the actor knowingly cause exactly this event?

Only the sixth and last question leads to the actual motivation and reads simply:

(6) Why did the person do that? What induced the actor to do that?

Plato and Aristotle already called attention to the central point: Nobody does anything voluntary and knowingly without rating it *good* in at least one respect, 'because nobody wants something except he thinks, that it

is good' (Aristotle, Rhetoric 1369a3-5; cf. Plato, Gorgias, 467e1- 468b8). So we have to ask: Why did the actor think it was good to knowingly and wilfully bring about an event that is either enjoyable and useful or unpleasant and detrimental for the people affected? Only then we have posed the actual question of motivation.

The responsibility discourse already reveals three main classes of motifs: I do or don´t do something

   (a) because I expect the induced result to be useful for myself or others,
   (b) because I expect the induced result to be pleasant for me or others,
   (c) because I have the opinion that I, by doing or not doing this, do something good.

Aristotle shows that, with regard to friendship, the moral motivation includes the two other forms of motivation, but not vice versa: A morally motivated friendship is also useful and pleasant for the friends, but a friendship simply based on pleasure or utility is not simply a completely good friendship. Those two forms of friendship are called friendship only because of the partial resemblance to a friendship based on moral motives. For the two other motivations can occur without dependency on a moral motivation. (Cf. Aristotle, EN VIII, 3-5)

2.2. Willing and Doing

In his dialogue *Gorgias*, Plato draws a basic profile of human action that completely reflects the just-mentioned structure of the responsibility dialogue. (Cf. Plato, Gorgias, 466a5- 467d10) The fundamental aspect is the distinction between what people want and what they do: No one wants to swallow bitter medicine, and yet people do. Merchants who go to sea might not want to go to sea, and yet they do go out to sea. So there is a difference between what one does and what one wants: what you do is swallowing the medicine, but what you want is to live healthy. What merchants do is going to sea, but what they want is to trade and thereby to be rich. As Aristotle adds in the Nicomachean Ethics, what we want is always something general (to be healthy, to be rich). In contrast, what we do is always something particular (to swallow medicine, to go to sea).

Thereupon, Plato distinguishes between the *actor*, the *object* and the *rating* of the action, and he does that for both the side of willing and the side of doing: The willing is directed to a general aim, which is considered to be *good*. The doing however is directed to a single end

as a specific result, which one intends to bring about, and with the successful achievement of which one is satisfied. This also explains the tension between doing and wanting: No one wants to swallow bitter medicine. This is because you want it only participating. You only do it because factually it is conducive (*ôphelimon*) to the aim you pursue, and thus its being voluntary is only conditioned, because what you do appears to be required (*prosêkon*) for the good will you pursue. So here again we can identify a six-membered basic profile of human action:

The human act: willing and doing			
	practical attitude:	practical object	< practical valuation>
Willing:	I want to:	live healthy	\<because it seems to me to be good\>
Doing:	I want to: - participating -	swallow bitter medicine - in order to serve -	\<because it seems to me to be good\> - required -

2.3. Intention and Execution

As Thomas Aquinas emphasizes, both of these triads complement each other in the same way as *intention* (*ordo intentionis*) and *execution* (*ordo executionis*) (cf. Summa Theologiae I-II q.1, a.4c): The *intention* of action is the willing to pursue a general aim, which is assessed to be good. The execution of action, either acting or deliberately not acting, is done to cause a result which, in case the desired end is achieved, satisfies the actor.

The determining sense of the relation of intention and execution is threefold: The will *subjectively* determines actions or inaction, the general aim *materially* determines the intended specific results, and the valuation of the general aim determines the *moral* value of acting or not acting beyond the positive outcome. Execution is accordingly determined threefold: The act itself is executed *voluntarily*. The results achieved *serve* the general aim. The delight at the success of the action or inaction is only then *morally* delightful, if the success seems to be a required for the goodness of the general aim. (Cf. Summa Theologiae I-II q.5, a.1) Corresponding to the three determining relations of intention, we can identify the three moments of *subjective willingness*, *materially serving* and *normative requirement* on the side of execution.

The human act: intention and execution			
	Actor	What is done	< Value >
Intention: - determinates -	willing (voluntas):	aim (finis)	< good (bonum) >
Execution: - determined -	action/inaction (actio)	result (terminus)	< satisfied (delectatio) >
	'I'm acting voluntarily'	'It serves my purpose'	'Good insofar required'

The intention therefore motivates the execution is a triple sense: Willing subjectively motivates whether you act or do not act. The object of action or inaction is materially motivated through the general aim. And the value of the achieved result is morally motivated through assessing the general aim as good.

With this, the basic pattern of every morally attributable motivation is defined: Someone does or does not do something as a genuine human action exactly then, when it is intended in the named tripartite sense, namely that it appears to be fulfilling, meaningful and good. The same applies to every single action, but also to all sectors of human praxis and for the human life entirely. So the highest possible happiness of human beings would thus be achieved through an eventually fulfilled, meaningful and good life.

To align one's actions in respect to one's intentions, usually three intermediary personal capabilities are necessary: Whoever wants to do something needs the necessary courage (*fortis*) to do it. To adequately define one's general aims in each case, one needs the necessary orientating knowledge (*wisdom*). To not be misled by superficial aspects of pleasure and utility in the rating of one´s action, one needs to be sober-minded (*temperans*). Each of these three attitudes deals justly when it does its job in order to the common good. The highest possible misery would hence be a life in paralyzing fear, hopeless distress and unforgivable guilt. The best possible good news (tidings of joy) would thus be the one that overcomes all fear, distress and guilt and through this paves the way for a fulfilled, meaningful and good life.

III. THREE SOURCES OF MOTIVATIONS

3.1. *The Basic Form of Moral Judgment*

Both intention and execution exhibit the structure of an intentional capability: The intention assumes shape of a judgment (*logos, iudicium*). But also the execution of an action has the structure of an intentional performance, insofar as the execution always grounds on sensuous and motoric capabilities, which on their part have an intentional structure. Only together they constitute the *single* human action in the distinctness of intention and execution, whereat the intention gives the form (*what* is done) and the execution the matter (in what the action consists). (cf. Summa Theologiae I-II q.1, a.2; q.17, a.4c)

The basic form of any intentional capability encompasses an act and a specific object appertaining to the household of this act. The seen in the act of seeing are always colours, the heard in hearing are always sounds, the smelled in smelling are always fragrances, etc. With dianoëtical capabilities, a twofold object appears. Thomas remarks frequently: The specific object of the intellect is the being and the truth. The specific object of the will, however, is the aim and the good. (Summa Theologiae I-II q.1, a.1c; q.1, a.2 ad 3)

According to Thomas, this very point characterizes rational capabilities, because this twofold object the structure of judgment is substantiated. This leads to the elementary parallelism of theoretical and practical judgments: the realized in realizing are general states of affairs, valued in distinguishing between true or false. And the wanted in willing are general aims, valued in distinguishing between good and bad.[1]

---

[1] The parallelism of theoretical and practical judgments is laid out in the works of Plato and Aristotle, but systematically carried out by Aquinas and includes three main aspects: 1) In a theoretical judgment we judge a thought state of affairs with regard to the universal difference of true and false. In a practical judgment however we judge a general aim or a general option of action or life with regard to the universal difference of good and bad. 2) How a (not just hypothetical) theoretical judgment raises a conceptual act of conviction, which necessitates the one who judges in that way to a corresponding speech, thus a (not just hypothetical) practical judgment raises a deliberate conviction that necessitates the one who judges in that way to a corresponding action or inaction. 3) The respects and accounts, under which we assess in theoretical considerations the truth of the thought state of affairs, correspond in practical considerations the respects and accounts, under which the goodness of an action option is being assessed.

In his late article 'The Thought', also Gottlob Frege insinuates the parallelism of theoretical and practical judgments, when he writes as follows: 'Just as the word beautiful leads the way for aesthetics, and the word good does so for ethics, so leads the word true the way for logic.' (Frege 1918-1919/2010: 87) Hence both the word *good* in ethical judgments and the word *true* in theoretical judgments are classified within the same semantic category. Consequently, both forms of judgments can be pictured as follows:

Theoretical judgment:	I think:	state of affairs	< true / false >
Practical judgment:	I want:	a general aim	< good / bad >

3.2. Emotions as Subjective Motivation of Action

With this Frege ties on a classical understanding of judgment, which is not any more readily available in modern philosophy. On the one hand a theory of redundancy of truth is dominant, where the twofold object of theoretical acts of judgments is reduced again to one single proper object. On the other hand, generally only one capability of judgment is assumed. The only judging capability of human reason is the intellect; the intellect judges and in judging it merely distinguishes between true and false. David Hume concludes correctly: From states of affairs which are true or false an act cannot be reasoned to be good or bad. Consequently, the source of moral motivation has to be found elsewhere. If though the intellect is only capable of theoretical judgments, the only motivator left is emotion.

The motivation by emotions however is perceived fundamentally different to this in the classic Socratic ethics. The most consequential analysis can be found with Aristotle. Again, the three-membered structure of actor, object and valuation is fundamental: Each intentional act has a self-reference sensible to emotions. An unimpeded self-realization is received as pleasant and gratifying; a disturbed self-realization conversely is felt as unpleasant and annoying. Disturbances can occur both on the side of the capability and on the side of the object. Hence pleasure develops if a capability is undisturbed and successfully directed towards an object belonging to its specific household. So it seems to be an 'unimpeded activeness of a nature-conforming habit' (EN VII, 13, 1153a12-15, cf. VII, 14, 1153b9-12) Aristotle defines his findings more precisely in his second discourse of pleasure: Pleasure *is*

not an undisturbed act of doing, but *adds* to a successful act of doing and *completes* the act in its subjective affectivity. (EN X, 4, 1174b31-33)

In every intentional act, pleasure thus comes up as subjective motivation, treating every successful activity with a pleasant self-referential sensation. Precisely because of this circumstance pleasure is no reliable motivator. This is because human action usually is comprised of the activation of several intentional capabilities. The acts of different capabilities however can come to compete with each other: it is pleasant for the hearing to listen to enjoyable music, but enjoyable music might occasionally disturb successful thinking. (Cf. Summa Theologiae I-II q.5 a.1 ad 3) For the will it would be pleasing to reach the goal of being healthy again, but for the taste it is not delightful to take bitter medicine. However, in this sense the immanent delighting acts of single capabilities might come to compete with a holistic human joy of living a life perceived to be an overall fulfilling, meaningful and good life.[2]

3.3. Utility as a Secondary Motivation of Action

Utility as well is not a primary principle of moral motivation. It is important for the Socratic concept of action, that we consider the relation of wanting and doing not as a connection of cause and effect. If it were simply meant that the actor swallows the physical medicine to produce a special organic effect, Plato's text would not describe what is done as *to ophelimon* (what serves), but as *to chrêsimon* (what is useful). For the proposition rather says, that to swallow the medicine (what I particularly do) *serves* to be healthy (what I generally want).[3]

The concept of serving (*ophelimon*) signifies an *in-order-of*-relation (i.e. a *for-the-purpose-of*-relation), so correlates reciprocally to the *by*-relation between several terms which describe an action: I get wealthy *by* swallowing the bitter medicine, so taking the bitter medicine serves in that case to live healthy again. I become rich *by* going to sea, so going to sea serves to become rich. (Cf. Goldman 1971) Doing this by doing that

---

[2] So Jeremy Bentham is right when he says, that 'nature has placed mankind under the governance of two sovereign masters, pain and pleasure', but this fact doesn't establish a moral principle (Bentham 2003: c. I, n.1).

[3] In his dialog Gorgias Plato introduces his action-profile in three steps: In the first step he exposed the difference between wanting and doing (466a5-467d10), in the second step the universal state of rating predicate good (467e1-468b8) and in the last step the concept of what serves (to ophelimon) (468b9-e6). There is no talk about what is useful (to chrêsimon) in this text passage.

refers to the order of intention and here to the inner constitution of one action in the unity of the two aspects what is done (general form) and in what does the action consist (single matter). However to use single means to cause another single event or effect refers to the causal relation between two single and different circumstances. Using something describes only a technical necessity, and in that view we consider our action somewhat from outside as a chain of several events. But to say *this serves to do that* in the sense that we do that by doing this signifies only one single event, and in that case we consider our action from the actor's intentional point of view.[4]

As has been shown already, the concept of moral motivation bonds with the normative concept of what seems required: Whoever wants a general aim the actor has considered to be good also believes the respective action to be required. From the profile of action outlined so far, now three basic forms of normativity can be defined more precisely:

(1) A certain action or inaction is *technically required,* when within an act of execution single measures need to be taken to accomplish single results: If I want to be on time for my appointment at the doctor´s, *I have to* leave now. In the profile of action, technical requirements emerge as a horizontal compulsion within the level of execution.

(2) Action or inaction is *pragmatically required,* if for the reason of an actual intention one feels compelled to act or not act: If I want to live healthy again, *I have to* swallow the medicine. In the profile of action, pragmatic requirements emerge as a vertical compulsion between the level of intention and the level of execution.

(3) Action or inaction is *morally required,* if one wants to act because the general aim has been assessed to be fulfilling, meaningful and good in regard to the human life at all: Insofar it is fulfilling, meaningful and good to care about human wellbeing, *I should* think about what one can do to maintain health and which measures I can set to realize these actions or inactions.

---

[4] The difference between to ophelimon (what serves, cf. opheilô: I should, owe, must) and to chrêsimon (what is useful) is lost in the course of the transition into the Latin terminology, because at least since Cicero both words are translated with utilis. Also the modern lexica give useful for both Greek words. Plato itself heeds the difference strictly in all his dialogs.

> Moral requirements therefore originate in a compulsion, which emerges on a horizontal level of intention, thus in the practical judgment itself.

So if someone wants to act in a way that his action is morally good (*eu prattein*), he will start off from the practical judgment considering three forms of thought: A first consideration is concerned with the *forming of will*: Are the general aims along which I orient my living and doing, fulfilling, meaningful and good? The second consideration is concerned with *consulting*: How I can effectuate my general aims by means of achievable single results? The third consideration concerns the *execution*: Which measures, competences and resources can I draw on to achieve those results? Hence, a morally good acting or not acting reveals, as was said as early as by Demokrit, the form of a threefold-birth (*tritogeneia*). So the earliest definition of prudence (*phronêsis*) as moral reason given by Demokrit is: The one is prudent who has good speech (*eu legein*), deliberates correctly (*eu logizesthai* resp. *eu bouleuein*) and then does what is to do (*práttein hê dei*). (FVS 68, B2) Thus the moral quality of a human action is decided by the moral quality of the leading intention.[5]

The objects of our needs, desires and interests as well as our self-concept or corporate identities may thus be motives which inspire us pragmatically to our actions. However, it is up to us whether we let us move from these motives or not move, and we are always able to subject these motives to the critical question of whether it is good to me to meet these needs, desires and interests or to follow those personal or social self-concepts.

---

[5] Two persons give alms, but in a different purpose: one wants practice charity, because he considers it good. The other wants to be admired by the people because this seems to him to be good. The first person has reached his goal of action when the need is help, the second actor, however, only when he is also admired for his act. Usually we make a difference between the moral qualities of these two cases. But both actions are the same in execution and differed only on the intention. So, the reason for the difference in moral evaluation is only found in the different intention. Cf. Summa Theologiae I-II q.1, a.3 ad 3.

IV. AUTONOMY AND HETERONOMY

4.1. Heteronomy and Autonomy of Moral Motivation

The classical difference between autonomous and heteronomous morality relates to nothing else but to the question of the general aims or intentions one has to follow. Indeed, one can ask: Do I follow maxims set by someone else, or should I only follow maxims I set myself? This question is as old as the Socratic ethics itself. Already in the Apology of Socrates, Plato distinguishes two forms of speech, i.e. the rhetorical speech and the Socratic speech: Rhetoric speech is violent, because it is the goal of rhetoric to actuate the audience to act in accordance with the speech of the rhetor. Socrates begins his speech for the defence with the words:

> What my prosecutors have done to you, men of Athens, I don´t comprehend (*ouk oîda*); I for myself almost forgot myself [!] over them; as they talked so suavely (*pithanôs elegon*). (Plato, Apology, 17a1-3)

Furthermore, in the Dialogue *Gorgias* the title character extols precisely this power of rhetoric 'that one is able to persuade by words', specifically as its greatest good (*megiston agathon*):

> For you have this in thy power, so the doctor will be your slave, the gym teacher will be your slave, and from that acquisition man (*chrêmatistês*) it will be shown that it acquires another (*allô*) and not himself (*ouch hautô*) but thee, who canst thou talk (*legein*) and to persuade the amount (*peithein*). (Plato, Gorgias, 452d3; d9-e7)

The power of the Socratic speech, however, lies solely in the insight whether the *logos* that guides the action is good and just. In this sense Socrates checks along with Crito, whether he should follow the speech of the people or the speech of Crito or the speech of the laws. Hence he continues:

> Jointly (*koinêi*), you Good, let's think about that, and if you oppose something (*antilegein*) to what I speak, oppose, and I will obey you (*peisomai*). If not, so stop it to tell me the same speech again and again therefore that I should against the Athenians will (*akontôn Athenaiôn*) go away from here. For it is surely of worth to me if you persuade me to do this, but not involuntary (*mê akontos*). (Plato, Kriton, 48d8-e5)

The Socratic speech stresses the utmost importance for each interlocutor to not consent to a demand of action before he himself has carefully examined it. So the Socratic speech treats the audience as a human being, who does not follow any demand but one he has appropriate to himself and has made it his own leading intention after an authentic process of thorough review. Precisely this motif is carried via Aristotle and Cicero to Thomas Aquinas and via Rousseau finally to Kant.

4.2. Non-cognitive Objections

As far as we assume that general aims or maxims mainly arise from accustoming to the moral standards of our educators and our social environments, moral convictions coincide with what in psychology is called *superego* (Freud) or *parent I* (Harris). But insofar as we are not masters in our own house, as Freud says, this entails just a moral heteronomy. In contrast, in order to preserve moral autonomy, other authors emphasize the freedom of human beings and defend pointedly the free choice of maxims, for which we can give only more or less plausible reasons (Hume, Tugendhat). John Stuart Mill combines both views in the conviction that what we refer to as conscience is nothing else but an emotional barrier fundamentally determined by education, and he sees no reason why those feelings 'may not be cultivated to as great intensity in connection with the utilitarian, as with any other rule of morals'. (Mill 1861: III 5)

The demand to explain not only the technical needs and pragmatic requirements, but also to justify the choice of his maxim seems to be avertable with the argument that this would only shift the problem, because such a justification would in turn be justified such that the reasoning gets only into the aporias of each final justification (Popper, Albert, Habermas). But if there are no objective criteria for justifying general moral standards, then internalized or freely chosen maxims are the last instances of human action justifications. In this case moral autonomy seems preserved, but at the cost of a lost generally binding morality. On the other hand, whoever argues that a justification of objective moral criteria, which binds the human will, is quite possible exposes himself to the accusation of moral heteronomy.

The concept of moral judgment in the tradition of the Socratic ethics agrees in a certain sense with the non-cognitive doctrine, that moral statements are neither true nor false and insofar they are not truth-apt;

a practical judgment does not say that a state of affairs is *true or false*, because it says that to act in a certain way is *good or bad*. Strictly speaking it also denies the claim that 'moral judgments are capable of being objectively true [!], because they describe some feature of the world' (Garner & Rosen 1967: 219-220), but this fact does not imply that there can't be any generally binding criteria of morality.

Given moral decisionism it is seemingly implied that moral maxims are based on decisions. But from the fact that maxims are set by the actor, it does not follow that there can't be any objective criteria for such settings. For until proven to the contrary the possibility remains open, that the choice of maxims does not appear randomly (conventional) or merely *de facto* (pragmatic), but in a certain sense inevitable (categorically). Maxims quite get their subjective validity for the actor by an act of consent. But this does not exclude the possibility that a maxim also has an objective validity, in the sense that there is no chance to deny it in a serious way, for it can be represented in an inescapable sense as reasonable and good towards any human being. If, however, the moral quality of an individual action or inaction were finally measured on the moral quality of the leading intention, it would be able to show that in certain cases there is no maintainable maxim that allowed a serious justification of the intended action or inaction.

4.3. Three Absolute Principles of Autonomy

This leads to a last pointed emphasis of our question: Do criteria exist to determine which leading intentions one can, must, or must not accept? Also in this point Aquinas refers to the parallelism of theoretical and practical reason. Commonly accepted are the main conditions of theoretical judgment. So anyone who wants to testify anything is necessitated to take into account three basic conditions of the theoretical judgment in his statement:

(1) No suspension of propositional attitude: Thinking that I do not just think would suspend the thinking at all. (Cf. Summa Theologiae I q.76, a.1c) From this first principle derives the Cartesian cogito argument and all transcendental arguments for self-assurance.

(2) No cancellation of what is said: Statements such as *The circle is square* or *The part is greater than the whole* are not conceivable because in these links one term cancels the other. Conversely, in sentences such as *The circle is round* or *That part is not greater*

> *than the whole* the combination of the terms will appear necessary, since in this case it seems not possible to think one term without the other term. (Cf. Summa Theologiae I q.2, a.1 arg.2) On this second principle are based all analytically true and false sentences.

(3) No waiver of truth reference: Who claims that there is no truth, must refer at least for this claim to the difference of true and false. Otherwise he would not seriously claim that there is no truth. (Cf. Summa Theologiae I q.2, a.1 arg.3)

These hints make it clear that each of the three constitutive moments of theoretical judgment has its own necessary condition that must be met for a serious statement. A claim that explicitly or implicitly nullifies one of these three principles thus inevitably destroys the basic conditions of a theoretical judgment at all. Therefore such a statement can no longer be represented as a serious judgment.

So an answer looms as well to the question according to the modalities of human wanting. To this subject Thomas dedicates a whole *quaestio* (Summa Theologiae I-II q.10), and it also leads Kant to the basic formulas of the categorical imperative. But it is also already apparent in the earliest Socratic ethics, that there are exactly three absolute limits of a rational consent. The answer stems from a simple observation: Whoever imposed upon me a moral standard, whether I myself or my educators or my social environment or my divine creator, must consider three necessary and jointly sufficient conditions.

Dignity: The first principle relates to the subject of intent: No one can seriously accept a general standard which denies the actor any right to an independent, voluntary consent. Aristotle uses this motif in the friendship books when he compares the relation of a lord to his slave by the relation of a craftsman (*technítês*) and a tool (*organon*). If the lord looks at the movements of the slave only as the execution of his own act, the lord looks at the slave merely as a tool. But toward a mere tool there can be neither justice nor friendship. Now the slave is at the same time also a human being. But then toward him there is justice:

> For it seems that there is a certain justice (*ti dikaion*) in particular for all [!] who have the capacity to participate in legal and contractual communities (*ton dynamenon koinônêsai nomou kai synthêkês*). (Aristotle, EN VIII, 13, 1161b6-8)

But if justice, then also friendship. So there can be no justice for the slave, as long as the slave is used as a mere tool, but rather as being a human

being (*kath` hoson anthrôpos*). Cicero connects this argument with the notion of dignity that each person belongs by virtue of his humanity. Therefore this principle finds its expression in the concept of humanity (*humanitas*). Adam Smith finishes the main chapter of his *Theory of Moral Sentiments* with the words: 'No action can properly be called virtuous, which is not accompanied with the sentiment of self-approbation.' (Smith 1759/1985: III, c.6; last sentence). Finally Kant dresses this principle as is well known in the so-called self-purpose formula: Nobody can seriously accept a maxim that looks on him only as a means and does not respect him as an aim in itself.

Sense: The second principle concerns the object of intention. No one will seriously accept a principle by which his own actions would become meaningless, be it contradictory or utopian. Wolf Biermann, a political songwriter in the former East Germany, said in a quite critical remark: Who preaches people hope is a liar. But he added immediately: Who takes peoples' hope is a bastard. The classical ethics therefore is compelled to have a highest common goal that can be willed itself as a fulfilling, meaningful and good. Without such an aim held eternally for its own sake and therefore able to lend meaning and orientation to all other aims all human doings would become 'senseless and void'. (Aristotle, EN I 1, 1094a20-21; cf. Summa Theologiae I-II q.1, a.4c) Especially Camus clearly underlines the principal character of this point: Where people experience the lack of a deeper reason for life, all action and suffering appears as useless and even human life banal. So the question remains whether life is worth the effort to be lived 'the fundamental question of philosophy. [...] Everything else comes later'. (Camus 1950/1997: 10) Kant addresses this second condition in the natural law formula: Nobody can seriously accept a maxim that leads to an absurd practice.

Justifiability: The third principle relates to the valuation of the intended object: Nobody is going to accept a principle that can be represented in any way as well. Aristotle points out that even the hedonist must maintain necessarily that pleasure is a good. (Cf. Aristotle, EN X, 2, 1172b35-1173a4) For Thomas, every moral judgment based on the fact that a human being, in contrast to other animal, is able to refer to the *bonum universale*. No one will accept an ultimate goal if he could not value that goal as a good target. Kant dresses this third moral principle in the formula of the common dominion of aims[6]: In a common practice in which everyone is affected by the maxims of the other, only those maxims can be accepted that appear compatible with all other maxims,

whose goodness cannot be seriously denied. So no one can seriously accept a moral intention which cannot be justified in a universal space of good intentions over all other general intentions whose goodness cannot be seriously disputed.

The top three limiting conditions, under which a leading intention can be seriously accepted, are therefore: dignity, meaning and justifiability. The point that puts Kant in the centre is that these three conditions just apply to each maxim that you choose yourself:

> Our own will, so far he would act only on the condition of a possible universal legislation by its maxims, this potential will in us in the idea, is the real subject-matter of respect, and dignity of humanity consists just in this ability, to be general lawgiving, although with the conditional, to be subjected to this same legislation. (AA IV, 440. 5-10, transl. in Egger 2014: 273)

The moral autonomy, therefore, constitutes not only all absolute rights, but also all absolute obligations of human beings. For each practical judgment which satisfies these three criteria is morally possible in a universal sense; if it does not meet these criteria in any way, it is morally impossible; and any practical judgment, in regard to which the denial does not meet these criteria in no way, is morally necessary. So, human action and inaction can be morally rated as far as they are suitable to realize dignity, sense and justifiability within the bounds of human contingency. But above all, the principle of autonomy ensures an essential self-respect that is due to the human being as a morally acting being.

That leads us back to the thesis of this paper: Human beings ultimately owe their moral autonomy to a divine creator. So this autonomy is an expression of secondary-cause autonomy. All other forms of religious motivation as human motivation can only be justified if they are compatible with the basic principles of this secondary-cause autonomy that are at the same time the basic principles of morality at all.

---

6 The term used 'kingdom of ends' is considerably misleading, if end is considered as result (terminus) instead of aim (finis).

BIBLIOGRAPHY

Bentham, Jeremy. 2003. 'An Introduction to the Principles of Morals and Legislation', *Utilitarianism and On Liberty. Including Mill's 'Essay on Bentham' and selections from the writings of Jeremy Bentham and John Austin*, ed. by Mary Warnock (Oxford: Blackwell), pp. 17-51

Camus, Albert. 1950/1997. *Der Mythos von Sisyphos. Ein Versuch über das Absurde* (Reinbek bei Hamburg: Rowohlt)

Egger, Mario. 2014. *Philosophie nach Kant* (Berlin: De Gruyter)

Garner, Richard & Bernard Rosen. 1967. *Moral Philosophy. A Systematic Introduction to Normative Ethics and Meta-ethics* (New York: Macmillan), pp. 219–220

Frege, Gottlob. 1918-1919/2010. 'Der Gedanke. Eine logische Untersuchung', *Die Philosophische Logik Gottlob Freges. Ein Kommentar*, ed. by Wolfgang Künne (Frankfurt am Main: Klostermann), pp. 87-112

Goldman, Alvin. 1971. 'The Individuation of Events', *The Journal of Philosophy*, 68.1: 761-774

Mill, John Stuart. 1861/1998. *Utilitarianism* (Oxford: Oxford University Press)

Smith, Adam. 1759/1985. *Theory of Moral Sentiments* (Oxford: Oxford University Press)

'CAN A PHILOSOPHICAL JUSTIFICATION OF ETHICS BE AUTONOMOUS WHILE ACKNOWLEDGING THE ROLE OF GOD IN GROUNDING MORAL FACTS?'

CHRISTOPH HALBIG

University of Zurich

Abstract. Autonomy and ethics are related to each other in complex ways. The paper starts by distinguishing and characterizing three basic dimensions of this relation. It proceeds by arguing for the compatibility of moral realism with a due respect for human autonomy. Nevertheless, supernaturalist moral realism seems to pose a special challenge for the autonomy of ethics as a self-standing normative realm. The paper ends with some considerations on the role of divine authority both in metaethics and in the general theory of value.

'Can a philosophical justification of ethics be autonomous while acknowledging the role of God in grounding moral facts?' Before addressing this question, which was raised at a workshop with the title *Divine Motivation vs. Human Autonomy? Metaethics Between Autonomy and Heteronomy* (and for which this paper has been written), the multiple ways in which it can be understood must first be disentangled. This is due to the fact that autonomy plays at least three different roles in thinking about the problem of how God and ethics might be related to each other:

Firstly, human autonomy is undoubtedly a genuine good that ought to be respected by an adequate religious ethics. But how is human autonomy compatible with a conception of moral norms that grounds itself in the will of another, moreover, omniscient and omnipotent being, i.e. God?

Secondly, whereas in this first role it is *human* autonomy that is at stake, it might be the autonomy of *morals* itself that is cause for concern: Richard Hare for instance claims that the autonomy of morals has become part of the very meaning of the concept of morality: "Ever since Kant, it has been *possible* for people to insist on the autonomy of morals – its independence of human or divine authority. Indeed it has been *necessary*, if they were to think morally, in the sense in which that word is now generally understood." (R. Hare 1992, 30). If Hare's thesis holds true, a theonomic conception of morals must be misguided on conceptual grounds – it is incompatible with how we post-Kantians use the very term 'morals'.

A *third* way of understanding the problem does not ask – unlike the second formulation of the issue – how to defend the *autonomy of ethics* (independently of the concrete form that ethics may take), but rather how an *ethics of autonomy* might be compatible with attributing God a crucial role in the constitution of moral facts. An ethics of autonomy considers moral normativity as being constituted by the self-legislating activity of practical reason; it thus belongs to the constructivist paradigm in meta-ethics, according to which the task of practical reason is not to bring us into contact with an independently existing realm of moral values, but rather to constitute moral normativity by the very activity of determining oneself according to principles that people prescribe to themselves qua rational beings. A theonomic conception of moral normativity seems even more difficult to reconcile with the idea of such an ethics of autonomy than with the idea of the autonomy of ethics: If moral facts are construed by some process which owes its authority to practical reason, there seems no place left for God. After all, a divine will intervening in that process seems to constitute a paradigmatic example of heteronomous interference that threatens to distort the core of moral normativity with external rewards and punishments.

Although these three dimensions of the autonomy-problem should be addressed separately, they are highly intertwined: An ethics of autonomy for instance seems ideally suited to preserve individual autonomy and offers an attractive way of defending the independence of ethics against both human and divine authority. On the other hand, moral realists who reject the constructivist idea of an ethics of autonomy often struggle with the so-called argument from autonomy, i.e. the charge that the existence of an independent realm of moral values necessarily undermines our autonomy as agents. Therefore, the problem exists even before a divine

agent enters the scene, an agent that, according to some versions of supernaturalistic meta-ethics, plays a crucial role in constituting moral reality. The grounding of moral facts in a divine agent seems however to aggravate that problem. In addition, such a supernaturalistic position also stands in tension with the idea of the autonomy of morals itself – instead of being autonomous, moral normativity seems to be swallowed up by divine authority.

Of the three respects in which autonomy takes center stage in ethics mentioned above, it is certainly the first that is the hardest to give up. A theonomic conception of ethics might very well join non-naturalist moral realists in rejecting the constructivist approach of an ethics of autonomy. At the least, it might qualify the idea that ethics itself has to be an autonomous normative sphere. But it nonetheless needs to show why it does not require the sacrifice of human autonomy: Not just for substantive moral reasons – a benevolent God seems to owe our autonomy due respect – but also because the idea of addressing autonomous agents seems to be part and parcel of moral normativity itself. A theory that disregards this constraint would hardly count as a theory of moral normativity at all. It is no coincidence that philosophers like James Rachels have taken human autonomy as the starting-point for an argument that is meant to prove the nonexistence of God. The idea is that any being who fills the role of being God is necessarily worthy of worship. This in turn requires total subservience on the part of God's creatures. Being subservient in this way however, as Rachels argues, is incompatible with the concept of an autonomous agent (cf. Rachels 1971, 325-337).[1] It goes without saying that sacrificing the autonomy of moral agents is not a promising way to refute such an argument.

In what follows I will make some tentative suggestions as to how to sort out the complicated issues that make up the three-dimensional problem just outlined: In a first step, I am going to argue that a moral realism which considers moral values as entirely independent from human attitudes is compatible with human autonomy. There is thus no need to embrace an ethics of autonomy as the only way to accommodate human autonomy. Nonetheless, there might of course be other reasons for a constructivist ethics of autonomy that are beyond the scope of this paper. For instance, it might be argued that constructivism carries less metaphysical baggage than moral realism since it avoids commitments

---

[1] For a critical discussion of this argument see Quinn 1973.

to evaluative or deontic facts and is therefore not subject to Mackie-style arguments from queerness. But even if it can be established that moral realism in general respects human autonomy, the claim that supernaturalistic versions of moral realism are also exempt from the criticism needs a separate defense. In a second step, I will explore the stance a theist should adopt towards the autonomy of ethics-thesis. As will be shown, both adopting and rejecting this thesis comes at a considerable price. I will conclude with some suggestions regarding the role divine authority should play in the theory of value and in meta-ethics. Addressing that issue will require introducing some further distinctions concerning the kinds, sources and scope of divine authority.

I.

An influential argument against moral realism (especially popular among those who, like Kantians or Neo-Kantians, subscribe to an ethics of autonomy in the sense defined above) claims that acknowledging the existence of evaluative or deontic moral facts completely independent of human attitudes (be they rational or orectic ones) *ipso facto* undermines human autonomy. But why think that? One might argue that subjecting human will to any authority that is not the product of rational self-legislation is incompatible with human autonomy. But such an argument fails for various reasons:

Firstly, it just restates the constructivist position and therefore begs the question against the realist. In areas besides morals, we are ready to acknowledge the existence of truths not of our own making without feeling compromised in our autonomy. As Russ Shafer-Landau memorably puts it: "It is not a restriction on autonomy that one can't make two and two equal five." (Shafer-Landau 2003, 44). One might be tempted to protest that this analogy downplays a crucial difference between truths in, say, mathematics or geology on the one hand and moral truths on the other. Unlike moral truths, mathematical facts are not intrinsically normative, as they do not imply reasons for action and thus do not impinge on our will. We are not allowed to *believe* anything incompatible with those facts, but it is up to us which practical attitudes we adopt towards them. The case of moral truths is different: They do seem to make normative demands on us that constrain our autonomy

from the outside. In response the realist may admit that our autonomy is constrained by external normative truths, but deny that this undermines in any way human autonomy. Quite to the contrary: autonomy seems to be enabled by those constraints. Rather, what is paradoxical is the constructivist program according to which human beings generate normative truths that are binding on those very same agents whose rational activity is responsible for their existence.

Secondly, in response to such a rejoinder one might argue in the spirit of Richard Rorty that realism in all its forms is authoritarian. Rorty explicitly highlights the close link between "the idea that truth is a matter of correspondence to the intrinsic nature of reality" and "the idea that morality is a matter of correspondence to the will of a Divine Being." (Rorty 2006, 257). According to Rorty, realism in all its forms should be replaced by the guiding idea of solidarity among human beings. Yet this Rortyean thesis is much too strong. Some forms of realism might indeed be authoritarian, for instance if they imply the existence of an inscrutable realm that exceeds our epistemic reach and is nonetheless making demands on us which we will never even be able to know. Admitting the existence of an outside reality not of our own making on the other hand seems – far from being a figment of misguided philosophical imagination – part and parcel of our common sense approach to reality. As John McDowell puts it: "Acknowledging a non-human external authority over our thinking, so far from being a betrayal of our humanity, is merely a condition of growing up." (McDowell 2000, 120).[2] Why should it be different in the case of moral truths? In addition, even a Kantian proponent of an ethics of autonomy will find Rorty an uncongenial ally: Insofar as it relies on the authority of practical reason, a Kantian constructivism strikes Rorty as being just as authoritarian as a moral realism relying on the existence of independent moral facts.

Thirdly, even if one eschews such a sweeping, Rortyean critique of realism in *all* of its forms, one might still insist on the special threat to autonomy that is posed by independent *moral* truths. But here again the realist might turn the tables on his opponent. On the one hand, autonomous decisions seem to presuppose some constraints that

---

[2] For a critical discussion of the "Rortyean argument from autonomy" against realism see Stern 2012, 126-129.

guarantee their own rational intelligibility as opposed to mere acts of whim. On the other hand, moral realism rightly understood leaves ample breathing space for autonomy: It is by no means relegated to a purely epistemic role in finding out about moral truths;[3] a moral realist is not *ipso facto* committed to the idea that, for instance, the truth of each and every all-things-considered moral judgment is already fixed by an independent moral reality. There might be moral ties, incommensurability between different values etc. which provide considerable challenges for the exercise of human autonomy. Claiming that our own autonomously selected moral commitments exhaust the realm of ethical truths would of course once again simply beg the question at stake.

In conclusion, the argument from autonomy against moral realism ultimately fails. There is no need to embrace an ethics of autonomy as the only meta-ethical position able to accommodate proper respect for autonomy. As mentioned above however, there might of course be other independent reasons to opt in favor of Kantian constructivism instead of an ontologically more demanding moral realism.[4]

<div align="center">II.</div>

How should a theonomic ethics respond to the autonomy of ethics-thesis? First, I will consider how moral realism deals with this worry and then I consider the special case of supernaturalist moral realism.
Non-naturalistic moral realism seems ideally suited to account for the autonomy of ethics. After all, it claims that moral facts do simply exist, even though they might supervene upon facts of some other kind (either natural or supernatural). Ontologically speaking, those facts are *sui generis*. The wrongness of lying might prove just as metaphysically primitive as the basic truths of logic. Those brute moral facts would then neither be amenable nor in need of a reduction to something non-moral.

---

[3] Shafer-Landau rather misleadingly suggests this when he writes: "In such areas [i.e. where truths are not of our own making], our autonomy, well utilized, consists in discerning the paths to gaining such truths, rather than in creating it." (Shafer-Landau 2003, 44)

[4] According to Korsgaard moral realists miss the point of moral problems in the first place by turning ethics into "a theoretical or epistemological subject" (Korsgaard 1996, 44). As a consequence, they have to shoulder an ontological burden that inevitably invites Mackie-style arguments of queerness.

Within such a framework, the autonomy of ethics comes as little surprise.

What about a supernaturalist version of moral realism? At first glance, such a position seems hard to reconcile with the autonomy of ethics. If evaluative properties like goodness and badness or deontic properties like rightness and wrongness are given a supernaturalist interpretation – the property of goodness might be considered identical with the property of resembling God (cf. Adams 1999, ch. 1), the property of wrongness identical with being contrary to the commands of a (loving) God (cf. Adams 1987a, 133-142) – then ethics seems to have lost its autonomy. But is the autonomy of ethics something one should try to preserve in the first place? That seems to me rather hard to say since there are good arguments on both sides:

On the one hand, the phenomenology of moral experience seems to support the autonomy of ethics-thesis: Some things like pain strike us as intrinsically good or bad and others like lying or torturing as intrinsically wrong – "intrinsically" meaning here *in virtue of their intrinsic, non-relational properties*. But "resembling God" or "being commanded by God" are paradigmatic examples of relational properties. Of course, there might be cases of normative overdetermination: lying might be wrong both in itself and because of its being forbidden by God's commands, where either of the two is sufficient to make it wrong. Common sense, however, is likely to insist that there is simply no need to add such a supernaturalistic story to the picture; it seems to be a superfluous add-on to moral experience.[5] Even worse: Providing such a supernaturalistic account of the ontology of normative truths seems not only dispensable, but might even distort their normativity. That lying is wrong might provide not only sufficient reason for refraining from doing so, but might also provide the *only* legitimate moral reason against lying. That lying violates God's commands and is likely to be punished by him might be a perfectly respectable prudential reason, but rather doubtful as a moral one.

On the other hand, claiming ethics to be autonomous seems not just a key *strength* of non-naturalistic moral realism (as it fits so well with

---

[5] For such an account of goodness see Adams 1999, 28-38. Erik Wielenberg in his review of Scott A. Davison's book *On the Intrinsic Value of Everything* correctly points out that a theory like Adams's is hard to reconcile with the independently plausible idea that some things beside God are simply intrinsically good or bad, quite independently of how they are related to other things including God. (Wielenberg 2012,145)

moral experience), but also its crucial *weakness*. The idea here is that moral values or norms are (unlike, for instance, logical truths) unlikely candidates for the status of brute, ontologically primitive facts. Moral truth seems to be in desperate need of *external* grounding (if moral truth is not to be abandoned in favor of non-cognitivist or error-theoretical positions in meta-ethics), and theism might be able to provide such grounding. If theism were the *only* candidate able to do the trick, the objectivity of moral claims could even serve as the starting point of an argument for the existence of God. Some philosophers, for instance William Lane Craig, consider the moral argument for the existence of God as "the most effective" argument for theism (quoted in Wielenberg 2009, 23). It is of course a hotly debated issue as to whether such an argument – even if one accepts its premise, i.e. the objectivity of moral truths, and in addition accepts that moral objectivity is in need of an external grounding – is likely to succeed. In any case, as a minimal requirement, God's existence needs to be a more plausible candidate for the status of a metaphysically primitive, brute fact than the existence of moral values; one reason to think so might be that the existence of God, unlike that of moral values, is metaphysically necessary.[6]

Besides the alleged need for an external grounding of ethical normativity, there might be a further reason for rejecting the thesis propounding an autonomy of ethics that derives from theism itself: Although the autonomy of ethics is, as has been shown above, compatible with the autonomy of finite human beings, it seems harder to reconcile with a divine person who is not just autonomous but also omnipotent. As Robert Johnson puts it: "The value that provides a reason for God to love it would be a constraint on God's love in the sense that God must respond to reasons provided by the value of things or else fail to have the requisite response" (Johnson 2007, 140). For many theists, even such a normative 'must' undermines God's omnipotence. Divine Command-theories that reject the autonomy of ethics-thesis avoid the problem of limiting God's omnipotence by making moral values and norms constitutively dependent on God's will. Although even according to

---

[6] As against such an asymmetry-thesis, Erik Wielenberg for instance argues that there is perfect parity between theists and moral realists: "To ask of such facts [i. e. basic ethical facts], 'were do they come from?' or 'on what foundation do they rest?' is misguided in much the same way that, according to many theists, it is misguided to ask of God, 'where does He come from?' or 'on what foundations does He rest?' The answer is the same in both cases: They come from nowhere [...]." (Wielenberg 2009, 26)

the divine command theorist, God's will might be constrained from the inside (i.e. by God's nature, in particular by his attribute of goodness), there is no external constraint on God's will exercised by an independent normative authority.

The theist cannot have it both ways: If he accepts the thesis of an autonomy of ethics and thus admits the existence of moral truths independent not just of human but even of divine attitudes, then he has to deny himself moral arguments for God's existence: Moral reality cares for itself and is not in need of any constitutive activity on God's part (God might of course lend a helping hand as a moral teacher, for instance, thus facilitating epistemic access to moral reality by revealing it to less than omniscient human beings). Furthermore, the theist has to show that not only finite human autonomy but even divine omnipotence is compatible with the existence of such an independent moral realm. If, however, the theist denies the autonomy of ethics-thesis, he escapes this twofold burden of proof, but at the price of facing another challenge. The dependence of moral truths on God (either on His essence or on His will) seems (i) to fit ill with the phenomenology of moral experience and (ii) poses serious problems for theism itself:

Ad (i): As already mentioned, moral phenomenology seems hardly transparent to the existence of a divine being playing a constitutive role for moral normativity. A statement like John Hare's at the outset of his theistic account of metaethics, "We want to say that value is created by God and is there whether we recognize it or not" (J. Hare 2001, ix), will not find the approval of someone not already committed to theism. The disvalue of pain for instance just seems to be there whether we recognize it or not, *and* it seems to be independent of God's creative activity (He might have created pain, but not the disvalue of it as part of a distinct act of creation in addition to the first one).

Ad (ii): Whereas acknowledging the existence of standards independent of God's essence or His will might compromise His omnipotence, *not* acknowledging such standards might compromise His supreme goodness. Living up to some arbitrarily self-imposed standards does not seem sufficient for laying claim to such an attribute. So even apart from saving the appearances of moral phenomenology, a theist has excellent reasons to keep moral normativity at least sufficiently independent from God so as to save the intelligibility of 'goodness' as one of His key attributes.

III.

Nearly all of the issues mentioned in the last section of this paper are the subject of intricate debates in contemporary metaethics and/or the philosophy of religion. My goal was not to solve those issues, but to provide the outlines of a dialectical framework in which their respective positions and complex interrelations become easier to focus on. In conclusion, let me offer some suggestions as to the role that should be attributed to divine authority in the theory of value and in metaethics. Before addressing this problem, some distinctions concerning (i) various *kinds* of divine authority, (ii) its exact *sources* and (iii) its *scope*, i.e. the area in which it is supposed to be exercised, are in order.

Ad (iii): As to the *scope* of divine authority, it might either cover the normative realm as a whole, or it might be restricted to certain parts of it. But how to justify making such distinctions within the normative realm? Two ways of carving up the normative realm recommend themselves in the present context: the first one according to whole *categories* of normative items, the other one according to the *content* of normative claims.

As to the first method, Elizabeth Anscombe, in her paper *Modern Moral Philosophy,* famously argued that deontic notions like 'morally right or wrong', 'moral oughts' or 'moral obligations' presuppose a divine law-giver.[7] In that respect, they are quite unlike aretaic notions like 'courageous' or 'temperate' that might lend themselves to an account that is based on the intrinsic normativity of human nature which is independent of any legislative act. Even if one hesitates to accept a sweeping thesis like Anscombe's, one has to admit that even non-theistic contemporary theories like Stephen Darwall's suggest that it takes a special, inter-personal (or, in Darwall's lingo, second-personal) framework to make sense of key deontic categories such as obligation. Stepping on the foot of one's fellow commuter might, according to Darwall, give one a "state-of-the-world-regarding" reason for removing one's foot and to bring the pain of the fellow commuter to an end (Darwall 2006, 5-10). This kind of reason however is quite distinct from the "second-personal" reason, which puts one under the authority of the person whom one has made to suffer. By harming y, x has conferred a special kind of authority to y;

---

[7] Cf. Anscombe 1958, 176: "Naturally it is not possible to have such a conception unless you believe in God as law-giver."

he is obliged to make amends to y as opposed say to turn the world into a better place by removing twice as much pain as that felt by y, but leaving his foot where it is. By presupposing such an intersubjective framework, the realm of deontic categories – quite unlike evaluative categories – might be non-arbitrarily singled out as special within the overall normative realm, and might lend itself to a theistic interpretation in a way that does not apply to intrinsic values or disvalues.

As to the second method, one might take the *content* of normative claims as one's starting point and distinguish for instance between non-religious und religious moral duties. Religious moral duties might in some way or other refer to God either directly as the object of duties of, for instance, devotion or gratitude, or indirectly insofar as, for instance, the violation of places or items dedicated to God might be a special offence (*sacrilegium*) on top of being merely, say, a run-of-the-mill sort of burglary.

Ad (ii): As to the *sources* of divine authority, all I can do here is point to some of the relevant options in the present context without being able to comment on their respective plausibility or importance. One might consider God's essence as constitutive of evaluative properties like goodness (by resemblance of finite entities to God, for instance) and his will as constitutive of deontic properties. In the latter case, it is still a hotly debated issue whether x's being morally obligated to phi depends on God's actual command that x phis or on God's willing that x be morally obligated to phi or just on God's willing that x phis.[8]

Ad (i): As to the *kinds* of divine authority, I suggest distinguishing between *epistemic*, *motivational* and *constitutive* authority. God's epistemic and motivational authority may be passed over without further comment in the present context. Suffice it to say that both kinds of authority are easily compatible with acknowledging the autonomy of ethics: God as omniscient has of course unique access to all ethical truths; he is familiar with the subvenient basis of such truths, with the ethical principles and the supervenience relations that apply in the moral field. God as a benevolent being might be expected to guarantee suitable epistemic access to those truths even to finite human beings; creating human beings but depriving them of access to at least fundamental moral

---

[8] For a critical discussion of those three options see Murphy 1998; Murphy himself argues in favor of the third option.

truths seems hardly compatible with Divine benevolence (the epistemic consequences of the self-inflicted Fall are of course another matter).

From a human point of view, however, God's authority is not confined to that of the teacher of moral truths; *qua* creator and judge of all things he carries a special motivational authority: Human beings will feel the motivational pull to obey God's commands even if their content is in itself sufficient to motivate them to act accordingly. Overdetermination seems possible if not required in such cases and threatens human autonomy in no way. As Robert Adams puts it (using a term originally coined by Paul Tillich): "The theonomous agent, in so far as he is right, acts morally because he loves God, but also because he loves what God loves. He has the motivational goods both of obedience and of autonomy" (Adams 1987b, 126).[9]

What about God's *constitutive* authority towards normative truths? It is here that the key challenge to the autonomy of ethics is located.[10] Let me conclude with four rather dogmatic theses on this crucial issue:

First of all, I think there are some basic axiological truths that are both necessary and not determined by either God's will or his essence. As mentioned above, any account of goodness incompatible with the idea

---

[9] This however leaves open the question of how both kinds of motives are related to each other: Is it that the acting from an explicitly theological motive takes normative precedence (a possibility discussed by Robert Audi 2007, 130); or is it a mark of the truly virtuous agent that the question of precedence actually never poses itself in the first place, since both motives always work in tandem?

[10] At this point it is of course the Euthyphro-dilemma that looms large: Is something right because God commands it (then God seems open to the charge of acting arbitrarily and we lose our grip on goodness as an attribute of God Himself) or does God command it because it is right (then God's omnipotence seems to be severely limited)? Audi's recent proposal to reject both horns of the dilemma actually strikes me as of little help: Audi argues that "God commands certain acts not because they are right but (at least in part) because of why they are right; i.e., because of the elements in virtue of which they are the right thing to do." (Audi 2007, 126) Thus God does not command something because it is obligatory – what makes him command it is the subvenient basis that grounds its being obligatory; nor does his commanding it make it obligatory – it is the subvenient basis that makes it so. Now, first of all, it would hardly be much of a relief to the defender of divine omnipotence to learn that is not normative reality itself that puts limits on God's omnipotence but the subvenient base properties on which that reality supervenes. Second, it seems mysterious why God should issue his commands not in virtue of their being right but because of those subvenient properties; these might be responsible for some commanded acts being right; but what recommends them for being commanded in the first place quite is obviously their normative status itself, i.e. that they are right, not those base-properties.

that at least some finite things have intrinsic worth, which supervenes on their non-relational properties and is thus independent of the relation in which they might stand to God, strikes me as implausible: It both distorts our sense of what is good or bad about instances of, say, pleasure and pain, and it leaves one puzzled as to whether any clear sense is to be made of an intrinsic goodness thus monopolized by God.

Second, the same goes for basic moral truths: Rossian *prima facie*-duties such as the prohibition of lying or the duty of reparation after having harmed someone are not in need of recourse to a divine authority in order to ground their normative force. *Pace* Anscombe, it is very well possible to believe in those deontic moral truths without presupposing a divine law-giver just as it is possible to have laws of geometry without considering them as the outcome of some legislative activity. This is of course bad news for someone who puts high hopes in the moral argument for the existence of God. Obligations like those of reparation might even constitutively presuppose interpersonal relations (unlike axiological truths that simply state for instance the badness of pain), but it does not take a divine law-giver to make such a relation possible: the person harmed and the person responsible for the harm will do perfectly for the job.

Third, even if God Himself figures in the content of some deontic moral truths (for instance in that we owe gratitude and love to our creator), it does by no means follow that those truths have to be constitutively dependent on Him. If for instance Duns Scotus is right, there is no possible world in which even God could will us to face Him with ingratitude or hatred.[11] It will therefore not do to consider necessary deontic moral truths like these – even if they have God as their content – as constitutively dependent on God (quite unlike truths that concern inter-human relationships that Duns Scotus considers as merely contingent).

Fourth, some deontic moral principles are indeed constitutively dependent on God's authority. If God has created the universe and keeps it in being at every single moment, then He does have the necessary

---

[11] Cf. Duns Scotus's interpretation of the first table of the Decalogue which contains our duties towards God: "So the commandments that tell us to love God have the kind of necessity required for natural law in the strict sense, but the commandments that tell us how to love our neighbor do not. They are extremely fitting Scotus says, but still contingent." (J. Hare 2001, 67)

authority to make some demands on His creatures – just like a human legislative body might be authorized to make its subjects drive on the left side of the road. So there is room for divine discretion. A classical example is the rules of worship: Why is it that the God of the Old Testament detests graven images in His worship? Even if there is no satisfactory theological explanation of the reasons God might have for this, it is certainly up to Him to determine which kinds of action are suitable to embody attitudes of devotion towards Him.

Where does all this leave us with respect to the three dimensions of the autonomy-problem distinguished above? *Firstly*, there is no reason to accept an ethics of autonomy as the only way to pay due respect to human autonomy; this holds independent of any theistic or anti-theistic assumptions. *Secondly,* ethics is mostly autonomous – not just key evaluative truths, but also deontic ones. Even many of the normative facts that cover the relationship between God and human beings are necessary and constitutively independent of any divine attitude towards them. *Thirdly,* the autonomy of ethics thus understood does not undermine God's omnipotence: (i) As 'companions in innocence'-style arguments show, necessary truths in areas other than morals do not restrict God's omnipotence: Necessary moral truths put as little limits on divine omnipotence as the *modus ponens* or truths of geometry do. (ii) In addition, even those necessary moral truths do not necessarily subject God to an *external* constraint: They might be part of his nature. Hence there is, as Kant already held for the 'Holy Will', no conceptual space for putting God under an *obligation*. Violating moral rules would simply be inconsistent with the divine character/nature. Since there are not even potentially counteracting forces in God's nature (like unruly inclinations or desires in the case of human beings), the very idea of moral demands putting pressure on God's will and thus undermining his sovereignty proves incoherent. (iii) It is well within the rights of God to issue moral laws purely out of his discretion: In this He is not acting arbitrarily. As much as an act of human legislation is only valid if backed by the authority of the legislative body to pass such legislation, God's authority to impose moral demands on human beings is not merely a function of His Omnipotence, but is grounded in His role as for instance Creator and Sustainer of all things. Just as the office of being fire-warden gives one the authority to oblige other people to leave their houses or even tear them down to prevent the fire from spreading, God's unique position towards

humanity is the source of legitimate moral obligations as opposed to pure, autonomy-undermining compulsion.

Acknowledging the autonomy of ethics to the extent suggested above is not only compatible with divine omnipotence, it has the additional advantage of buffering human autonomy against divine pressure. Moral demands are either obligatory in themselves (which is the case for most of them), or they are constituted by the exercise of God's legitimate authority. God could, of course, make the moral order and His own role as an omnipotent judge of all human beings so epistemically and motivationally overwhelming that there would be little room left for human autonomy. But this would in turn be incompatible both with God's nature and with His basic interest in free human beings who are able to freely cooperate with God in realizing His providential intentions. This corresponds to our deeply held moral convictions: Human beings have the responsibility to find out about their moral obligations, to think through their implications and to creatively apply them to new situations. They are even free to resist submission to the moral order both in its autonomous and its theonomic dimensions. But *not* resisting submission to rightful authority is not in the least less autonomous – indeed it is the rational way to respond to the normative demands that both created and uncreated divine reality makes on us.

BIBLIOGRAPHY

Adams, Robert. "Divine Command Metaethics Modified Again." *The Virtue of Faith and Other Essays in Philosophical Theology*. Oxford: Oxford University Press, 1987a. 128-143.

Adams, Robert. "Autonomy and Theological Ethics." *The Virtue of Faith and Other Essays in Philosophical Theology*. Oxford: Oxford University Press, 1987b. 123-127.

Adams, Robert. *Finite and Infinite Goods*. Oxford: Oxford University Press, 1999.

Anscombe, Elizabeth "Modern Moral Philosophy." *Human Life, Action and Ethics*. Eds. Mary Geach & Luke Gormally. Exeter/Charlottesville: St Andrews Studies in Philosophy and Public Affairs, 1958. 169-194.

Audi, Robert. "Divine Command Morality and the Autonomy of Ethics." *Faith and Philosophy* 24.2 (2007): 121-143.

Darwall, Stephen. *The Second-Person Standpoint*. Cambridge: Harvard University Press, 2006.

Hare, John. *God's Call*. Grand Rapids/Cambridge: William. B. Eerdmans, 2001.

Hare, Richard. "The Simple Believer." *Essays on Religion and Education*. Oxford: Clarendon Press, 1992. 1-36.

Johnson, Robert. "Value and Autonomy in Kantian Ethics." *Oxford Studies in Metaethics*. Vol. 2. Ed. Russ Shafer-Landau. Oxford: Oxford University Press, 2007. 133-148.

Korsgaard, Christine. *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.

McDowell, John. "Towards Rehabilitating Objectivity." *Rorty and His Critics*. Ed. Robert Brandom. Oxford: Wiley, 2000. 109-122.

Murphy, Mark. "Divine Command, Divine Will, and Moral Obligation." *Faith and Philosophy* 15.1 (1998): 3-27.

Quinn, Philip. "Religious Obedience and Moral Autonomy." *Religious Studies* 11.3 (1973): 265-281.

Rachels, James. "God and Human Attitudes." *Religious Studies* 7.4 (1971): 325-337.

Rorty, Richard. "Pragmatism as Anti-Authoritarianism." *A Companion to Pragmatism*. Eds. John Shook & Joseph Margolis. Oxford: Wiley-Blackwell, 2006. 257-267.

Shafer-Landau. Russ. *Moral Realism. A Defence*. Oxford: Oxford University Press, 2003.

Stern, Robert. "Constructivism and the Argument from Autonomy." *Constructvism in Practical Philosophy*. Eds. James Lenman & Yonatan Shemmer. Oxford: Oxford University Press, 2012. 119-137.

Wielenberg, Erik. "In Defense of Non-Natural, Non-Theistic Moral Realism." *Faith and Philosophy* 26.1 (2009): 23-41.

Wielenberg, Erik. "Review of *On the intrinsic Value of Everything* by Scott A. Davison." *Ethics* 123.1 (2012): 141-146.

DIVINE MOTIVATION THEORY AND EXEMPLARISM

LINDA ZAGZEBSKI

University of Oklahoma

Abstract. In this paper I summarize two versions of a new form of ethical theory in which all basic moral terms are defined by direct reference to exemplars of goodness. I call the Christian form Divine Motivation Theory in a book by the same name (Cambridge University Press, 2004), and the more general form I call Exemplarist Virtue Theory (Gifford Lectures 2015) or Exemplarist Moral Theory (forthcoming 2017, Oxford University Press). In the Christian form the supreme exemplar is God. In exemplarist virtue theory exemplars are superbly admirable persons or fictional characters, whose goodness is identified through the emotion of admiration rather than through the satisfaction of descriptive properties. In both versions of the theory the terms 'good person', 'virtue', 'good life', 'admirable act', and 'right act' are defined by the acts, motives, judgments, and attitudes of exemplary persons.

I. DIVINE MOTIVATION THEORY

In my book, *Divine Motivation Theory*, I proposed a theologically based moral theory according to which all moral properties of persons, acts, and outcomes of acts are defined by reference to the motives of God. I think of motives as states like love and compassion. They are either emotions or as similar to emotions as divine states can be. Human persons and their qualities are good in so far as they are like God or imitate God in the relevant respect. Human motives are good in so far as they are like the divine motives as those motives would be expressed in finite and embodied beings. Human virtues are those traits that imitate God's virtues as they would be expressed in human beings in human circumstances. Outcomes of acts get their value by their relation to good

and bad motivations. For example, a state of affairs is a merciful one or a compassionate one or a just one because the divine motives that are constituents of mercy, compassion, and justice respectively aim at bringing them about. Acts get their moral value from the acts that would, would not, or might be done by a being who imitates God in the relevant circumstances. God's own goodness and the rightness of God's own acts follow immediately from the theory since God himself is the supreme standard of all moral value.

Divine Motivation Theory (DMT) defines moral rightness and wrongness in a way that is structurally parallel to a form of Divine Command Theory (DCT), with the important difference that DMT defines rightness by reference to divine motives like love and compassion rather than to the divine will or divine commands. I believe that this allows DMT to avoid famous objections to Divine Command theory.

One problem for DCT is that if morality is grounded in God's commands, and if God can command anything, then it appears that God could command brutalizing the innocent. But that means that brutalizing the innocent could have been morally right, a very implausible consequence. In contrast, it is impossible that brutalizing the innocent is right in DMT as long as being loving is one of God's essential motives. The right thing for humans to do is to act on motives that imitate the divine motives. It is not possible that brutalizing the innocent imitates the divine motives as long as it is impossible for such an act to be an expression of a motive that is like the motives of God. If God is essentially loving, God's nature makes it impossible for him to have a motive that is imitated by brutalizing the innocent. Therefore, it is false that brutalizing the innocent could have been morally right according to DMT and this problem does not arise.

DMT also avoids the famous Euthyphro dilemma. The problem for DCT is that God's commands are either based on a reason or they are not. If they are, then the reason is the ground for moral rightness, not God's commands. If they are not, then God's commands are arbitrary. This problem does not arise in DMT. Although a command needs a reason, a motive *is* a reason. A divine motive does not need to be based on some *other* reason. A divine motive provides not only the impetus for an act, but is the justification of the act. So if God acts from a motive of love, there is no need to look for some further reason for the act. On the other hand, a divine command requires a reason, and if the reason is or includes divine motivational states such as love, then DCT needs

to refer to divine motives also.[1] DMT grounds morality directly in the motivational aspect of God's nature – God's love and other emotions, not his will.

In my original version of DMT, a moral duty is an act a person who imitates the divine motives would do in like circumstances, a morally wrong act is an act a person imitating the divine motives would not do in like circumstances, and a morally permissible act (a right act in one of its senses) is an act that a person imitating the divine motives might do in like circumstances.

DMT need not be distinctively Christian, but the Christian form of it makes the imitation of Christ central. In the doctrine of the Incarnation, Christ is the perfect revelation of the Father, and for us, the great metaphysical gulf between God and fallen humans is bridged in the person of Jesus Christ. We can imitate God because we can imitate Christ. Given the doctrine of the Incarnation, we can say, roughly, that the virtues are the traits that imitate Christ, good outcomes are the states of affairs at which persons with motives like those of Christ would aim, the wrong thing to do is what an imitator of Christ would not do, and so on. In *Divine Motivation Theory* I proposed that the way we make the imitation of Christ relevant to our individual natures and circumstances is by studying narratives of saints – imitators of Christ. So in the process of moral self-improvement we try to become imitators of imitators of Christ.

DMT also has advantages in solving puzzles in philosophical theology. There are a number of interconnected problems of potential inconsistency between the attribute of perfect goodness and other attributes. One problem is that perfect goodness traditionally includes the inability to do evil (or wrong). But that implies that God is not free in the sense we want for moral praiseworthiness since we do not call someone morally good if that being could not be other than good. Perfect goodness, then, seems to be incompatible with divine freedom in the morally significant sense of freedom. A second problem is that perfect goodness seems to be incompatible with omnipotence since the inability to do evil implies that God lacks a power. These problems lead to a third problem. If the concept of perfect goodness is meant to entail goodness in all its forms, including moral goodness, and if the concept

---

[1] DCT has the proviso that moral rightness is what is commanded by a loving God in the version defended by Robert Adams (cf. Adams 1975).

of perfect goodness is inconsistent with the concept of moral goodness because the latter entails the ability to do evil and the former entails the inability to do evil, then it seems to follow that the concept of perfect goodness is internally inconsistent.

These puzzles are solved if the goodness of all of God's attributes (indeed, the goodness of anything) derives from God's motives. Perfectly good power is the kind and degree of power God is motivated to have. If God is not motivated to have the power to do evil, then perfect power does not require the power to do evil. Perfect freedom is the kind of freedom God is motivated to have. If God is not motivated to have the freedom to do evil, perfect freedom does not require the freedom to do evil. God is perfectly good because God is motivated to be what he is. Perfection of all kinds derives from the perfect motives of a perfectly good God, not by reference to an independent standard.

DMT also leads to a reformulation of the problem of evil. What happens in the world is *ipso facto* something that is compatible with the divine motives, and cannot be deemed evil *if* evil is the conceptual opposite of good. Possible states of affairs that are directly contrary to the motives of God do not exist. However, many things occur in our world that seem to be incompatible with the motives of a loving God. I think that means that the problem of evil is not actually about the opposing concepts of good and evil and the apparent inconsistency between a good God and the existence of evil. It is about the apparent conflict between the motives of a loving deity with whom we would want to have a loving relationship, and the motives apparently exhibited in a suffering world. This way of looking at the problem of evil calls our attention to a metaethical assumption generally made by both sides of debates on the problem of evil. Typically these arguments assume that what makes a person and his motives good or bad is the goodness or badness of the states of affairs that person aims to bring about or to prevent. So the goodness or badness of a person's motives is derivative from the goodness or evil of states of affairs. DMT maintains the opposite. A state of affairs is a merciful one or a just one or a loving one because the divine motives of mercy or justice or love respectively aim at bringing about those states of affairs. This approach makes it crucial that we investigate the Christian tradition of revelation of the nature of God, as well as philosophical work on the divine nature in order to get an understanding of God's motives as they relate to the created world. In any case, my position is that we cannot infer the moral status of God's motives from our independent judgment

of the goodness or badness of states of affairs.[2] Narrative insights on the nature of a good person and a good life not only show the falsehood of that inference, but they give us vivid models of what the imitation of Christ looks like.

II. EXEMPLARIST VIRTUE THEORY

Divine Motivation Theory makes God himself the ultimate standard of goodness as the supreme exemplar. If an actual being is the standard, I argued that we can take the Putnam/Kripke theory of direct reference as a model for the initial move in constructing DMT (cf. Zagzebski 2004: 40-50), although I did not explore the implications of direct reference very far in that book. The basic semantical point is that the term 'good person' does not refer through a descriptive meaning – a person with certain descriptive qualities. Instead, we should think of 'good person' as referring to persons *like that* – Jesus, Confucius, Gandhi, Socrates, etc. We find out what persons like that are like by investigating them in personal experience and narratives. Virtues, right and wrong acts, and good outcomes of acts can be defined by reference to these persons. So roughly, a virtue is a trait of those persons, virtuous acts are acts that express those persons' traits, right and wrong acts are acts that persons like that would or would not do, and so on. The descriptive content of the concepts of a virtue, a right act, and a good outcome are determined after investigation of exemplary persons – persons in history or in fiction.

This approach has numerous advantages, and I subsequently decided to work on a general style of ethical theory I call exemplarist virtue theory, based on direct reference to good persons. I added the important element that exemplars of goodness are the persons who are most admirable, and that we pick out these people through the emotion of admiration (that has withstood reflection). In developing this theory, I have incorporated empirical studies on admiration and on exemplars, as well as numerous narratives of exemplars, such as stories about Holocaust rescuers and interviews with them, and psychological research on members of L'Arche communities who create and live in communities for the mentally disabled. I have linked the theory with a theory of moral education based on emulation of admirable persons, and I have argued that this kind of theory can bridge the gap between the theoretical purposes of moral

---

[2] I argue for this point in detail in Zagzebski 2016.

philosophy, and the practical purpose of motivating us to live good lives.

The theory includes a map of moral terms that is similar to the ones I proposed in DMT, with some additions and one important alteration:

(1) A *virtue* is a trait we admire in an exemplar. It is what makes *that* person admirable in a certain respect.

(2) A *good motive* is a motive of an exemplar. It is a motive of a person like that.

(3) A *good end* is a state of affairs at which exemplars aim. It is a state of affairs at which persons like that aim.

(4) A *virtuous act* is an admirable act, an act we admire in a person like that.

(5) An *admirable life* is a life lived by an exemplar, a person like that.

(6) A *desirable life* (a life of well-being) is a life desired by an exemplar. The components of a good life are *good for* a human person.

(7) A *right act* for person A in some set of circumstances C is what the admirable person would take to be most favoured by the balance of reasons for A in C.

(8) A *duty* in some set of circumstances C is an act an admirable person demands from both herself and others. She would feel guilty if she did not do it, and she would blame others if they do not do it.

These definitions are not intended to give the content of a series of concepts, nor are they intended to reveal the 'deep' nature of virtue, right action, or a good life. They do not tell us what a virtue, a right act, or a good life *is*, but they give us directions for finding out. They are like defining 'water' as 'stuff like that', or 'tiger' as 'a member of the same species as that', where the determination of the deep nature of water or tigers is left for empirical investigation. Similarly, the purpose of the definitions I have given is to permit us to identify the reference of moral terms in such a way that we know what to investigate to find out what virtue, right action, and a good life are.

There are more moral terms in the above list than the ones I proposed in DMT, and there is also a change in the definition of a right act. If we mean by a right act an act that is best supported by the moral reasons, I no longer think that a right act is the act that an exemplar or supremely virtuous person would *do* in some set of circumstances. Granted, there is

a tradition in many different cultures of using the actions of an exemplar as the touchstone for the act that ought to be done – a right act in the above sense. So in Zen Buddhism, the disciples are set a problem: What would the master do? Similarly, it was popular some years ago for many Christians to make moral decisions by asking themselves, 'What would Jesus do?' But there are reasons why this approach will not suffice for the purposes of constructing a way to define a right act by direct reference. For one thing, some circumstances are such that no virtuous person would be in those circumstances, so there is nothing the exemplar would do in those circumstances. What should you do if you break a promise? What should you do if you lie? What should you do if you failed to learn various virtues and have to live with the consequences? Furthermore, what the exemplar would do might be something that would put you in the way of a temptation that the exemplar could resist but you could not. It would not be right for you to imitate the exemplar in that case. Or maybe you could do what the exemplar would do, but you would not do it with the right spirit, and that could be worse than not doing it at all. For instance, Martin Luther King said that if people are unable to abide by the totality of the nonviolent approach to bus integration, they should stay off the bus and keep walking. (King 2003: 459, rule 9). In other words, it is better not to engage in protest at all than to do it without the spirit of non-violence.

I think this means that the right thing to do in the sense of the act that is best supported by the balance of reasons in some circumstances is determined by the judgment of the exemplar, not by the exemplar's behaviour. To repeat, this approach is not telling us what the property of rightness is, nor is it giving us the content of the concept of rightness. It is telling us what we should investigate to find out what is right and wrong. We should investigate the judgment of exemplars, and that is revealed in narratives about them and other forms of observation of them. What the exemplar does in coming to a judgment is left open. If the exemplar reasons from principles, we ought to be able to find that out. If the exemplar takes certain emotions as reasons for making a certain judgment, we ought to be able to find that out too. It is likely that there is more than one way to get to a judgment of what a person should do in a certain situation, and if so, it is helpful to us both as theorists and as moral learners to find out what the different paths to moral judgment might be.

I think that there are some interesting differences among three categories of exemplars: the saint, the hero, and the sage. Probably all three kinds of exemplars exemplify virtues, and we would refer to all of them to find out what constitutes different kinds of admirable lives, but some moral terms are most plausibly defined by reference to one kind of exemplar rather than another. For instance, 'right act' is most plausibly defined by reference to wise persons rather than, say, heroes. But wise persons, or sages, may reveal little about the variety of virtues and the scope of virtuous ends. There are also some differences among the exemplars recognized in different cultures and in different historical periods. Sages exist in many cultures, as do heroes, and the hero is especially important in ancient Greece, but arguably, the category of the saint made its first historical appearance with Christianity.

Exemplarism is a successor to the generic form of DMT I called motivation-based virtue theory in my 2004 book. The idea driving motivation-based virtue theory is that the motives of exemplars are metaphysically and conceptually basic. Moral goodness flows from the goodness of motives, which are emotion states. So motivation-based virtue theory might as well be called emotion-based virtue theory. It is a radical form of virtue ethics because all moral properties of persons, acts, and states of affairs are defined by reference to good motives or emotion states. In the general non-theological version, emotion states are intrinsically good or bad. The theological form of the theory has a metaphysical anchor in the motives of God. Human motives derive their goodness from their imitation of God's motives. But both the theological and non-theological versions are motivation based because motives are the most basic bearer of moral properties.

I have added several elements to exemplarism that make it different from motivation-based virtue theory and its theological version. I have already mentioned one difference, that in exemplarism, exemplars of goodness are picked out through the emotion of admiration. I argue that it is through reflecting on what we admire that we are able to say that motivational structure is the basic feature of good persons in virtue of which they are good, or admirable. This is parallel to the position that it is the deep physical structure of water that makes it water. We pick out water by superficial properties of taste and appearance, but we think that those properties are not what makes it water. Rather, deeper physical properties both explain the existence of the superficial properties, and are what makes water what it is. Similarly, we pick out good persons

by easily observable behavioural properties, but we think that deeper psychological properties both explain the behavioural properties, and are what makes a good person good. The importance of motivational structure is determined by reflection on what we admire about admirable persons. Exemplarism therefore has an underlying explanatory layer missing from motivation-based virtue theory.

There are also additions to the semantics of the theory. Exemplarist virtue theory is externalist in ways that parallel the externalism of Putnam. Putnam maintained that what we mean when we use a natural kind term (and some other terms as well) is determined by something outside of us in two ways. First, it is partly determined by the world because the indexical feature of meaning has the consequence that (part of) what we mean by a term is *that*. Water is stuff like that, dogs are animals like that, and so on. So a difference in extension is sufficient for a difference in what we mean in the use of a term.[3] Second, what we mean is partly determined by a social linguistic network that links us to the extension of the term. Even though we are expected to grasp a vague description (what he calls a 'stereotype') in order to be a competent user of a term, a description in the head is not necessary to fix the extension, and is far from sufficient. Given his principle of the Division of Linguistic Labour, ordinary speakers defer to experts to both identify the objects in the extension, and to find out what the deep structure of a given kind is. In this way ordinary speakers are dependent upon others in the network for their semantic success. What we mean when we use natural kind terms is not up to us. It is up to the world and it is up to certain other people.

I think that 'good person' and virtue terms are externalist in both of these ways. What 'good person' refers to is partly determined by the way the world is – by the features of exemplars awaiting our discovery. It is also externalist because I think there is a Division of Moral Linguistic Labour. We refer to good persons through a network that connects us to admirable persons through other users, some of whom have a privileged

---

[3] That was the point of Putnam's famous Twin Earth example. In that thought experiment, we imagine a planet exactly like earth except that the liquid we drink and in the oceans and rivers and falling from the sky is not H_2O, but is another substance, XYZ. XYZ has the same properties of taste and appearance as H_2O, it is indistinguishable to the ordinary observer. Putnam says that if the Twin Earthians point to XYZ when they say 'water', while we Earthians point to H_2O when we say 'water', we are not disagreeing about the nature of water. We are talking about two different substances. In fact, Putnam thinks we *mean* something different by 'water' than they do (Cf. Putnam 1975: 139-144).

function. Each user of the term 'good person' needs to be able to grasp
the stereotype of a good person. Virtue terms function as descriptors
that are part of the stereotype of a good person, and they are important
for communication among the members of a community, but they do
not provide necessary and sufficient conditions for membership in the
kind. It is not necessary that an ordinary user can give an account of
virtue or particular virtues in order to acquire the use of a virtue term.
Different people have different roles in the network. I don't think there
are moral experts in the same sense as scientific experts, but there are
specialized functions for religious leaders, psychologists, philosophers,
and others. The use or disuse of a term in the media can change the
meaning of a virtue term. I think that the meaning of 'virtue' already has
changed and is on the verge of going out of use. A linguistic community
can expand, and sometimes the extension of a term can change.[4] I also
argue that the overlap of moral linguistic communities is a necessary
condition for agreement in contested areas of moral judgment.[5]

III. GOD IN EXEMPLARIST VIRTUE THEORY

Divine Motivation Theory proceeded from an assumed background
of Christian monotheism. That included acceptance of the traditional
divine attributes and divine personhood, acceptance of the doctrine of
the Incarnation and its place in Christian ethics, and a desire to integrate
the tradition of philosophical reflection on divine goodness and the
problem of evil into the theory. I did not give an account of how we
identify exemplars, nor did I discuss semantic externalism. I would like
to end with a short discussion of the advantages to Christian ethics of
blending the new features of exemplarism with DMT.

In exemplarism we determine what is admirable through the emotion
of admiration when it survives reflection. I have argued in another place
(Zagzebski 2012: ch. 4) that there is a general problem of psychic circularity
that parallels the problem of epistemic circularity, and this applies to the
connection between what we admire and the admirable. Just as we have
no way to tell that our beliefs are true without relying upon our faculties

---

[4] An example of the expansion of a linguistic community occurred when the Universal
Declaration of Human Rights was adopted by most of the countries of the world. The
community expanded with respect to the use of the term 'rights', and the stereotype of
a right changed. It is possible that the extension changed also.

[5] I argue for the Division of Moral Linguistic Labour in Zagzebski 2017: ch. 7.

as a whole, we have no way to tell that what we admire upon reflection is admirable without relying upon our faculties as a whole, including our emotion dispositions. We need trust in our emotion dispositions for the same reason that we need trust in our epistemic faculties. In both cases we have a psychic state with an external object that it can fit or not fit. But we can never tell that either our emotion states or our belief states are fitting without relying upon the dispositions that produce those states together with our other faculties. I argue that such trust is rational as well as inescapable.

Exemplarism makes admiration the power we use to identify exemplars, the most admirable persons. The theory depends upon trust in that power. I believe it is rational to trust that power, but blending DMT with exemplarism strengthens the theory. In DMT exemplars are imitators of God, the perfectly good being. God is the object of the highest admiration, which we normally designate by a different word – adoration.

If God is the supreme exemplar, that aids exemplarism in at least two ways. First, it gives us a stronger ground for trust in our emotion of admiration, as well as a way to critique that emotion. In a Christian world view, our faculties are designed to fit their objects and so they are generally trustworhty, but there are lots of ways we can make mistakes. Likeness to Christ is a test for the appropriateness of our admiration in particular cases.

Second, DMT gives us reason to think that even though there are many different kinds of exemplars recognized by individuals and different communities, exemplars are connected to each other by their likeness to the divine Exemplar. This supports my contention in *DMT* that dialogue between exemplars in different cultures is the best way to secure moral agreement. It is a tremendous challenge in a pluralistic world to find grounds for optimism in getting agreement when there are enormous historical and psychological differences among the people who disagree. The theological metaphysics of DMT explains why hope for agreement is not in vain. There are genuine commonalities among the exemplars of different cultures and different historical periods that are explained by their relation to the supreme Good.

The semantic externalism of exemplarism also strengthens DMT. In the first book I implied an externalist semantics arising from the indexical feature of moral terms, but I did not discuss externalism for its support of moral realism. Although exemplarism is intended to be neutral on

meta-ethical issues of realism and cognitivism, I argue that it can be used to support at least a moderate ethical realism for the same reason that the theory of direct reference has been used to support scientific realism. Further, I argue that there can be necessary *a posteriori* truths about morally good persons for reasons that parallel similar arguments by Saul Kripke, Keith Donnellan, and Nathan Salmon (cf. Zagzebski 2004: ch. 8; Zagzebski 2017: ch. 8).[6] The blending of DMT with the semantics of exemplarism yields a stronger argument for the possibility of necessary *a posteriori* moral truths than we get from exemplarism alone. In DMT, as in traditional Christian theism, God's goodness is essential to him, and the components of that goodness – lovingness, compassion, etc. – are necessary to his goodness. It is a necessary truth that a good person is loving because it is a necessary truth that the ground of goodness – that divine being, is loving. We find out that God is loving *a posteriori*; it is not something we know *a priori* by reflection upon the concept of God. At least, that is my hypothesis.

Exemplarist virtue theory is a more developed theory than motivation-based virtue theory was in my 2004 book, and I think that exemplarism leads to a fuller and more interesting version of Divine Motivation Theory. I hope that my sketch of these theories indicates some ways in which they strengthen each other for the purposes of Christian moral philosophy.

BIBLIOGRAPHY

Adams, Robert M. 1975. 'A Modified Divine Command Theory of Ethical Wrongness', *Religion and Morality*, ed. by Gene Outka and John Reeder (Garden City: Anchor Press)

King, Martin Luther. 2003. 'Stride Toward Freedom', *A Testament of Hope: The Essential Writings and Speeches of Martin Luther King Jr.*, ed. by James Melvin Washington (New York: HarperCollins, 2003), pp. 417-490

Putnam, Hilary. 1975. 'The Meaning of "Meaning"', *Mind, Language, and Reality* (*Philosophical Papers* vol. 2), ed. by Hilary Putnam (Cambridge: Cambridge University Press)

Salmon, Nathan. 1981/2005. *Reference and Essence* (Amherst: Prometheus Books)

---

[6] For a subtle discussion and overview of the mechanisms used for deriving necessary *a posteriori* truths by Kripke and Donnellan, with Salmon's reconstruction and commentary, cf. Salmon 2005.

Zagzebski, Linda. 2004. *Divine Motivation Theory* (Cambridge: Cambridge University Press)

Zagzebski, Linda. 2012. *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (New York: Oxford University Press)

Zagzebski, Linda. 2016. 'Good Persons, Good Aims, and the Problem of Evil', in *Applying Contemporary Ethical Theory to the Problem of Evil*, ed. by James Sterba (Bloomington: Indiana University Press, forthcoming)

Zagzebski, Linda. 2017. *Exemplarist Moral Theory* (Oxford: Oxford University Press, forthcoming)

EVIDENTIALLY COMPELLING RELIGIOUS EXPERIENCES AND THE MORAL STATUS OF NATURALISM

TRAVIS DUMSDAY

Concordia University of Edmonton

Abstract. Religious experiences come in a variety of types, leading to multiple taxonomies. One sort that has not received much attention as a distinct topic is what I will call 'evidentially compelling religious experience' (ECRE). The nature of an ECRE is such that if it actually occurs, its occurrence plausibly entails the falsity of metaphysical naturalism. Examples of ECREs might include visions / auditions / near-death experiences conveying information the hearer could not have known through natural means, later verified; unambiguously miraculous healings; fulfilled prophecy; supernatural rescues; inter-subjective religious experiences (e.g., multiple people simultaneously having the same vision of the Virgin Mary), etc. After presenting a representative set of published case studies of ECREs, I argue that for most settled metaphysical naturalists (though not all), the combination of a settled metaphysical naturalism with an awareness of the relative commonality of testimony to ECREs is either irrational or immoral. This is because that conjunction entails either an unjust and uncharitable judgement on a great many of those testifying to ECREs (namely that they are liars), or an irrational refusal to acknowledge this entailment.

I. INTRODUCTION

There are many competing definitions of 'religious experience,' just as there are many of 'religion'; Franks Davis' (1989: 29) assessment seems plausible: 'Because there are so many religious traditions and so many types of experiences within those traditions, I look upon the quest for a neat, precise definition of "religious experience" [...] as fruitless. Most people have a working idea of what counts as a religious experience, based on the many uncontroversial examples available.' I will take the

presence of such a working idea for granted in what follows. Still, even without providing necessary and sufficient conditions constituting an experience as 'religious', we can develop taxonomies of distinct types of experiences plausibly counted as such. One such distinction is that between religious experiences the content and context of which leave open the question of a supernatural origin versus those the content and context of which seem to imply a supernatural origin. It is safe to say that most reported religious experiences fall within the first category, but it is undeniable that there are also many, many instances of testimony to what I will here refer to as 'evidentially compelling religious experience' (ECRE). If these really are experienced, then metaphysical naturalism is plausibly falsified.

I will provide a number of examples of published claims to ECREs below, but to help clarify the distinction initially one might contrast (1) someone's claim to have heard a message from God during intense prayer, a message of moral support (which audition could perhaps, in theory, have been the product of self-delusion, mental illness, or a variety of other naturalistic factors), and (2) someone's claim to have heard a message from God during intense prayer, warning her to dodge to the left because a gunshot was about to come through the wall (fired from a passing vehicle perhaps), which warning was taken to heart and which prediction immediately and exactly came true, saving the individual from harm. Claims of the latter kind are by no means unheard of. If the individual did not hallucinate the event / is not mentally ill / has not unconsciously manufactured a false memory / is not lying about it, etc., such that her report of the experience is accurate, then plausibly metaphysical naturalism has been falsified. ('Metaphysical naturalism' is understood here as involving at least the claim that there are no *supernatural entities*, which in turn implies minimally that there are no *wholly non-physical / spiritual agents* – thus no gods, ghosts, angels, Cartesian egos, etc.) Experiences of that sort are ECREs.

Of course, for those who take the Duhem / Quine thesis and related ideas seriously, the data could *somehow* be made to fit with naturalism, just as any data can be creatively re-interpreted to accord with any hypothesis. On such a view no theory is ever definitively falsified. Even in the example just cited, one could posit that there may be a bizarre, heretofore undiscovered but thoroughly natural human capacity for perceiving future oncoming dangers (a 'spider sense'?) that a person unconsciously recasts as the voice of God. Or perhaps one could maintain

it was just a truly one-in-a-million coincidence.[1] For the Duhem / Quine proponent, such a claim cannot be definitively disproved. But even if one adopts such a strong understanding of the underdetermination thesis, there can still be *plausible* falsification of an hypothesis, or in slightly different terms, *rationally adequate* undermining of an hypothesis, however difficult it may be to lay out criteria for when precisely that takes place. So, wishing to leave debates about underdetermination aside, I will here stick to talk of 'plausible' or 'rationally adequate' falsification.

What sort of attitude must a settled naturalist take to testimony of ECREs? (And by 'settled' naturalist I simply mean someone with a firm commitment to the truth of metaphysical naturalism, in contrast to someone adopting it merely tentatively. To make that rather vague characterization a bit more precise: a necessary condition for being a settled metaphysical naturalist is a belief that there is no rational or other requirement to reconsider seriously the reality of the supernatural.) He/she must believe that all such testimony is unreliable, that the subject was either honestly mistaken about the content of the experience (again, perhaps via some sort of false memory), or hallucinating due to drug use or physiological defect, or mentally disturbed, or lying. A great many claimed ECREs do not plausibly admit of the first few sorts of explanation (as will become evident through a number of the published case studies below), leaving the final option – intentional deception – as the only rationally available naturalistic explanation.

I will argue that an awareness of the widespread nature of testimony to ECREs places most settled naturalists in a difficult position: they must either judge a very great number of individuals claiming ECREs as liars, which is immoral (for reasons I will elaborate on), or they must suspend judgement on the question, which in this context would be irrational (since inconsistent with *settled* naturalism). Consequently, for most (not all) settled naturalists who are well-informed concerning contemporary testimony to religious experience, it is either immoral or irrational to maintain their settled naturalism. They should instead shift from settled naturalism to tentative naturalism, where 'tentative naturalism' involves at least the belief that there is a rational or other requirement to reconsider seriously the reality of the supernatural.

---

[1] That option could plausibly be undermined by further tweaking the hypothetical case – e.g., specify that the warning voice said it would be a gunshot from a .357 magnum, which calibre the police later verify upon digging the bullet out of the wall.

The paper is structured as follows: in the next section I present a range of published testimonies to ECREs. This is designed to familiarize the reader with these experiences, and to provide some idea of just how common they are in the literature. (Those presented below constitute but a small sampling.) Then in section three I develop the argument that most settled naturalists, upon becoming aware of the commonality of such testimony, face a dilemma between immorality and irrationality. Section four sees a discussion of various objections, and in the concluding fifth section I briefly draw some links between the present argument and some other treatments of the supposed immorality of naturalism.

II. PUBLISHED CASE STUDIES OF ECRES

The following selection of published ECREs is culled from a much wider set that the author has encountered. They are taken from a variety of kinds of publication: academic sources, popular / devotional religious literature, etc. A selection of six ECREs is provided here; I expect some readers will not wish to wade through all six. However, in order to grasp the nature and import of these experiences there is no substitute for a broad exposure to actual testimonies.

First Case Study

This case is excerpted from van Lommel, van Wees, Meyers, and Elfferich (2001: 2041), a study on cardiac patients' reports of near-death experiences published in one of the world's leading medical journals:

> During the pilot phase in one of the hospitals, a coronary-care-unit nurse reported a veridical out-of-body experience of a resuscitated patient: 'During a night shift an ambulance brings in a 44-year-old cyanotic, comatose man into the coronary care unit. He had been found about an hour before in a meadow by passers-by. After admission, he received artificial respiration without intubation, while heart massage and defibrillation are also applied. When we want to intubate the patient, he turns out to have dentures in his mouth. I remove these upper dentures and put them onto the 'crash car'. Meanwhile, we continue extensive CPR. After about an hour and a half the patient has sufficient heart rhythm and blood pressure, but he is still ventilated and intubated, and he is still comatose. He is transferred to the intensive care unit to continue the necessary artificial respiration. Only after more than a week

do I meet again with the patient, who is by now back on the cardiac ward. I distribute his medication. The moment he sees me he says: 'Oh, that nurse knows where my dentures are'. I am very surprised. Then he elucidates: 'Yes, you were there when I was brought into hospital and you took my dentures out of my mouth and put them onto that car, it had all these bottles on it and there was this sliding drawer underneath and there you put my teeth.' I was especially amazed because I remembered this happening while the man was in a deep coma and in the process of CPR. When I asked further, it appeared the man had seen himself lying in bed, that he had perceived from above how nurses and doctors had been busy with CPR. He was able to describe correctly and in detail the small room in which he had been resuscitated as well as the appearance of those present like myself. At the time that he observed the situation he had been very much afraid that we would stop CPR and that he would die. And it is true that we had been very negative about the patient's prognosis due to his very poor medical condition when admitted. The patient tells me that he desperately and unsuccessfully tried to make it clear to us that he was still alive and that we should continue CPR. He is deeply impressed by his experience and says he is no longer afraid of death. 4 weeks later he left hospital as a healthy man.'

Second Case Study

The following is taken from Morse and Perry (1990: 1-5). Morse is a paediatrician who, among his other duties, has studied near-death experiences among children. He here tells of his first encounter with a child reporting an NDE:

> I stood over Katie's lifeless body in the intensive care unit and wondered whether this little girl could be saved. A few hours earlier she had been found floating facedown in a YMCA pool. No one knew how long she had been unconscious or exactly what had happened to cause her to lose consciousness .... I didn't really expect to find out what had happened. The machines to which she was now hooked up told a grim story. An emergency CAT scan showed massive swelling of the brain. She had no gag reflex. An artificial lung machine was breathing for her .... Looking back even now, I would guess that she had only a ten percent chance of surviving. I was the doctor who resuscitated her in the emergency room .... [Morse then recounts a prayer vigil held by the child's immediate family.] Three days later she made a full recovery.

Her case was one of those medical mysteries that demonstrate the power of the human organism to rebound …. When she was feeling well enough, I had her come in for a follow-up examination. One of the things I wanted to know was what she remembered about her near drowning. The answer was important to the type of treatment she would receive as an outpatient. Had she been hit on the head? Had someone held her under water? Had she blacked out or experienced a seizure? …. Katie clearly remembered me. After introducing myself, she turned to her mother and said, 'That's the one with the beard. First there was this tall doctor who didn't have a beard, and then he came in.' Her statement was correct. The first into the emergency room was a tall, clean-shaven physician named Bill Longhurst. Katie remembered more. 'First I was in the big room, and then they moved me to a smaller room where they did X-rays on me.' She accurately noted such details as having 'a tube down my nose', which was her description of nasal intubation. Most physicians intubate orally, and that is the most common way that it is represented on television. She accurately described many other details of her experience. I remember being amazed at the events she recollected. Even though her eyes had been closed and she had been profoundly comatose during the entire experience, she still 'saw' what was going on. I asked her an open-ended question: 'What do you remember about being in the swimming pool?' 'Do you mean when I visited the Heavenly Father', she replied. Whoa, I thought. 'That's a good place to start. Tell me about meeting the Heavenly Father.' 'I met Jesus and the Heavenly Father', she said. Maybe it was the shocked look on my face or maybe it was shyness. But that was it for the day. She became very embarrassed and would speak no more. I scheduled her for another appointment the following week. What she told me during our next meeting changed my life.

Third Case Study

John White is a psychiatrist and Christian author. He describes an apparently miraculous healing in which he and his spouse participated (1988: 122):

I remember praying with my wife for a two-year-old child in Malaysia. Her body was almost completely covered with raw, weeping eczematous areas. She ran around the room restlessly so that her parents had to catch her to bring her struggling to us. We began to pray and extended our hands to lay them on her. The instant our hands touched her she fell into

profound and relaxed slumber in her parents' arms. But there was more to follow. I shall never forget our sense of exhilaration and excitement as the weeping areas began to dry up, their borders shrinking visibly before our eyes like the shores of lakes in time of drought ... A person who has never experienced the impact of such a sight has no idea of its effect on one's emotions ....

Being a medical doctor, White presumably is a competent judge of the apparently non-natural manner in which the healing took place (esp. with respect to speed and timing).

Fourth Case Study

Leanne Hadley is a United Methodist minister and former chaplain at a children's hospital. In a book on her chaplaincy experiences, she recounts an interaction with a ten-year old boy who was awaiting surgery to remove a brain tumour. His attendant nurse was concerned that he manifested a fear of death and poor attitude to the surgery, and asked Hadley to talk to him (2013: 59-60):

I went into his room, expecting to see an upset or angry child. But James was sitting up and smiled as I walked into the room. We talked for a few minutes, and there was nothing abnormal about him. He wasn't upset, he wasn't angry, and he wasn't depressed. I asked him how he felt about the surgery, and he said fine. I asked if he was upset because they would have to shave off even more of his hair, and he wasn't. I saw no sign of a bad attitude. Finally I told him that I had been called in because his nurse was worried that he had a negative attitude about the surgery. He had no idea what I was talking about. I decided I needed to ask his nurse what he had said that made her so upset. As I was leaving the room, James said, 'Leanne, I need to ask you something.' 'Yes?' I replied. 'Is there time for me to be baptized before my surgery?' 'I think so', I said. 'But why? Is there any special reason you want to be baptized?' 'Because God told me to in my dream', he replied nonchalantly. It began to dawn on me that this dream might have been what he had shared with the nurse.

He goes on to tell her about meeting Jesus in a dream, who shows him a glimpse of heaven (ibid.: 61-63):

Jesus asked me if I wanted to cross the bridge with him, I did! I wanted to go wherever he went! So he took my hand and we crossed the bridge, and he showed me heaven. Leanne, heaven is so beautiful. It isn't what

you expect, because we have never seen anything as beautiful as it is .... I don't know how long I stayed in heaven, but Jesus told me it was time to go home. Then he told me that he showed me heaven because I would be coming here in three months and he didn't want me to be worried or be afraid. He said my surgery will go fine, but I'll still die in three months. He promised to meet me on the bridge again .... And he told me that I needed to be baptized before surgery. So can I be baptized?' [James and his family were Baptists, who delay administration of that sacrament until the age of reason.] I was speechless .... 'Yes, of course', I said, coming back to reality. 'I mean, we have to ask your parents, but I can't imagine they won't agree.' I went back to the nurse and told her that James wanted to be baptized before surgery. We had only a few hours until he would be taken to the operating room. She didn't want him to be baptized. 'He's preparing to die, and he needs to have a positive attitude for surgery', she said. 'He has no fear of dying in surgery', I reminded her .... The nurse reluctantly called his parents .... They, of course, agreed .... And the surgery was a success. James recovered quickly and was released from hospital within a few days .... About two months after his release from the hospital, he had started having trouble swallowing and speaking. He'd been taken into the cancer centre for more tests, and it was discovered that his brain tumour had some fingers that were embedded deeply in his brain. They could not be removed, because doing so would have destroyed his brain and the surgery itself would have killed him. James's mother said that he never showed any signs of anxiety or fear .... It was three months after his surgery when he died.

Fifth Case Study

The Religious Experience Research Centre, formerly housed at Oxford, now at the University of Wales, has for more than forty years collected and archived reports from the general public. They periodically publish collections of these reports. Here is one from the anthology edited by Maxwell and Tschudin (1990: 77-78):

M. & F. 81 (26) #3015[2] This happened in the year when Lord Derby was calling for army recruits, before conscription, for World War I. We were 'boy and girl' friends, and married when the war was behind us.

---

[2] These numbers record the gender of those involved in the experience, the age of the individual at the time the report was made, the age of the individual at the time the experience occurred, and a # for archival reference of the report.

My husband-to-be knew he would have to enlist and he dreaded the idea, but he did not want to be a conscript, so decided to join the Derby scheme that summer. It was a Saturday evening, a glorious evening, with a most brilliant moon, when he asked *my* opinion about this 'joining up', but I refused, as I said if anything happened to him, I should always blame myself! We were on a country lane near my home, now built up, but then pasture land on both sides, which I knew well on each side of the road. As we walked along, scarcely speaking, as the decision was little nearer, we suddenly saw a brilliance most unusual, even more than the loveliest moon I have ever seen: there was an opening in the stone wall, with much more light than the moon, the width of a farmyard gate, which I knew a gate did not exist for several yards further down the road. Then a figure emerged, a most brilliant sight. We were both speechless, but not afraid, it was so beautiful. The figure, Jesus Christ, glided on to the centre of the road while we were on the rough pavement. We were spellbound as the figure walked up and we were walking down. We could see the white gown with a broad, twisted girdle around his waist, knotted and falling down his left side. The figure glided along, but we could see no feet, and as it got nearer, we tried to make out his face features, but could not, and as it got level with us, it gradually faded away from the bottom of the gown up to the head, and it had vanished! When we got down to the 'wall opening', it was the solid wall, as I knew it for many years, further down, and it had got to near midnight; but, still, we never spoke to each other, we were spellbound! When we got to my home, as he was about to leave me, we at last spoke, cross-questioned each other on what we had seen, without a hint on the matter, each and every answer coincided. We still remember every detail, but our views on religion have deepened; although, still, we are not *too* religious. When I got in, home, I told my mother about this strange happening, and her remark, 'Oh! If he goes a soldier, he will come back all right!' Funnily enough, he got a 'nasty packet' of 'poison gas' out in France one night when sleeping. He was sent to hospital for treatment for six months, then back on service, and left the army A1, we are thankful to say, when war ended.

Sixth Case Study

For her doctoral dissertation in anthropology at the University of Birmingham, Emma Heathcote-James decided to study contemporary belief in angels. As part of her research, she solicited reports of encounters with what people took to be angels; initially the requests

appeared just in local church newsletters, then circulated much more broadly throughout the UK as public interest in the project mounted. She received more than 800 accounts from across the country. There is actually quite a variety in the reports; some tell of seeing beings that do accord with Biblical and other traditional depictions (about 30% of the reports); others relate audible messages, supernatural rescues or warnings, etc., that are interpreted by the recipients as being of possible angelic origin. Here's an example of the latter kind (2002: 41-42):

> Paul Dunwell lost control of his motorbike going round a bend at 70 miles per hour in the dark. Minutes before, he had passed a red sports car (also going at speed). Paul wrote: 'All I could think of was I'd risked [the life of] the driver now that my bike lay in his path on the blind bend. Telling myself that my own idiocy had brought all of this about, I resolved [to try] to move the bike despite its weight and my hands being a bit pulped. Time was, I knew, running out .... then an odd thing happened .... [the car] appeared on the bend at a snail's pace, his hazard warning lights already on. He stopped there and ran to me and said "You'll never guess what's just happened .... there was a light in my car. And I was told, like there was somebody there in the car with me, and actually shown that you were lying there in the road .... I was told I mustn't hit you."

III. SETTLED NATURALISM AS IMMORAL OR IRRATIONAL

The experiences listed above all count as ECREs because for each the following holds true: if the event really transpired as the person reports having experienced it, then metaphysical naturalism is plausibly falsified. Moreover, for all of these examples it is implausible to suggest that the subject of the experience honestly misinterpreted the nature of the experience. Likewise it is implausible to think that all the subjects of these experiences suffer from a mental illness serious enough to prompt major delusions. And where a substantial amount of biographical information is available, the lack of relevantly serious mental illness seems empirically established. Unconscious implanting of false memories also will not work for many such cases, assuming the basic mental health of the individuals reporting them. Individuals of normal mental health may misremember past events in certain respects, but do not unknowingly confabulate entire series of events, let alone events of substantial existential significance (excluding cases where false memories

have been implanted via hypnosis). Excluding those explanations, what can a settled metaphysical naturalist say about such cases? It seems the only plausible option left is to accuse the subjects of deception.

And what's wrong with that answer? If deception is the only plausible explanation remaining, surely we ought to have no hesitation in calling these people liars. They invite the designation when they make claims to ECREs, and make those claims in such a way that we are forced to choose between affirming naturalism and affirming the possibility of their good character. Moreover, in the past a number of prominent philosophers have not hesitated to attribute deception in such cases. Hume, for instance, famously argues that it is always more likely that a person is lying than that a miracle took place (or at the very least that the two probabilities cancel each other out), such that we can always dismiss any testimony to a miracle.[3]

Indeed, there is a certain sort of naturalist for whom such a reply is perfectly acceptable, namely the sort for whom deception literally is the only rational explanation left. That is, there are some naturalists who are not merely settled in their naturalism, but who think that any sort of non-naturalism is absurd – who think that belief in the God of classical theism (for instance) really is equivalent to believing in the Flying Spaghetti Monster, or that belief in an immortal human soul is equivalent to belief in leprechauns. Correspondingly, from their perspective, to make an accusation of deception in response to a claimed ECRE is no more unjust, uncharitable, slanderous or whatnot than to make it in the case of an alleged leprechaun sighting. In both cases, the accusation is justified, morally and rationally, by the absurdity of what is being claimed.

To those naturalists, the following argument has no application. However, it does apply to most self-professed settled naturalists, because most would claim that while belief in theism (for instance) is mistaken, nevertheless it is neither absurd nor a proposition whose falsehood is absolutely certain. Most would even admit to having occasional, fleeting doubts about their own atheism, or speculative thoughts about the possibility of a life after death, etc. God and the soul are rejected by them, but not rejected as absurdities, and not rejected with airtight confidence. Nevertheless, they remain *settled* naturalists because think

---

[3] See section 10, part 1, of *An Enquiry Concerning Human Understanding* (1975/1748: 115).

there is no rational or other requirement to reconsider seriously the truth of non-naturalism. They feel no need to pick up a book of miracle-testimonies and subject it to serious scrutiny, or to keep up with the latest developments in natural theology. They have a firm, basically stable confidence in their worldview. The following argument is aimed at that broad category of settled naturalists:

Assumption 1 – For a great many ECREs, deception is the only rationally available naturalistic explanation.

Assumption 2 – It is wrong (because unjust, uncharitable, and a violation of the Golden Rule) to attribute deception to someone (especially on a matter of great importance) without adequate evidence of deception having been committed.

Premise 1 – *If* most settled naturalists view metaphysical non-naturalism as neither absurd nor dismissible with certainty, *then* for them the mere claim to an ECRE is not by itself adequate evidence of deception (even where deception is the only rationally available naturalistic explanation).

Premise 2 – *If* for them the mere claim to an ECRE is not by itself adequate evidence of deception (even where deception is the only rationally available naturalistic explanation), *then* it is wrong for them to attribute deception to the person claiming an ECRE without further, adequate evidence of deception.

Premise 3 / Conclusion 1 – Therefore *if* most settled naturalists view metaphysical non-naturalism as neither absurd nor dismissible with certainty, *then* it is wrong for them to attribute deception to the person claiming an ECRE without further, adequate evidence of deception.

Premise 4 – Most settled naturalists view metaphysical non-naturalism as neither utterly absurd nor dismissible with certainty.

Premise 5 / Conclusion 2 – Therefore, it is wrong for them to attribute deception to the person claiming an ECRE without further, adequate evidence of deception.

Premise 6 – The settled naturalist who is aware of the relevant facts about ECREs can only retain settled naturalism by either acting against the moral stricture just laid, or by ignoring premise 5 / Conclusion 2 altogether.

Premise 7 – The previous premise entails the following dilemma: if the first of the two options is taken, an immoral act will be committed, since that strategy involves attributing deception where it is immoral to doso.

If the second of the two options is taken an irrational attitude will be adopted, since Premise 5 / Conclusion 2 implies that ECREs require further investigation prior to dismissal, and recognizing a need for such investigation is incompatible with settled naturalism.

Final Conclusion – Therefore, for most settled naturalists, it is either immoral or irrational to continue being a settled naturalist (once aware of the relevant info regarding ECREs).

A justification for assumption 1 has already been presented (if briefly), in the form of the observation that innocent misinterpretation or unconscious confabulation / false memory are just not plausible options in the cases provided above (and many other like cases). With respect to honest misinterpretation, consider for example case #3: one could not honestly mistake seeing the spontaneous disappearance of serious lesions before one's eyes (especially if one is a medical doctor, as in this case). As to false memories, wholesale manufacture of complex false memories does not occur in individuals of normal mental health (outside of unique situations like hypnosis); in particular, wholesale manufacture of complex *existentially significant* false memories does not occur in individuals of normal mental health, and ECREs are of undeniable existential significance. Relatedly, physically sound non-drug users are not subject to spontaneous, detailed hallucinations. Along these lines consider case #2. It is simply not plausible to maintain that Morse, a paediatrician, unconsciously manufactured a false memory of this child's NDE report, especially given the magnitude of its importance in his later personal and professional life. And in at least some of the cases above, enough biographical facts are known of the reporters of the experience to ascertain that they are clearly not mentally ill nor suffering from relevant physical defects (brain trauma, etc.) or addictions.

Further empirical support for assumption 1 would require further case studies of testimony to ECREs. Having already taken up a good bit of space (perhaps excessive space) with case studies in this paper, I cannot expand on this here. But the reader can easily access further cases for him/herself by turning to some of the primary sources cited here, and many, many others not cited. Personal inquiry among trusted family and friends may also turn up cases. (It has done so for me, and I doubt that my own sphere of acquaintance is much of a statistical outlier in this respect.)

The idea behind assumption 2 is fairly commonsensical: it is unjust and uncharitable to attribute a bad deed or character trait to someone without adequate evidence that the deed was done or that the trait is really present. It amounts to slander. Moreover, to do so is a violation of the Golden Rule: no one likes being called a liar, especially on a matter of great importance. A settled naturalist can certainly sympathize with the situation that the subject of an ECRE would find herself in; we can all imagine what it would be like to have such an experience, and how we would feel if we were dismissed as liars upon mentioning it.

While the moral principle at play here should, I think, be generally acceptable,[4] the injunction to refrain from making harsh judgements of others, except where absolutely demanded by the facts, is particularly important within Christian ethics (for Biblical precedent see especially Matthew 7:1-5 and Luke 6:37). This injunction has of course been the subject of much commentary and discussion in the subsequent history of moral theology; Aquinas provides an influential treatment, writing that 'from the very fact that a man thinks ill of another without sufficient cause, he injures and despises him. Now no man ought to despise or in any way injure another man without urgent cause: and, consequently, unless we have evident indications of a person's wickedness, we ought to deem him good, by interpreting for the best whatever is doubtful about him'.[5]

The claim of premise 1 is that because most settled naturalists still view non-naturalism as neither absurd nor certainly false, they cannot properly take testimony to an ECRE as *ipso facto* adequate evidence of deception, even where deception is the only plausibly available naturalistic explanation. This premise is simply a more specific instance of the general principle that viewing a proposition as neither absurd nor certainly false is incompatible with being completely dismissive of testimony relevant to the truth of that proposition (whether via an allegation of deception or some other means).

---

[4] This is not to say that it is devoid of complications, or free from possible objections (some of which are treated briefly in section 3 below). See Oderberg (2013) for a recent explication and defence of the ethical perspective adopted here. (Actually, in certain respects Oderberg's stance is stricter than the one employed here, in that he thinks that being charitable can in some cases demand assuming someone's innocence even in the face of adequate (but not decisive) evidence. While such a stance is certainly defensible, I will not make use of that more robust thesis.)

[5] *Summa Theologiae*, IIa IIae, q. 60, art. 4, resp.

Premise 2 follows readily from the combination of assumption 2 and premise 1, while premise 3 / conclusion 1 is deduced from premises 1 and 2. Premise 5 / conclusion 2 is in its turn deduced from premises 3 and 4. And what is the justification for premise 4? In the absence of any social science literature that looks into the content of people's naturalism in depth (i.e., not just polling rates of belief in naturalism but also providing detailed data on the degree of confidence with which that belief is held, whether they see non-naturalism as a live option in the sense employed here, etc.) I have to rely on my own (admittedly unscientific) interactions with many thoughtful and self-reflective naturalists and engagement with much recent and past naturalistic literature. Based on that experience I believe premise 4 is true and would be supported by rigorous social science data, if such studies were to be undertaken. However, I grant that I could be wrong about this – certainly some of the so-called New Atheist literature promotes a version of naturalism in which non-naturalism is seen as utterly absurd. It may also be that while I am right about this at present, I may not be right about it years down the road, if the sort of naturalism pushed by the New Atheist becomes much more common. If therefore I am wrong about this, or become wrong about this, then the formulation of the argument will have to change, with the target class switching from 'most settled naturalists' to 'many settled naturalists'.

Regarding premises 6 and 7, the two options laid out seem exhaustive. Given that there are many testimonies to ECREs the only rational naturalistic explanation for which is deception, upon becoming aware of these the settled naturalist must either go that route and attribute deception, or decline to put forward a response. The former is option is immoral, while the latter is irrational. The latter would not be irrational for tentative naturalists, for whom suspension of judgement pending further investigation into individual ECREs seems a workable response. But it is not a response open to the settled naturalist, since a suspension of judgement regarding the reality of ECREs is inconsistent with maintaining naturalism as a firm, stably held view, implying as it does that non-naturalism warrants further investigation.

Let me emphasize again the restricted nature of the resulting conclusion. I am not claiming that all settled naturalists are doing something either immoral or irrational in holding to naturalism as a settled belief. The argument just provided applies only to settled naturalists who still see non-naturalism as neither absurd nor certainly false and who are aware of the commonality of ECREs – more specifically, the commonality of

ECREs the only rational naturalistic explanation of which is deception. At present therefore the argument may apply only to a small number of naturalists, which I freely admit. (Hopefully that number will increase after this article is widely read!) Moreover, the argument conclusion does not entail that the settled naturalist must abandon naturalism; rather, it entails that at the least the settled naturalist must shift to a tentative naturalism.

IV. OBJECTIONS

(4.1) Maybe it is immoral to dismiss as liars those claiming ECREs for which deception is the only rational naturalistic explanation. Nevertheless, it is rational to do so, given the prior balance of probability of naturalism over non-naturalism. That balance may not be so tilted as to render non-naturalism absurd or certainly false, but it renders it sufficiently improbable that it is perfectly rational, if immoral, to dismiss the relevant testimony as deceptive. We live in a messed up world where sometimes we're faced with a conflict between morality and rationality, and when that happens we should go with rationality.

This objection touches on a much larger philosophical debate concerning whether moral reasons always trump other sorts of reasons. I lack the space to review adequately the arguments in favour of considering moral reasons as overriding, and will not attempt to do so. However, I will say that I believe moral reasons are indeed trumps. In fact I would go a bit further and maintain that seeing other reasons as competitors is a problematic way of framing the issue – moral reasons constitute a different and incommensurable class of reasons, such that other sorts could not function as competition. Again, I cannot hope to make a case for this here; suffice it to say that a view of morality as overriding is widely advocated in the ethics literature.[6] In consequence, the present objection is a risky one for the settled naturalist.

As a counter-reply, one might re-conceive the present objection, such that the situation is not one in which one sort of reason is competing with

---

[6] See for instance Fairbanks (2012), Hare (1981), and Stroud (1998). Pojman (1991) makes an interesting case that while the moral overridingness thesis holds on theism, it does not hold on any kind of atheism. That would add an interesting complication in the present context; if correct, it would imply that my argument would be an effective critique of settled naturalism *from a theistic perspective*, but not from a naturalistic one. In that case it would retain some interest, but its intended audience would be badly curtailed.

morality, but rather a situation in which one moral duty (the duty to seek the truth) conflicts with another (the duty to follow the Golden Rule). Seen in that light, the situation is really that of a moral dilemma, such that the moral overridingness reply would not automatically address it.

By way of a counter to the counter: the demand to switch from settled to tentative naturalism in the face of ECREs does not conflict with the duty to seek the truth. Indeed, by prompting someone to reconsider seriously the evidence for the supernatural, it might be truth-conducive.[7] Since the settled naturalist who does not consider belief in the supernatural absurd or certainly false cannot rule out the possibility of that shift's being truth-conducive, it is not the case that the shift would violate the duty to seek the truth.

(4.2) Surely to judge harshly someone whom one has never met, indeed a person whose name may not even be known (as in the anonymous cases included among the case studies above), does that person no clear harm. That speaks against its immorality. Moreover, if such judgement is immoral it must be a very minor immorality – even if moral reasons do usually trump other sorts of reasons, surely the import of a minor immorality can be outweighed by the importance of preserving one's settled naturalism.

A person can be harmed without knowing she is harmed, and a person can be wronged without knowing she is wronged. Still, it is wrong to wrong people. It is wrong to judge your neighbour a jerk on inadequate grounds, even if you don't inform your neighbour of this judgement. And it is wrong to judge someone a liar prematurely even if that person is unaware of it.

And is it true that such a judgement would be merely a minor wrong whose import would be outweighed in this case? This gets into the tricky question of how to weigh wrongs. To the extent that people take their reputations to be important, slandering someone might be seen as a serious wrong. Doing so in public would be worse than doing it merely in thought, but the latter still wrongs the person.[8] As to whether that

---

[7] A related point will come up at the end of the reply to objection (4.4).

[8] Doesn't this amount to a kind of thought-police, inimical to liberal values? Oderberg (2013) takes this up in his general discussion of judging others, writing that 'the application of morality to states of mind is hardly novel. Even liberal-minded people disapprove morally of hatred, spite, jealousy, and other corrosive states of mind – and presumably not just because of their tendencies to outward manifestation. We can make sense of a society of hate-filled people who nevertheless managed to get along well due to certain firmly built-in codes of proper conduct. But would the neutralization of external

wrong is outweighed by the good of preserving one's settled naturalism, one might question whether that really is a good, on the grounds that naturalism might be false, and indeed might perhaps be demonstrably false depending on how one comes to evaluate testimony to ECREs. One might also question such a reaction on the more general grounds mentioned in the previous reply: if morality is a trump, presumably it is always a trump, even where the competing non-moral reasons seem particularly important. Trumps are only meaningful, after all, if they retain their force even when the temptation to ignore them is substantial.

(4.3) If we adopt the perspective advocated here then we will rapidly get an expansion problem. Do we have to maintain an open mind about flying saucers, Yetis, ghosts, etc. because there may be instances of testimony to these the only rational naturalistic explanation for which is deception? How far do we take the injunction not to judge others harshly?

Based on the criteria I've used above, if one finds the proposition 'ghosts are real' to be absurd or certainly false, one can dismiss all such testimony as deceptive and not wrong anyone in doing so. If on the other hand one thinks of that proposition as neither absurd nor certainly false, then to judge someone a liar simply because he/she has given testimony to a ghost sighting (the only rational naturalistic explanation for which is deception) would indeed be immoral.

(4.4) But if we follow the demands laid out here, important beliefs might be held hostage to liars. Why should one's degree of belief in naturalism (or any important belief) be left vulnerable in this way?

Aquinas, in the same article of the *Summa* quoted earlier, considers a similar objection: 'It would seem that doubts should not be interpreted for the best. Because we should judge from what happens for the most part. But it happens for the most part that evil is done, since "the number of fools is infinite" (Ecclesiastes 1:15), "for the imagination and thought of man's heart are prone to evil from his youth" (Genesis 8:21). Therefore doubts should be interpreted for the worst rather than for the best.'

Aquinas replies by reiterating the relevant moral injunction: 'He who interprets doubtful matters for the best, may happen to be deceived

---

manifestation equally neutralize the internal states themselves, morally speaking? .... In any case, whether you concur with this latter consideration or not, it remains that every rash judgment puts a dent or hole in someone else's reputation (given that a reputation just *is* the sum total of opinions everyone has about an individual), and if reputation is a highly valued good, that good is thereby, however slightly, under- mined.'

more often ... yet it is better to err frequently through thinking well of a wicked man, than to err less frequently through having an evil opinion of a good man, because in the latter case an injury is inflicted, but not in the former.' This fits in well with the notion that moral reasons override reasons of personal self-interest. It also accords with the Platonic idea that it is always better to suffer evil than to do evil, and indeed with a basic tenet of our own justice system: namely, that people are presumed innocent until proved guilty, on the assumption that it is better to risk letting some guilty men go free than to condemn the innocent. Such views are difficult and dangerous to maintain and put into practice. But then, if morality were easy everyone would do it ....

To soften the blow, recall that avoiding the habitual attribution of lying motives to others when they testify to ideas one disagrees with is surely not only virtuous but also beneficial. It is part of being open-minded, and may be conducive to learning important new truths one would otherwise remain ignorant of. The habit of refraining from harsh, inadequately supported judgements of others can thus receive further support (not that it needs it) from its being potentially truth-conducive.

(4.5) Must the settled naturalist really shift to tentative naturalism? Mightn't she instead rationally and morally adopt the following stance: I don't know how to explain what's going on with ECREs. These experiences do indeed seem compelling, and the witnesses sincere and reliable. It must be granted that at present naturalism has no workable explanation for them. But that needn't weaken the evidential status of naturalism, certainly not to the point where one is obligated (morally or otherwise) to reconsider seriously the reality of the supernatural. Every theory, no matter how plausible and well-supported, has to contend with anomalies, certain findings that are difficult to integrate into the theory. For now, ECREs are an example of such an anomaly. The naturalist can admit this while still rationally maintaining that one day the anomaly will be explained.

This objection underestimates the evidential significance of ECREs. Think again of the experiences related in our six sample case studies. These experiences, if they really occurred as reported, are not merely difficult to explain on the assumption of naturalism, or in tension with naturalism; rather, they are to all appearances *incompatible* with naturalism (keeping in mind of course the points made regarding the Duhem / Quine thesis in the Introduction). As such, mere faith in the possibility of a future workable naturalistic explanation is not a tenable strategy of reply for the

settled naturalist. Such faith may suffice to permit the settled naturalist rationally to remain a naturalist, but it will not suffice to avoid the need to shift from settled naturalism to tentative naturalism. This is particularly apparent given that ECREs are not isolated, terribly rare experiences, but rather are reported often enough that any researcher could easily assemble a much larger set of case studies than I have presented here, simply by dipping into the relevant existing literatures.

(4.6) Shouldn't this sort of argument give pause to many non-naturalists? How could one be a settled Christian (for instance) if one had to take seriously every claim to an ECRE with unambiguously Hindu content (a vision of Krishna perhaps), or Buddhist content, or some other religious tradition whose principal tenets are inconsistent with Christianity?

There are claims to religious experience within a variety of religious traditions, not all of which are compatible in all of their main teachings. But they are all agreed on the falsity of naturalism, and any ECRE, in any tradition (or none), tells against naturalism. As the relationship between ECREs and naturalism is the concern of the present paper, this point does not constitute an objection to the argument of section 3 above. Consequently, while the general issue of how to understand religious experiences across conflicting traditions is extremely important, and one that has rightly received sustained attention in the literature (from a variety of perspectives),[9] I will not take it up here. Certainly the Christian has resources within his/her tradition for explaining at least some experiences had by Hindus, Buddhists, etc., and I expect that the latter also have resources for accounting for experiences had by Christians. The question then becomes one of whose explanatory resources are most effective and plausible, a question that has to be resolved, in part, by reference to broader issues of dogma and morality, etc. The most relevant point for our purposes is simply that the Christian needn't dismiss as deceptive testimony to an ECRE with content appropriate to Hinduism, nor need the Hindu dismiss as deceptive testimony to an ECRE with content appropriate to Christianity.

However, one might press the present objection further by asking whether ECREs could be counted against naturalism if in fact they did *not*, collectively, point to some single, coherent alternative. If indeed they did not, mightn't one just see them as brute facts, fundamentally

---

[9] See for instance Alston (1991: 255-285), Heim (2000), Maritain (1944: 225-255), and Yandell (1993: 279-321).

meaningless indications of the world's underlying absurdity? Well, it must be granted that it would be a problem were there not indeed some single worldview at which ECREs collectively pointed. However, it would not necessarily be a problem if we could not (or not yet?) *discern* such a worldview *through* them. They might be coherent in the relevant respect, despite seeming diversity, and yet not in a way that was discernible to us. (As a matter of fact I think that ECREs, even in the midst of their diversity, do collectively point toward a single worldview, though to argue for this would be require a very different and much lengthier project.)

V. CONCLUSION

Various cases have been made against naturalism on the grounds of its moral implications. Some have argued that naturalism is inconsistent with moral realism, or at least is liable to undermine belief in moral realism. Others have argued that naturalism undermines moral accountability, given its attendant denial of the possibility of post-mortem judgement. The case made here is quite different. It assumes that naturalism is consistent with moral realism, and relies on what is taken to be a commonly shared moral intuition, namely that premature harsh judgement of others is morally problematic. Some past naturalists have been too quick to dismiss certain kinds of ECREs as the products of deception, without considering in detail the moral implications of such dismissal. The argument made here is designed to draw attention to those neglected implications, and to the extent it does so it ought to prompt serious reflection on the part of naturalists.

BIBLIOGRAPHY

Alston, W. 1991. *Perceiving God: The Epistemology of Religious Experience* (Ithaca: Cornell University Press)

Aquinas. 1920. *Summa Theologica*. Translated by the Fathers of the English Dominican Province. Available at: <www.newadvent.org> [accessed 09/08/2016]

Fairbanks, S. 2012. 'A Defense of the Overridingness Thesis', in: *Should We Always Act Morally? Essays on Overridingness*, ed. by Sebastian Schleidgen (Marburg: Tectum Verlag)

Franks Davis, C. 1989. *The Evidential Force of Religious Experience* (Oxford: Oxford University Press)

Hadley, L. 2013. *Touching Heaven: Real Stories of Children, Life, and Eternity* (Grand Rapids: Baker Books)

Hare, R. 1981. *Moral Thinking* (Oxford: Oxford University Press)

Heim, S. M. 2000. 'Saving the Particulars: Religious Experience and Religious Ends', *Religious Studies*, 36: 435-453

Hume, D. 1975/1748. *An Enquiry Concerning Human Understanding*, in: *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. by L. A. Selby-Bigge, (Oxford: Oxford University Press)

Maritain, J. 1944. *Redeeming the Time* (London: Geoffrey Bles)

Maxwell, M., and Tschudin, V. (eds.) 1990. *Seeing the Invisible: Modern Religious and Other Transcendent Experiences* (London: Penguin)

Morse, M., and Perry, P. 1990. *Closer to the Light: Learning from the Near-Death Experiences of Children* (New York: Ballantine Books)

Oderberg, D. 2013. 'The Morality of Reputation and the Judgment of Others', *Journal of Practical Ethics*, 1: 3-33

Pojman, L. 1991. 'Is Contemporary Moral Theory Founded on a Misunderstanding?', *Journal of Social Philosophy*, 22: 49-59

Stroud, S. 1998. 'Moral Overridingness and Moral Theory', *Pacific Philosophical Quarterly*, 79: 170-189

van Lommel, P., van Wees, R., Meyers, V., and Elfferich, I. 2001. 'Near-Death Experience in Survivors of Cardiac Arrest: A Prospective Study in the Netherlands', *The Lancet*, 358: 2039-2045

White, J. 1988. *When the Spirit Comes with Power: Signs and Wonders Among God's People* (Downers Grove: Intervarsity Press)

Yandell, K. 1993. *The Epistemology of Religious Experience* (Cambridge: Cambridge University Press)

GOD'S BEING IS IN BECOMING:
AN ESSAY IN THEOLOGICAL IDEALISM

HARTMUT VON SASS

University of Zurich

Abstract. God's being is becoming – the title is the thesis. The first section of this paper will be dedicated to the problem of radical historicity in sketching three dogmatic approaches dealing with the relation between God and history. After critically introducing the concept of relational – in contrast to intrinsic – properties in the second section I will apply a revised version of this concept theologically in integrating it into the architecture of Trinitarian thinking. Accordingly, and on that basis, the last section can address the ambivalent as well as precarious question in which sense God's ultimate being is in real (be) coming.

Theologians nowadays are 'idiots concerned with salvation'. That's at least what the German philosopher Peter Sloterdijk holds – leaving the closely related question, unfortunately, open, as to what kind of 'idiots' philosophers currently are.[1] Theological 'idiots' show their true character not only by being concerned with salvation, but rather in speaking of God, i.e. not only by being focused on the effects of religion, but mainly in being engaged in thinking about the 'happiest and most lovely substance', as Leibniz once called God as the ultimate.[2]

Now, it might turn out to be an unhappy starting point to distinguish sharply between God and what God does, between the divine reality and its bearing on its true believers. It is one of the most relevant issues

---

[1] Peter Sloterdijk, *Zeilen und Tage. Notizen 2008–2011* (Frankfurt a.M.: Suhrkamp, 2012), p. 103; trans. mine.

[2] Gottfried Wilhelm Leibniz, 'Vernunftprinzipien der Natur und der Gnade', in idem, *Vernunftprinzipien der Natur und der Gnade. Monadologie*, ed. by Herbert Herring (Hamburg: Meiner, 1969), pp. 2–25 (p. 23).

in Christian dogmatics and its philosophical company to settle, clarify and elaborate on that very topic. One classical label under which this topic is dealt with is the radical historicity that faith in the Christian tradition endorses, embraces, or presupposes. The first section will very briefly be dedicated to this problem in sketching three approaches dealing prominently with God and history (1). This makes the subsequent question unavoidable whether this historicity also concerns and touches on God Himself. After and by introducing and discussing critically the concept of relational – in contrast to intrinsic – properties (2) I will apply a revised version of this concept theologically in integrating it into the architecture of Trinitarian thought and, accordingly, in meeting the complicated as well as ambivalent question in which sense God's being is in (be)coming (3). However, before I begin, I have to confess, that the emphasis in the title is neither put on 'idealism' nor on 'theological', but on 'essay', which turns this paper into a tentative outline to a dogmatic problem whose further ramifications exceed the capacities of the modest pages that are following now.

I. RADICAL HISTORICITY

According to the Christian understanding God does not only and necessarily relate to history and His creatures and creation, but is Himself – in the twofold sense of almost every genitive – a *God of history*.[3] Hence, historicity is, then, not only the human condition beyond which God is leading His solitary 'existence', but God too is essentially subjected to that very change. Whereas 'historicity' means that all earthly things, facts, and beings have a tradition – coming from somewhere, going anywhere ('anything goes') – 'radical historicity' amplifies this 'traditional' approach by including literally everything, also God, into this constant open-endedness.

There are different and at least three prominent projects attempting to come to terms with this precarious scenario that, obviously, lead us into a realm *after* (or *post*) classical metaphysics.[4] The first of these endeavours

---

[3] As *locus classicus* see Wolfhart Pannenberg, 'Dogmatische Thesen zur Lehre von der Offenbarung', in *Offenbarung als Geschichte*. In Verbindung mit R. Rendtorff, U. Wilckens, T. Rendtorff hrsg. von Wolfhart Pannenberg (Beiheft 1 zu *Kerygma und Dogma*) (Göttingen: Vandenhoeck & Ruprecht, 1961), pp. 91–114.

[4] Insofar as 'classical metaphysics' is taken to be an attempt to identify an invariant structure beyond temporal and cultural change; cf. Hartmut von Sass and Eric Hall,

is the Hegelian framework whose emphasis lays on a processing dynamics of counterbalance and its hopefully higher resolution.[5] This kind of dynamics is a *structural* one based on either real or substantial respectively virtual or conceptual items mediating theologically the thinking of God into God's self-thinking as the divine essence.[6]

The second project carries its ambition in its name: process thinking in the wake of Alfred North Whitehead. Here, God is identified with His history of self-completion or self-perfection at the price that God is no longer temporally before His creation. It is, then, only consistent, as Whitehead actually does, to infer from this premise that it is equally true that God created everything as well as holding that God was created by the world.[7] This dynamic is a *cosmological* one integrating God into the process of everything in a revitalization of a Spinozian pan*en*theism.[8]

The third theological version of thinking historicity is the *hermeneutical* one. It deserves its name not only because of promoting 'understanding' to the central feature of faith (faith as understanding act of believing) and not by considering itself as a hermeneutical enterprise (theology as hermeneutics), but rather by understanding God Himself as an event of understanding. God is, then, to be thought of as hermeneutical reality (God as the reality of the new understanding of or as faith).[9]

'Groundless Gods. Metaphysics, its Critique, and Post-Metaphysical Theology. An Introductory Essay', in idem (eds.), *Groundless Gods. The Theological Prospects of Post-Metaphysical Thought* (Eugene OR: Wipf and Stock Publishers / Pickwick, 2014), pp. 1–37, 13–17; see also Friedrich Hermanni, *Metaphysik. Versuche über letzte Fragen* (Tübingen: Mohr Siebeck, 2011), p. 1.

[5] Cf. Jörg Dierken, *Fortschritte in der Geschichte der Religion? Aneignung einer Denkfigur der Aufklärung* (Leipzig: Evangelische Verlagsanstalt, 2012), pp. 72–81.

[6] See Ludwig Feuerbach, who thinks that Hegel considers God's essence as God's acting and were, therefore, 'genötigt, das Gedachtwerden Gottes zum Sich-selbst-Denken Gottes zu machen' (Ludwig Feuerbach, *Das Wesen des Christentums*. Ausgabe in zwei Bänden, ed. by Werner Schuffenhauer (Berlin: Dietz, 1956), p. 348).

[7] Cf. Alfred N. Whitehead, *Process and Reality*. Corrected edition, ed. by David Ray Griffin and Donald W. Sherburne (New York and London: The Free Press, 1978), p. 348; see also Daniel A. Dombrowski, *A Platonic Philosophy of Religion. A Process Perspective* (Albany NY: SUNY, 2005), ch. 3 and 4.

[8] See Philip Clayton and Paul Davies, *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion* (Oxford: Oxford University Press, 2006).

[9] For details see Hartmut von Sass, 'Faith and Being. Hermeneutical Theology as Post-Metaphysical Enterprise', in idem and Eric E. Hall (eds.), *Groundless Gods. The Theological Prospects of Post-Metaphysical Thought* (Eugene OR: Wipf and Stock Publishers / Pickwick, 2014), pp. 214–241.

Evidently, all three versions of thinking historicity in theological terms suffer from, at least, the danger not to do justice to God's sovereignty and independence of the world. In the first version God comes to Himself by and only by the process of thinking and being thought. In the second version the development of the world and God's 'biography' are to be identified. The third and hermeneutical option, finally, has the inclination to turn God into a function of human understanding.

On the one hand there are good reasons for an incarnational religion to invest theologically in clarifying the historicity of faith *and* God not as two, but *one* single question;[10] on the other hand it is equally understandable to implement concepts to prevent God from collapsing with his creation.[11] There has to be genuine reservation against that danger of equation traditionally expressed in the idea of a hidden God, the distinction between immanent and economic Trinity, the God that is beyond our grasp and knowledge. All these attempts – the *Deus absconditus*, the Trinitarian immanence, or negative and apophatic theology – try to save the God of history from being a merely historical God. Isn't it true that, to put it in Schleiermacherian terms, the idea of God's being *a se* (aseity) belongs essentially to the religious consciousness of sin and mercy, of fallenness and divine redemption?[12]

All this leads to the fundamental question of how to find (or even to detect) the appropriate balance between God's own historicity and His independence, more specifically: is His 'story' only understandable as derived from His being *a* and *per se* (as, roughly speaking, Calvinism teaches) or is it the other way round, namely, that His independence is an integral element of his being *pro nobis* (as the Lutheran tradition seems to be defending)? Turning now to the concept of 'relational' properties amounts to the attempt to make sense of this second option.

---

[10] Cf. John A. T. Robinson, *Honest to God* (London: SCM Press, 1963), pp. 49 and 61.

[11] See Wolfhart Pannenberg, 'Der Gott der Geschichte. Der trinitarische Gott und die Wahrheit der Geschichte', in idem, *Grundfragen systematischer Theologie. Gesammelte Aufsätze 2* (Göttingen: Vandenhoeck & Ruprecht, 1980), pp. 112–128 (p. 119).

[12] See Notger Slenczka, 'Das Dogma als Ausdruck des religiösen Selbstverhältnisses. Trinitätslehre bei Schleiermacher, Troeltsch und Tillich', in Ulrich Barth, Christian Danz, Wilhelm Gräb, Friedrich Wilhelm Graf (eds.), *Aufgeklärte Religion und ihre Probleme. Schleiermacher, Troeltsch Tillich* (Berlin / Boston: De Gruyter, 2013), pp. 661–684, esp. 682–684.

II. RELATIONAL PROPERTIES

There are several ways of thinking through changing objects. Either the object in question changes due to an internal development, or the person conceiving this object changes, or the link between both is processing in a new direction. Under Kantian (or post-Kantian) conditions, the first way is a mere abstraction since change is itself a category that is introduced by an observer for whom change is conceivable or, as stronger version, without whom there is no change at all. Hence, the object leads back to its observer and his or her relation to that very object amounting to the second and third way. Thus, change is only change for someone in a particular context. Nevertheless, we would like to speak of change that is a feature of the object, in other words, an objective change.

There are again two options to clarify this objective change – and this alternative brings us to the distinction I am interested in here, the difference between intrinsic and relational (or extrinsic) properties. It has been developed within the discussion of defending the hermeneutical thesis by Hans-Georg Gadamer on the essential underdeterminedness of every interpretation, but could also be used outside that important and influential debate.[13]

Intrinsic properties are taken to be those features that an object has in the way that this thing is 'in itself', such as shape, size, colour (if one wants to count secondary properties to that object 'itself').[14] Relational properties, however, are those features that an object has and that depend (wholly or partly?) on something other than that very thing, such as being an uncle of someone or being married to or divorced from Queen Mary.[15] So, the two ways mentioned above consist in either the change of the object in question (intrinsic) or the relation between this object to other objects (relational).

---

[13] Cf. esp. David Weberman, 'A New Defense of Gadamer's Hermeneutics', in *Philosophy and Phenomenological Research* 60:1 (2000), 45–65; Weberman, however, relies here on David Lewis' 'Extrinsic Properties', in *Philosophical Studies* 44:2 (1983), 197–200; see also Jeff Malpas, 'The Origin of Understanding: Event, Place, Truth', in *Consequences of Hermeneutics. Fifty Years After Gadamer's Truth and Method*, edited by Jeff Malpas and Santiago Zabala (Evanston Il: Northwestern University Press, 2010), pp. 261–280.

[14] Cf. Andy Egan, 'Second-order predication and the Metaphysics of Properties', in *Australasian Journal of Philosophy* 82:1 (2004), 48–66.

[15] See also Peter F. Strawson, *Individuals. An Essay in Descriptive Metaphysics* (London: Methuen, 1959), ch. 5.

Consider, for instance, the French Revolution. Our understanding of that historical event today is (or might be) crucially different compared to the past instances of understanding and contextualizing it. The importance of temporal distance consists mainly in the way in which more recent events have brought out new aspects of or 'redetermined' the earlier phenomena. In the case of the French Revolution, there is the occurrence of Napoleon Bonaparte, the ramified creation of the French Republic, the tensions up to the war with Germany, and, of course, numerous more subtle and less prominent occasions.

However, the temporal element is only one ingredient that constitutes the relational shifting. There are also relational properties that derive from the distance between the object of understanding and the vantage point of the interpreter distanced not by its temporal, but by its cultural specificity. Therefore, we have the *diachronic* as well as the *synchronic* version of relational properties and, of course, a mixture of both. What this precisely means is that the object of understanding is underdetermined, because its relational properties are shifting according or even due to the temporal or cultural point of the observing person. If it is correct that *Hamlet* (I was sitting in Southern Denmark while writing this) is a different object for us here and now than for a 17th-century reader in England because of its delayed relational properties, it might be no less true that the same play is a different object for different readers today because of the different relational features Shakespeare's masterpiece has as a consequence of its relation to divergent cultural points of view. *In summa*: The object in question is not complete (or, with Kant, 'in itself') but rather in a state of constantly being formed, shifted, adjusted – changed, because of its relational properties, both temporal and cultural.

So far my description of what has been introduced as 'relational properties' by a few authors like David Weberman. It is, to begin with, clear that relational properties do not follow the logic of dispositions that unfold themselves as time marches on, but that they represent something truly new due to sometimes unexpected links to other entities. There are no defined and definable limits for these connections, rather a potentially *holistic network* of clusters, connections, and alliances that is synthetic (and not analytic) in character. Moreover, the concept of relational properties belongs to the hermeneutic tradition of reception theories turning initially aesthetic claims into ontological ones. Rudolf Bultmann's famous claim according to which the reception belonged to

the received text is only the theological specimen of that very doctrine starting off with Gadamer's main work *Truth and Method* from 1960 and being an 'import hit' for literary theories thereafter.[16]

However, different objections may arise at this point and discussing them will lead us to make some conceptual modifications. First, does not this conception allow for a helpless relativism? No, I think, it doesn't, since we have to distinguish between being relativistic and relational. Whereas the former might express something like the absence and lack of any criteria, the latter underlines only a kind a dependence on specific contexts that do have criteria and set limits.[17] Second, it is not entirely clear in Weberman's paper on which ontological classes these properties are applicable. Sometimes it looks as if we are only talking about properties of things; at other times, however (as with the example of the French revolution), we are dealing with historical events; more to the point, one could argue that relational properties are also relevant for other properties of any kind. If we do not subscribe to the (debatable) claim, that the universe is ontologically furnished only by things, events, and properties, then we might include our minds as changing in relation to their new relational properties too. So, everything is subjected to that change and everything could be changed in the way described.

And last and most importantly, is not the whole scenario sketched so far based on a confusion of hermeneutical changes with ontological ones, of shifting *significance* on the one hand with stable *meaning* on the other, of features of *our* understanding an object with features of that object *itself*? Not necessarily! Consider, for example, that a person might describe the French revolution differently because she has undergone a political conversion; then, this descriptive shift is a result of an upheaval in the personal epistemic or attitudinal makeup – and not in the event 'itself'. But as soon as a person describes the revolution in a new way because the revolution has adopted new relations to other

---

[16] See Rudolf Bultmann, 'Das Problem der Hermeneutik', in idem, *Glauben und Verstehen. Gesammelte Aufsätze II*, 4th ed. (Tübingen: Mohr Siebeck, 1965), pp. 211–235 (p. 229); for the comparatistic background see Norman N. Holland, *The Dynamics of Literary Response* (New York / Oxford: Oxford University Press, 1968).

[17] Cf. Hilary Putnam, *Renewing Philosophy* (Cambridge MA / London: Cambridge University Press, 1992), ch. 4; Dewi Z. Phillips, *Faith After Foundationalism. Plantinga-Rorty-Lindbeck-Berger – Critiques and Alternatives*, 2nd ed. (Boulder CO / San Francisco / Oxford: Westpoint Press (1988) 1995), esp. ch. 4 and 17; in fact, this relative, but non-relativistic position is a Wittgensteinian one.

and new events, then it is not the person that went through a change, but the event did so, Weberman holds. Relational properties may be tantamount to new descriptions, but they are not merely changes in the epistemic or hermeneutical setting or the attitude and orientation of a person. It follows from this consideration that relational properties possess *ontological* weight.[18]

However, it might be the case that Weberman exaggerates his case and, moreover, that he claims more than what is required to make successfully his point. The confrontation between in- and extrinsic properties is, for sure, a crucial one; but it seems not to be necessary to hold that relational properties are *entirely* extrinsic to dispel the critical impression above, namely, that we are just playing here with the confusion between hermeneutical and ontological categories, with significance and meaning. Instead, one could argue in favour of the 'softer' supposition that relational properties are neither wholly intrinsic nor wholly extrinsic, but that they *combine* elements of both in constituting, therefore, a third category, a 'mixed bag'.

Coming back to the example of the French revolution, it was true that this historical event itself changed by new relational properties, but that this was not completely independent of what we as observer take that very event to be for us nowadays. It is analogous to what Kant described as 'perspective' on an object: I see the table as I am seeing it under certain parameters; there is, so far, nothing subjective (emotions, etc.) involved, just the unavoidable and hardly regrettable standpoint of a particular observer; however, there is, obviously, no perspective without the observing person, and this perspective could be shared by others if they see that table under similar conditions.[19] Hence, here we have the mixture between in- and extrinsic aspects as well; and nevertheless, we could defend the claim that is the core element of Weberman's considerations, namely that relational properties have indeed ontological bearings.

III. GOD'S BEING IS IN (BE)COMING

This final section has only one task, to combine the initial exposition of the problem of God *a se* and God *pro nobis* with the concept of relational

---

[18] Cf. Weberman, 'A New Defense of Gadamer's Hermeneutics', p. 55.

[19] Cf. Markus Gabriel, *Warum es die Welt nicht gibt*, 8th ed. (Berlin: Uhlstein, 2013), p. 114 (engl. trans. in preparation with Polity Press, Cambridge).

(or extrinsic) properties. In other words: how could that very concept help us to understand a bit better God as a changing reality without trapping Him in the cage of human minds?

We should remember that this topic leads back to the medieval debates on God's existence in relation to God's essence. According to the classical doctrine to be found in the Thomist and, later, the nominalist tradition, there is 'no metaphysical distinction' between the divine existence and essence.[20] Thus, for God it is impossible to bring about His own nature without having already all essential properties. What is, following the existentialists, characteristic for human beings, namely to exist and, then, to develop by creating one's own essence is not the way in which God relates to His creation.

The constructive (but semi-heretical) conclusion from identifying God's existence and God's essence consists, however, in equating God with His acting on us, since His essence is nothing but His pure act or God Himself as *actus purus*. Insofar, Ludwig Feuerbach is completely correct in saying that God is not *per se* if He is not at the same time there *pro me*; or to put it in even stronger terms: God does not exist if no one believes in Him; He is only insofar as He is the object of faith.[21] And this relation is expressed in the institution of prayer and confession where God and men come together in men being surrounded by the 'ultimately concerning' reality called God.[22]

But again, how to make sense of God essentially acting on us without being essentially depending on us? Or, to borrow Eberhard Jüngel's phrase: who is the God whose Being is in (be)coming? It is interesting to see, that Jüngel (paraphrasing Barth in a critical discussion with his German colleague Helmut Gollwitzer in the 1960s) deals with a similar problem as we are doing here. Whereas Gollwitzer emphasizes the God *per se*, the divine aseity,[23] Jüngel underlines the theological necessity to think even God – *pace* the ahistorical metaphysical tradition excluding (radical) historicity in favour of temporally invariant structures – in historical

---

[20] See Brian Leftow, 'Is God an Abstract Object?', in *Noûs* 24:4 (1990), 581–598, esp. 594.

[21] Cf. Ludwig Feuerbach, *Das Wesen des Christentums*, p. 312.

[22] See Eleonore Stump, 'Petitionary Prayer', in *Philosophy of Religion: The Big Questions*, ed. by Eleonore Stump and Michael J. Murray (Oxford / Malden MA: Blackwell Publishers, 1999), pp. 353–366, esp. 363–364.

[23] See Helmut Gollwitzer, *Die Existenz Gottes im Bekenntnis des Glaubens*, 2nd ed. (München: Chr. Kaiser, 1963), p. 175.

terms of change, affection, and relational conceptions.[24] God's acting is His essence, and God's being is a moved and moving Being, a Being in action.[25] Accordingly, God's being is in becoming.[26]

To conceptualize the divine coming and becoming it is, Jüngel holds in line with many other voices, unavoidable to think of God in Trinitarian terms, since the main intention behind that doctrine is to elaborate on God's radical historicity, i.e. even His self-confinement is to be thought historically and in a relational mode.[27] This could also be expressed, coming back to the concept of relational properties, as meaning that God is not 'in Himself', but that there are new relations to other and new events that change not only what we take to be God, but touch God's own Being.

Traditionally, the doctrine of Trinity serves as a hermeneutical model to think God's reality between the two poles outlined at the beginning, His sovereign aseity and His essential relation to the world. However, what exactly does this imply, how could we make sense of this doctrinal topic? My assumption is now, that we could paraphrase Trinitarian thinking by referring to the institution of relational properties, in other words and more precisely: we might use the concept of relational properties as ontologically transforming features for elaborating on the Trinitarian dynamics between the three 'persons' of Father, Son, and the Holy Spirit as well as their relation to human beings.

This implies that the Trinitarian modes of Being (Barth's 'Seinsweisen') represent three ways of gaining new relational properties and that God is the very dynamics of this divine enrichment. Either the doctrine of Trinity is, then, the theological expression of the general claim of change based on gaining (and losing or substituting) relational properties; or

---

[24] See Eberhard Jüngel, *Gott als Geheimnis der Welt. Zur Begründung der Theologie des Gekreuzigten im Streit zwischen Theismus und Atheismus*, 3th ed. (Tübingen: Mohr Siebeck, 1978), pp. 6–7, hereafter GGW.

[25] Ibid., pp. 46 and 103; see Karl Barth, KD I/1, 391; II/1, 288.

[26] See Eberhard Jüngel, *Gottes Sein ist im Werden. Verantwortliche Rede vom Sein Gottes bei Karl Barth. Eine Paraphrase*, 2nd ed. (Tübingen: Mohr Siebeck, 1967), p. 77; also GGW, pp. 213, 415.

[27] Cf. GGW, p. 472; see also Ingolf U. Dalferth, 'The Eschatological Roots of the Doctrine of the Trinity', in *Trinitarian Theology Today. Essays on Divine Being and Act*, ed. by Christoph Schwöbel (Edinburgh: T & T Clarke, 1995), pp. 147–170, esp. p. 147; Hartmut von Sass, 'Nachmetaphysische Dreifaltigkeit. Barth, Jüngel und die Transformation der Trinitätslehre', in *Zeitschrift für Theologie und Kirche*, 111:3 (2014), 307–331.

it is the other way round, that relational properties constitute an essential element within that doctrine that could not explicated without referring to the category of relation. In the one case we would be dealing with a new paraphrase of an old doctrine; in the other case we would just be confronted with unfolding of what is already entailed in the Trinitarian dogma.[28]

Now, there are different regards for integrating relational properties into the Trinitarian dynamics. The first one is expressed in the question of whether we are dealing with God's internal relation(s) between Father, Son and Spirit or whether His relation to His creature is concerned (or whether we want to defend that distinction in the first place).

The second regard touches on the problem in how far relational properties are relevant for all three Trinitarian 'persons' or whether only the second 'Seinsweise' of God's reality possesses relational properties whereas the both other 'persons' embody other properties of God, namely intrinsic (Father) and extrinsic features (Spirit). The former alternative means that God the Father denotes the divine reality to which one can relate (intrinsically independent); God the Son is the mode in which God relates to Himself in love and ultimate loneliness, in complete identification and desperate alienation on the cross (relational); and God the Holy Spirit is the medium in which we as humans relate to God as the relation between father and son (extrinsically dependent). The latter alternative would include that all three Trinitarian 'persons' possess *all three* properties (intrinsic, relational, extrinsic) emphasizing therefore, *pace* Arianism, the classical *homoousios* determined at the council of Nicea in 325.

The third regard creating divergent readings of relational properties within a Trinitarian framework concerns the two versions of what it 'is' that is gaining the new properties by collecting new relations. It could mean that God 'before everything else has come into existence' would adopt these new relations; contrary to that conservative reading, the proposal could also mean that this concept of God as the ultimate beginning is itself excluded by relational properties, since He Himself is and was changed by these very features, meaning His transforming reality is gaining these new properties. In the former case God would

---

[28] Cf. Wolfhart Pannenberg, 'Subjektivität Gottes und die Trinitätslehre. Ein Beitrag zur Beziehung zwischen Karl Barth und der Philosophie Hegels', in idem, *Grundfragen systematischer Theologie. Gesammelte Aufsätze 2* (Göttingen: Vandenhoeck & Rupprecht, 1980), pp. 96–111, esp. 100 and 107.

serve as metaphysical back-up existing at the very beginning initially without having relational properties; in the latter case that idea would present already a confusion because God himself is the process of having new relational properties that essentially change what God's reality actually *is*.

(NB: this third regard, has obviously, important implications for Christology: the former version still allows for sticking to the idea of a theological relevance of the historical Jesus in contrast to his later reception starting with His own disciples; the latter version, however, undermines that classic distinction between the Jesus of history from the very beginning and the 'kerygmatic' Christ of confession, sermon, and actual faith, since the reception of Jesus would – according to the basic idea of relational properties – change what Jesus Christ *is*.[29])

Since all these relations may imply new connections to other facts, things, persons, and events even God does not remain untouched from these new relational properties that do not create God himself or turn him to be essentially dependent on something other than Himself, but essentially change Him by creatures who relate to Him in prayer and are changed by God in that very address. The doctrine of Trinity is a theological model to elaborate on that very idea of radical historicity; the institution of relational properties might help us to clarify the God who is thought of in Trinitarian terms and is, therefore, a God whose Being is in becoming.

IV. NEAR THE END: A THEOLOGICAL IDEALISM?

God's Being is in becoming – this title does not only mean that God incarnates into the unstable and risky realm of change and becoming. It means, moreover, that God Himself is in that process without holding back an unchangeable backup that would only duplicate or separate God into two entities. A theological idealism is *theological* because it is God and not only a 'transforming religion' that is the object of interest. And it is a theological *idealism* because it is the transformative and in this transformation also transformed God who is the ground of the Christian

---

[29] See Martin Kähler's famous paper on the distinction between the Jesus of history and the Christ of proclamation and preaching: *Der sogenannte historische Jesus und der geschichtliche, biblische Christus* [1892]. Neu herausgegeben von Ernst Wolf, 2nd ed. (München: Chr. Kaiser, 1956).

hope – including the expectation that this idealism might be the real or even 'higher realism'[30] in philosophical theology.[31]

---

[30] Friedrich D. E. Schleiermacher speaks of a 'höheren Realismus' in *Über die Religion. Reden an die Gebildeten unter ihren Verächtern* [1799], ed. by Hans-Joachim Rothert (Hamburg: Meiner, 1958), p. 31; for the relevant background see, for instance, Dietrich Korsch, 'Höherer Realismus: Schleiermachers Erkenntnistheorie der Religion in der Zweiten Rede', in *200 Jahre Reden über Religion. Akten des 1. Internationalen Kongresses der Schleiermacher-Gesellschaft*. Halle 14.–17. März 1999, ed. by Ulrich Barth and Claus-Dieter Osthövener (Berlin / New York: De Gruyter, 2000), pp. 609–628, esp. 625–626.

[31] This article is the revised version of a paper given at the 20th conference of the European Society for Philosophy of Religion: 'Transforming Religion' in Münster, Germany, in August 2014.

COULD GOD FAIL TO EXIST?

JOSHUA RASMUSSEN

Azusa Pacific University

Abstract. I apply developments in modal reasoning to the question of whether God has necessary existence. My larger task is to assess the main reasons to think that God is *not* a metaphysically necessary being. I consider Hume's conceivability-based argument, and then I pay attention to more recent arguments, including Swinburne's neo-Humean argument and the subtraction argument. I show that such arguments face a 'parity' problem, since the very reasoning that gets them off the ground also launches parallel arguments for an opposite conclusion. In my closing section, I sketch an argument schema designed to illustrate a new, general strategy for deducing the necessary existence of God by building upon recent modal cosmological arguments.

I. INTRODUCTION

How stable is God's existence? Many theists have thought that God's existence is as stable as possible: God is not the kind of being that *could* fail to exist. Are they right?

Here is a quick reason to think so. God is supposed to be the greatest conceivable being, or something near enough. The greatest conceivable being enjoys the strongest conceivable grip on existence: after all, a being is intuitively greater if it cannot fail to exist than it if can. So: God enjoys the strongest conceivable grip on existence. Therefore, God enjoys necessary existence – assuming God exists at all.

But there is a problem. Imagine a world that is empty of all beings. Can't you do that? If so, then you can conceive of a world in which God does not exist. From here it follows that even if God *does* exist,

his non-existence is conceivable. But then God doesn't have the strongest conceivable grip on existence. His non-existence is not inconceivable in the way that it is inconceivable for (say) a sphere to be wholly inside itself. So, we face a problem if we suppose that God cannot, in the strongest sense, fail to exist.

We have a puzzle here. If God *can* fail to exist, then it is difficult to think of God as *maximally* great. Wouldn't a necessary God be greater? If, on the other hand, God *cannot* fail to exist, then we've got to somehow shake the sense that an empty world is perfectly possible. Either way is puzzling.

How one solves the puzzle affects one's metaphysics and philosophical theology in many ways. For example, if God has necessary existence, then one could theorize that necessary abstracta are constituents of God's necessarily instantiated nature. By contrast, if God does not have necessary existence, then abstract objects – such as, properties, propositions, numbers – either have a stronger grip on existence than does God, or they are somehow contingent beings, if they exist at all. Also, if God has necessary existence, then the fact that God exists may explain why there is something rather than nothing, since the option that there is nothing would then be impossible. Furthermore, the great-making status of necessary existence hangs in the balance.

I will show how developments in modal logic and explanation-based reasoning lay a foundation for a fresh defence of the position that God has necessary existence. My primary task is to show that the toughest objections to God's necessary existence, including the conceivability-based objection, Swinburne's related semantic objection, and the recent subtraction argument, fall prey to a parity problem. In the final section, I'll outline a new argument strategy for establishing the necessary existence of God.

II. DEFENCE

Let us begin by getting clear on what it means to say that God has *necessary existence*. I will follow both Plantinga (who thinks God has necessary existence) and Swinburne (who thinks God does not have necessary existence) by treating 'metaphysical necessity' as expressing the strongest form of necessity. I am interested in the strongest form of necessity since it is the form of necessity that God's existence has been

classically thought to exemplify. So, if '□' abbreviates 'it is metaphysically necessary that', then we may state our hypothesis as follows:

(N) God exists → □ God exists.

Note first that the hypothesis here is ontological, not epistemological: it is about what must *be*, not what a rational person must *think*.

We may further clarify the meaning of '□' by distinguishing it from strict logical necessity. I consider it an open question at the outset whether all metaphysical necessities are deducible from the cannons of logic (together with appropriate definitions of the terms involved). We will return to this question when we discuss Swinburne's semantic objection. For now, what matters is that we understand metaphysical necessity as the strongest form of necessity. We might express the notion this way: *p* is metaphysically necessary if and only if *p* is true at every world at which the basic rules of logic hold. Or more succinctly: *p* is metaphysically necessary if and only if *p* is true wherever there are no true contradictions.[1] In either case, I leave open whether there are metaphysically necessary truths that cannot be *shown* to be true via strict logical deduction.

I will assume that the axioms of S5 modal logic characterize metaphysical necessity.[2] In particular, I will assume that necessities are necessarily necessary and that possibilities are necessarily possible. These assumptions are in keeping with treating metaphysical necessity as the strongest sort of necessity: modal notions are so strong that they *cannot* vary. Moreover, the thesis that God has necessary existence in a sense that is consistent with S5 logic is interesting in its own right.

---

[1] Notice here that I have not reduced metaphysical necessity to strict (or narrow) logical necessity. To say that metaphysically necessary truths hold in the same worlds where strictly logically necessary truths hold is not to say that all metaphysically necessary truths *are* strictly logically necessary. Nor is it to say they are not.

[2] Specifically, I assume the following axioms (where '◊p' abbreviates 'it is not necessary that not p'):

M: □p → p
K: □(p → q) → (□p → □q)
4: □p → □□p
5: ◊p → □◊p

I leave out **N** (the necessitation rule) because **N**, together with standard non-free logic, implies that the theorem, ∃x (x = x), is necessary, and thus that there must be something. One may wish to avoid building into the *meaning* of 'necessity' anything that strictly implies the controversial metaphysical hypothesis that there must be something.

We are now ready to consider objections to (N). The main objections throughout history fall into three categories: conceptual, semantic, and ontological. Each objection is designed to show that there could be a world without any (concrete) beings. I will assess a representative instance of each category.

The Problem of Conceivability

David Hume famously writes, 'Whatever we conceive as existent, we can also conceive as non-existent' (1854: IX, 189). He goes on to argue that our notion of *necessity* cannot apply to God, since we can conceive of God as non-existent. He concludes: God is not a necessary being.

Let us slow down the objection. Hume suggests that any conceivable thing can be conceived to not exist. Now he doesn't really need to assume that it is possible to conceive of the nature of a necessary being. Perhaps such a nature is actually inconceivable. What's crucial for his argument is that we can conceive of worlds that are empty of all beings. Or, for those who think that it is impossible to coherently conceive of a world that is empty of *abstract* entities, such as numbers, properties, or worlds themselves, we may focus on *concrete* – causally capable – things. We can conceive of a world without any concrete things, it seems. So I think we can put Hume's argument in its best light as follows:

A1: A world empty of concrete things is conceivable.

A2: If a world empty of concrete things is conceivable, then such a world is possible.

A3: Therefore, a world empty of concrete things is possible.

A4: If a world empty of concrete things is possible, then there is no necessary concrete thing.

A5: Therefore, there is no necessary concrete thing.

There are reasons to like each premise. The first premise is plausible on reflection: just imagine a world containing nothing but empty space. Such a world has nothing concrete in it (assuming the empty space is not itself concrete). By imagining such a world, we verify that it is conceivable.

The second premise – that conceivability implies possibility – may be justified by David Chalmers' conceivability tests (2002: 145-200). Chalmers carefully distinguishes several different notions of conceivability and proposes how they might guide us into reliable

judgments about metaphysical possibilities. Conceivability is especially helpful, he argues, when it is 'positive' and 'ideal'. One *positively* conceives of a situation S when one is able to coherently imagine a situation in which S obtains. Chalmers understands 'imagination' as broader than visual imaging (like in a vivid dream), since an imagination can include a conceptual or intuitive representation, such as when one brings to mind the details of a logic or math proof. Conceivability is *ideal* if no amount of further scrutinizing would or could reveal incoherence in what one is imagining. So, for example, if a perfect cognizer is able to (positively) conceive the situation in question, then the situation is ideally (positively) conceivable. But, even without ideal conceivability, one can enjoy prima facie (upon initial inspection) or secunda (upon further inspection) conceivability. These notions provide varying degrees of justification for modal judgments, where the more we inspect, the more justification our modal judgments may enjoy.

We may apply Chalmers's epistemology to our modal judgments about the empty world. It seems we can positively conceive of an empty world. Furthermore, this conception is anywhere between prima facie to ideal conceivability. So if Chalmers's test is a good one, it seems we have good reason to think that an empty world is metaphysically possible.

The next premise is A4: if a world of concrete things is possible, then there is no necessary concrete thing. This premise is plausible given the definition of 'necessary concrete thing': such a thing exists and is concrete in all possible worlds if it exists in any.

The conclusion follows: there is no necessary concrete thing.

Reply:

Conceivability is a two-edged sword. Consider, first, what it takes to conceive of an empty world. I suggested that we can conceive of an empty world by imagining a world that has no concrete things. This imagination goes beyond mere visualization, though: we imagined that there are no concrete things, but no mental image contains what does *not* exist. A mental image of blank 'space', for example, is not a picture of nothing: such an image doesn't preclude the existence of things outside the imagined region; nor does it preclude invisible or non-extended things within it. To imagine a completely empty world, one must bring to mind the very abstract and general idea that there are no (concrete) things of any kind, visible or invisible, spatial or non-spatial. Such imagination

may count as 'positive' in Chalmer's sense, but the imagination involved is deeply conceptual.[3]

Once we allow conceptual imagination, one wonders why we cannot conceptually imagine that there is a necessary being. Try this. Imagine that there is a black particle that never comes into being or goes out of being. Then imagine that the reason the particle persists so long is that it cannot fail to exist. If you can do that, then you can imagine that there is something necessary.

Notice here that we didn't have to get into our minds all possible worlds in order to imagine that there is something necessary, just as we didn't have to get into our minds all possible concrete things in order to imagine that no concrete things are real. The question remains, then: if we can imagine that there is nothing, why can't we equally well imagine that there is something necessary?

Someone might answer that the proposition that something is necessary is importantly different because it includes a *modal* notion: it says that something is *necessary*. By contrast, the proposition that there is nothing is merely a statement about what doesn't exist. No modality is included in that statement. Perhaps conceiving of modal situations is a less reliable guide to possibility than conceiving of non-modal situations. So the stalemate is broken in favour of the argument against necessary beings.[4]

This answer is also two-edged, however. For there are conceivable non-modal situations whose possibility implies that there is necessary being. I'll give one example. Consider a situation in which there are some concrete things, such that for any of them, there is an explanation of the fact that those things exist. Call this situation 'E'. E is non-modal: it includes no statement about what must or could be. Furthermore, there is no obvious reason to think it is more difficult for one to conceive of E than it is to conceive of an empty world. In both cases, one entertains a general situation, either about universal non-existence or about universal explanation. By hypothesis, therefore, we have evidence for the *possibility* of E. But from here, we may deduce that there is a necessary being as follows:

B1: Suppose there is no necessary concrete thing.

---

[3] For more about the difficulties with imagining an absence of reality, see Pruss 2009: 47-49.

[4] I owe this reply to Mark Balaguer.

B2: Then there cannot be a necessary concrete thing.

B3: If there cannot be a necessary concrete thing, then there cannot be an explanation of the existence of all *contingent* (i.e. non-necessary) concrete things.[5]

B4: But, there can be an explanation of the existence of all contingent concrete things (because E is possible).[6]

B5: Therefore, the starting assumption is false: there is a necessary concrete thing.

Each step in the deduction follows from normal definitions. B2 follows from B1, given the modal system in play, which implicitly characterizes 'metaphysical necessity'.[7] B3 follows from an ordinary meaning of 'explanation' on which an explanation of the existence of some contingent things cannot be solely in terms of one or more of those very

---

[5] In terms of possible worlds: if no possible world contains a necessary concrete thing, then for every possible world containing contingent concrete things, there is no explanation of the existence of all the contingent concrete things in that world. (I am assuming for the sake of argument that in the possible world where E obtains, there are *contingent* concrete things. E includes the fact that there are concrete things. So if there are no contingent things, then in the world where E obtains, it straightforwardly follows that there is a necessary concrete thing.)

[6] In terms of possible worlds: there is a possible world containing contingent concrete things, where there is an explanation of the existence of all the contingent concrete things in that world.

[7] Here is a proof of the inference:

Let 'N' abbreviate '$\exists x\, (N(x)$', where '$N(x)$' reads '$\Box\,(\exists!(x)\ \&\ \Diamond\,(\exists y\,(x\ \text{causes}\ y)))$'. Then:

1. Assume $\Diamond N$.
2. Then: $\Diamond\Box N$. ($\Box(N \rightarrow \Box N)$, by axioms **4 & 5**)
3. Now suppose (for the sake of argument) that $\Diamond{\sim}N$.
4. Then: $\Box\Diamond{\sim}N$. (by axiom **5**)
5. Then: ${\sim}\Diamond{\sim}\Diamond{\sim}N$. (by substituting '${\sim}\Diamond{\sim}$' for '$\Box$')
6. Then: ${\sim}\Diamond{\sim}{\sim}\Box{\sim}{\sim}N$. (by substituting '${\sim}\Box{\sim}$' for the second '$\Diamond$')
7. Then: ${\sim}\Diamond\Box N$. (because '${\sim}{\sim}X$' is equivalent to 'X')
8. But (7) contradicts (2).
9. So: (3) is not true. ((3) $\rightarrow$ (7))
10. So: ${\sim}\Diamond{\sim}N$.
11. So: $\Box N$. (by substituting '$\Box$' for '${\sim}\Diamond{\sim}$')
12. So: N. ($\Box X \rightarrow X$, by axiom **M**)
13: So: if $\Diamond N$, then N.
13: So: if ${\sim}N$, then ${\sim}\Diamond N$.

Note that I am treating causal capacity as a sufficient condition for concreteness. But even without that assumption, it is plausible that concrete things are essentially concrete, which is all we need.

same things.[8] To be sure, if the things whose existence is to be explained are not contingent, then there may be an explanation in terms of the impossibility of their non-existence. But that option isn't available for contingent things.

The final premise, B4, is justified by the very conceivability test that gives life to Hume's conceivability-based argument. So we have a parity problem. The conceivability test that we used to garner evidence for the possibility of an empty world gives us evidence for the critical premise in an argument against the possibility of an empty world. The arguments are awash.

It is far from clear how a defender of a conceivability-based argument against a necessary being may avoid the parity problem. Three tasks are required: (i) explain what it means to *conceive* of an empty world; (ii) explain why we should think that conceiving, in this way, is a reliable guide to metaphysical possibility; and, perhaps most importantly, (iii) block 'parity' arguments that purport to show that there are conceivable situations whose very possibility is incompatible with the possibility of an empty world. I won't claim that these tasks 'cannot' all be accomplished in principle. But unless they are, the argument from conceivability ends in stalemate.

Seeing this stalemate is not a trivial matter. We teased it out with the help of post-Humean modal logic – especially to deduce the inference from B1 to B2.

On a final note, it might turn out that a necessary being has a nature, such that if one were to fully conceive it, one could not conceive of it as non-existent. Nothing we *do* conceive rules that option out.

The Problem of Semantics

Richard Swinburne (2012) adds teeth to the Humean principle that whatever can exist can fail to exist. His argument begins with the following semantic assumption:

---

[8] In case there are doubts, we could make the definition explicit by filling out the notion of 'explanation' in play: so, for example, we could run the argument in terms of 'non-circular explanation' or 'non-probabilistic explanation'. What matters for my purposes is that the resulting principle remains conceivable and non-modal. (Note that Maitzen's proposal that we can explain why there are *any* contingent concrete things in terms of mundane facts about certain contingent concrete things (2013: 264) causes us no problem, for the explanandum in my argument is a fact about particular concrete things rather than the generic fact that there are any contingent concrete things.)

Sentence Meaning: the meaning of a sentence includes the conditions of its truth and falsity.

Swinburne motivates Sentence Meaning from observations about our language use. Take a simple illustration. My two-year old says, 'There is a kitty.' I reply, 'No, that is a squirrel.' My reply reveals a condition of falsehood, and the effect is that my child gains a more accurate understanding of the meaning of 'There is a kitty'. In general, the more conditions or truth and falsehood we learn about a sentence, the more fully we grasp its meaning.

The next step in Swinburne's argument is to connect Sentence Meaning with a criterion for necessary truths. A necessary truth, he argues, is such that its sentential negation is self-contradictory. To say that a sentence is self-contradictory is to say this: anyone who fully grasps its meaning can see a priori that it entails a contradiction. Thus, we may put his proposal as follows:

Necessary Truth: p is a necessary truth iff there is a sentence s that expresses the negation of p, where anyone who fully grasps the meaning of s can see a priori that s entails ~s.

The trip from Sentence Meaning to Necessary Truth is a walk in the park. Let n be any necessary truth. There are no conditions on which n is false. So by Sentence Meaning, the meaning of a sentence s that expresses n includes no conditions on which n is false. Therefore, anyone who *fully* grasps s's meaning can see that s cannot be true: for they can see that 's is true' contradicts the consequent of all possible truth-conditions built into the meaning of s. The result, in short, is that necessary truths can be seen to be necessary a priori.

What about so-called *a posteriori* necessities, such as 'Hesperus is Phosphorus' or 'water is H_2O'? How can they be seen to be necessary a priori? Swinburne has an answer. What we discover a posteriori, he suggests, are the truth-conditions built into the meaning of the sentence in question. Take 'Hesperus is Phosphorus' as an example. The terms 'Hesperus' and 'Phosphorus' are what Swinburne calls 'uninformative designators'. Those terms designate an object, but they do not by themselves supply us with enough information about the designated object for us to see that they actually designate the same object. Through empirical observations, we learned that 'Hesperus is Phosphorus' is actually a statement of identity, which, like all statements of identity,

is necessarily true. Again, the very meaning of the sentence includes its truth conditions. So by discovering what those conditions are we thereby discover more about the meaning of the sentence. When we know enough about the meaning of 'Hesperus is Phosphorus' to see that 'Hesperus' designates the same object as 'Phosphorus', then from here we can see a priori that 'Hesperus is Phosphorus' is true. Therefore, 'a posteriori' necessities are not counterexamples to Necessary Truth.

The final step is to show that 'there is a necessary being' is not a necessary truth. Here is Swinburne's reasoning. Suppose there is a necessary being. Then 'there is no necessary being' is self-contradictory, according to Necessary Truth. Therefore, 'there is no necessary being' entails 'there is a necessary being'. But surely the mere non-existence of something cannot entail the existence of something. So: the initial assumption that there is a necessary being is false. There is no necessary being.

Reply:

In my judgment, the final step of Swinburne's argument is the most problematic. I realize that the first step – which involves justifying Sentence Meaning – may raise red flags for some. But we may treat Sentence Meaning as expressing a stipulation on what Swinburne *means* by 'meaning'. Swinburne himself identifies meanings with truth-conditions (2012: 345). So Sentence Meaning is actually true by definition: by seeing what he means by 'meaning', we see that the negation of Sentence Meaning entails a self-contradiction. The next step also seems to follow from the meaning of terms: if Sentence Meaning is true, then sentences that express necessary truths include *in their meaning* the fact that they are true on all possible conditions.

But the third, final step is vulnerable to a parity problem. Consider the following reasoning. Suppose there is no necessary being. Then it is not possible that there is a necessary being (assuming S5). It follows that 'a necessary being exists' is self-contradictory, according to Necessary Truth. Therefore, 'a necessary being exists' entails 'there is no necessary being'. But surely the mere existence of something cannot entail the non-existence of something. So: the initial assumption that there is no necessary being is false. There is a necessary being.

The above argument is exactly parallel to Swinburne's argument. Both arguments appeal to Necessary Truth in the same way. The weakest link in each argument is the premise about entailment. The parallel argument

has it that existence doesn't entail non-existence, while Swinburne's argument has it that non-existence doesn't entail existence. Neither premise is more plausible than the other. So the arguments are awash.

There is more. Anyone who wishes to defend the semantic objection faces the following dilemma: either (i) the sentence 'necessarily, there is an x, such that $x = x$' is self-contradictory, or (ii) it is not. Both options are problematic, however.

Suppose, first, that (ii) is true. Then by Necessary Truth, it is actually *possible* that it is necessary that there is something – that there is an x, such that $x = x$. But that implies that it is necessary that there is something (given S5),[9] which contradicts the conclusion of Swinburne's argument.

So suppose instead that (i) is true: the sentence in question is self-contradictory. Then, if Swinburne's semantic objection is sound, it follows that it is part of the *meaning* of 'necessarily, there is an x, such that $x = x$' that there are no conditions on which it is true. But how could that be? The sentence carries no such meaning on any ordinary sense of the term 'meaning'. Moreover, if we allow an unordinary – or stipulated – meaning of 'meaning', then we are left without a non-question-begging way to justify that (i) is true. The very reasoning Swinburne uses to justify the possibility that there is no necessary being is applicable here. For that reasoning can be reversed to construct a parity argument: just replace 'there is a necessary being' with 'there is no necessary being' and observe that neither is prima facie more likely to be self-contradictory. So at best the semantic objection ends in stalemate.

In reply to all of this, perhaps Swinburne could insist that those of us who grasp the meaning of 'necessity' that he has in mind will be able to see a priori that it is not necessary that there is something. On this basis, they may infer that 'necessarily, there is an x, such that $x = x$' is indeed self-contradictory (given Sentence Meaning). Furthermore, they may infer that there is no necessary being.

Is there a notion of 'necessity' on which one can see a priori that nothing is necessary? If there is, it is far from evident that such a notion is characterized by S5. Consider that there are S5-based arguments in the literature for the necessary existence of *propositions* (see: Plantinga 2003, Carmichael 2010, and Rasmussen 2014: 87-105), and Swinburne admits that his argument is in trouble if there are necessary propositions (2012: 259). Yet, Swinburne doesn't tackle any of these arguments. If one

---

[9] See note 7 for the deduction.

is to successfully argue that 'there is nothing necessary' follows from definitions of terms, it would greatly help to have a diagnosis of where the S5-based arguments for necessary propositions go wrong. And more importantly, we should be able to show that 'necessarily, there is an x, such that $x = x$' is self-contradictory. I don't see a way to extract these further results from the semantic objection. So far it hasn't been done.[10]

The Problem of Subtraction

Recent metaphysics gives us a new argument against the necessity of God's existence. It is the so-called 'subtraction argument', which is designed to show that concrete things could be subtracted, one by one, until there are none.[11] We may state a version of the argument as follows:

C1 (Finite): Possibly, there is a finite number of concrete things.

C2 (Subtraction): For any finite number of concrete things there might be in total, there could be fewer concrete things.

C3 (Leap): If Finite and Subtraction are true, then there is no necessary being.

C4: Therefore, there is no necessary being.[12]

There is a way to find each premise plausible. Start with Finite, which says that there could be a finite number of concrete things. I see a couple ways one might motivate this premise. First, there is an argument from conceivability. It may seem we can coherently conceive of a finite number of concrete things: imagine, for example, two blue spheres with nothing else. If conceivability is evidence of possibility, then it may seem we have evidence here that there could have been a finite number of concrete things. Second, there are the 'paradox-based' arguments against the possibility of an infinite number of concrete things. These arguments attempt to show that if there were an infinite number of concrete things,

---

[10] I have been granting for the sake of argument Swinburne's criterion for necessary truths. But for a recent mathematical challenge to that criterion, see Pruss 2015.

[11] See, for example, discussions by Baldwin (1996), Lowe (2002), Paseau (2002), Rodriguez-Pereyra (1997, 2000, 2002, 2013), Cameron (2006, 2007), Efird, et al. (2005, 2006, 2009), and Hoffman (2011).

[12] Versions of the subtraction argument that treat 'concrete' as synonymous with 'spatially situated' pose no obvious threat to the necessary existence of God, since God is classically understood as a non-spatial object. But I am assuming for our purposes here that *causal-capacity* is sufficient for concreteness.

then certain absurd – apparently impossible – situations would be possible.[13] So there are routes to premise 1.

The second premise – Subtraction – falls out of a principle of 'modal continuity'.[14] The basic thought is that any finite lower bound on the number of possible concrete things would be completely arbitrary. To help us appreciate this point, suppose there are exactly four blue spheres. This number of spheres is not a necessary number. After all, someone could create five blue spheres, perhaps out of playdoh. But if it isn't necessary that there are four blue spheres, then it equally isn't necessary that there are three blue spheres. By this same reasoning, it seems there is no finite number, such that it is necessary that there is exactly that number of blue spheres. Moreover, it seems there is no number, such that it is necessary that there is *at least* that number of blue spheres. There could be any number of blue spheres, including 2 or 1 or none. If we suppose instead that 3 is the least number of blue spheres there could be, then the lower limit on blue spheres is intolerably arbitrary. In this case, modal continuity is broken without any justification. Similarly, if there is some finite number of concrete objects, such that there could be no fewer, then the lower limit on concrete objects is intolerably arbitrary. Again, modal continuity is broken without justification.

One may also motivate Subtraction by inductive reasoning. We observe that for any type of concrete thing, instances of that type can be removed from reality. For example, there can be fewer cars, fewer trees, fewer planets, fewer Helium atoms, fewer solar systems, and so on. Subtraction predicts these observations, and it is far from clear that there is a competing, more virtuous principle (in terms of simplicity, predictive power, etc.) that predicts these same observations. So, one might infer that Subtraction is a plausible generalization from many cases.

Turn, next, to Leap, which connects the previous two premises with the conclusion that there is no necessary being. Leap is easy to demonstrate using mathematical induction. Here is that demonstration. Suppose Finite is true. Then there is a finite number k, such that there could be k concrete things in total. Let P(n) = 'there could be $k - n$ concrete things in total, where $k \leq n \leq 0$.' The base case, where $n = 0$, is true by hypothesis. The inductive step follows from Subtraction: if it is

---

[13] See, for example, Craig & Sinclair 2011: 106–16.

[14] Rasmussen (2014) introduces and defends a principle of modal continuity as a guide to possibility.

true that there could be $k - n$ concrete things in total, then it is true that there could be $k - (n + 1)$ concrete things in total, where $k - (n + 1)$ is finite. It follows, therefore, that P(n) is true in the case where $n = k$. In other words, it follows that there could be zero concrete objects in total. But if there were a necessary being, then there would necessarily be at least one concrete thing. Therefore, there is no necessary being.

Reply:

As with the previous two arguments against the existence of a necessary being, the subtraction argument faces a 'parity' problem. The problem, basically, is that the same sort of reasoning used to support the subtraction argument can be used to support a parallel argument for a necessary being. Here is the parallel argument:

D1 (Finite 2): Possibly, there is an explanation of there being at least n concrete things, for some finite n.

D2 (Subtraction 2): For any finite number n, if there could be an explanation of there being at least n concrete things, then there could be an explanation of there being at least n - 1 concrete things, where n - 1 is a positive number.

D3 (Leap 2): If Finite 2 and Subtraction 2 are true, then there cannot be zero concrete things.

D4: Therefore, there cannot be zero concrete things.

There is a way to find each premise plausible. Start with Finite 2, which says that there could be an explanation of there being at least a certain number of concrete things. So, for example, suppose a factory produces one billion cups. Then the operations of the factory explain why there are at least one billion concrete things.[15]

The second premise – Subtraction 2 – falls out of the principle of modal continuity. The basic thought is that any finite lower bound on how many things could be explained would be completely arbitrary. To help us appreciate this point, suppose there is an explanation of why there are at least four blue spheres. That seems possible: imagine, for example, a factory producing four blue spheres. In this example, we used the number four. But the scenario is no less plausible on any other finite

---

[15] I intend a tenseless reading of 'are'. Thus, when I say 'there are at least one billion cups', I mean to quantify over all cups that there have ever been, presently are, or ever will be.

number. Suppose instead that although there can be an explanation of why there are at least four blue spheres, there *cannot* be an explanation of why there are at least three blue spheres. Then we have an intolerably arbitrary modal boundary. In this case, modal continuity is broken without any justification. Similarly, if there is some finite number of concrete objects, such that there could not be an explanation of there being at least that many concrete things, then we have an intolerably arbitrary modal boundary. Again, modal continuity is broken without justification.

One may also motivate Subtraction 2 using a principle of induction. We observe that for any number of concrete things of any type, there is an explanation available for why there is at least that number of concrete things of that type. For example, there are explanations available for why there are at least various numbers of cars, trees, planets, Helium atoms, solar systems, and so on. Subtraction 2 predicts these observations, and it is far from clear that there is a competing, more virtuous principle (in terms of simplicity, predictive power, etc.) that predicts these same observations. So, one might infer that Subtraction 2 is a plausible generalization from many cases.

Turn, next, to Leap 2, which connects the previous two premises with the conclusion that there cannot be zero concrete things. Here is the reasoning behind Leap 2. First, it follows from the above premises (via mathematical induction) that there can be an explanation of there being at least one concrete thing. Second, the only possible explanation of there being at least one concrete thing is that there *cannot* be fewer concrete things. Therefore, there cannot be zero concrete things.

Someone may object that there is another way to explain why there is at least one concrete thing. For example, one might think that it is objectively unlikely for there to have been no concrete things, and that this unlikelihood explains why there is at least one concrete thing. So Leap 2 is false.

But this objection doesn't get at the heart of the parity argument. Although there may be many ways to explain a given fact, we may explicitly build into Subtract 2 the relevant notion of explanation, such as non-probabilistic and non-circular. The motivations for Subtract 2 (i.e., modal continuity and induction) are no less compelling on this more precise reading.

From all these premises we get the parity conclusion: there cannot be zero concrete things. This conclusion cannot be true if the original

subtraction argument is sound. For if that argument is sound, then there can be zero concrete things. So either the subtraction argument is unsound, or this parallel argument is unsound. Once again we've reached a stalemate.

To break the stalemate, it seems we need some way to see that the original subtraction argument is superior to the parallel argument. It is true that less work is required to defend the subtraction argument's Leap than is required to defend Leap 2. But this work is surely more than offset by the extra work required to defend Finite. After all, it is far more controversial that there cannot be an infinite number of concrete things than that there can be an explanation of there being at least four concrete things (say). In any case, anyone who wishes to press the subtraction argument against a necessary being faces the challenge of showing that no parallel argument for an opposite conclusion is as good. They face the parity problem.

III. OFFENCE

I will close this article by getting on the table a new strategy for arguing for God's necessary existence. The strategy, in a nutshell, is to motivate an instance of the following 'explanation-based' argument schema:

P1. If there is any necessary concrete thing, then God necessarily exists (if God exists).

P2. There is a necessary concrete thing.

P2.1. Any possible state of type T *can* be explained.

P2.2. There is a possible state of type T that cannot be explained unless there can be a necessary concrete thing.

C2.1. Therefore, there can be a necessary concrete thing.

P2.3. If there can be a necessary concrete thing, then there is a necessary concrete thing.[16]

C2.2. Therefore, there is a necessary concrete thing.

C. Therefore, God necessarily exists (if God exists).

There is a way to find each premise plausible. P1 falls out of a minimal conception of God as the ultimate source of all other concrete things.

---

[16] This premise is a theorem of modal system S5 (if concrete things are essentially concrete). See note 7 for a proof.

For if God exists and there is a necessary concrete thing, then in every world where God exists, either God is identical with that necessary thing or God is that necessary thing's cause. In the first case, God is a necessary thing, after all. And surely if God is the *cause* of a necessary thing, then God is also a necessary thing. So, in either case, it seems God is a necessary thing, if there is any necessary thing.[17]

The more controversial premise is P2. Yet one may find support for it from recent developments in modal cosmological arguments.[18] For example, my 'new' argument for a necessary being (2011) can be converted into an instance of the above schema, if we let T = 'state of a property beginning to be exemplified' (cf. Turri 2011). Other Ts include, for example: 'state of contingent existence' (cf. Rasmussen 2010), 'contingent state of concrete particulars' (cf. Weaver 2013), and 'state of things coming into existence' (cf. Pruss & Rasmussen Forthcoming).

One advantage of all these arguments is that, unlike traditional cosmological arguments for a necessary being, they do not require that any state of reality is, or must be, caused or explained. Each argument makes use instead of a premise about what states are *possibly* explained.

More importantly, modal explanation-based arguments do not fall prey to the sort of parity problems that afflict arguments against a necessary being. Consider that the soundness of a modal cosmological argument is fully compatible with the possible *failure* of an explanation for a given T. For example, it could be that there is no explanation of the exemplification of *being a contingent concrete thing* in our world, while in some other possible world there is such an explanation. The parity problem is not on the porch of these arguments.

Perhaps someone will devise a clever parallel argument for every instance of the explanation-based schema. The challenge here, however, is to find a parallel argument whose soundness is evidently *as plausible* (not ad hoc, etc.) as the original. I have no idea how that challenge could be met, but I leave it as an open challenge.

I conclude, then, that the inquiry into which Ts may generate plausible instances is wide open. Parity problems block the main objections to God's necessary existence, as I have argued. Meanwhile, recent modal

---

[17] This argument is given in Pruss and Rasmussen Forthcoming.

[18] I am thinking here of arguments given by, for example, Gale & Pruss (1999), Turi (2011), Weaver (2013), Pruss and Rasmussen (Forthcoming), and Rasmussen (2010, 2011).

cosmological arguments open up a new way to argue for the necessity of God's existence.[19]

BIBLIOGRAPHY

Baldwin, Thomas. 1996. 'There Might Be Nothing', *Analysis*, 56 (4): 231–238

Cameron, Ross. 2006. 'Much Ado About Nothing: A Study of Metaphysical Nihilism', *Erkenntnis*, 64 (2): 193–222

Cameron, Ross. 2007. 'Subtractability and Concreteness', *Philosophical Quarterly*, 57 (227): 273–279

Carmichael, C. 2010. 'Universals', *Philosophical Studies*, 150: 373-89

Craig, W. L., & James D. Sinclair. 2009. 'The *Kalam* Cosmological Argument', in *The Blackwell Companion to Natural Theology*, J. P. Moreland, & W. L. Craig (eds) (Malden, MA: Wiley-Blackwell), pp. 101-201

Efird, David, & Tom Stoneham. 2005. 'The Subtraction Argument for Metaphysical Nihilism', *Journal of Philosophy*, 102 (6): 303–325

Efird, David, & Tom Stoneham. 2006. 'Combinatorialism and the Possibility of Nothing', *Australasian Journal of Philosophy*, 84 (2): 269–280

Efird, David, & Tom Stoneham. 2009. 'Justifying Metaphysical Nihilism: A Response to Cameron', *Philosophical Quarterly*, 59 (234): 132–137

Lowe, E. J. 2002. 'Metaphysical Nihilism and the Subtraction Argument', *Analysis*, 62 (273): 62–73

Maitzen, S. 2013. 'Questioning the Question', in Tyron Goldschmidt (ed.), *The Puzzle of Existence* (New York: Routledge), pp. 252-271

Paseau, Alexander. 2002. 'Why the Subtraction Argument Does Not Add Up', *Analysis*, 62 (1): 73–75

Rodriguez-Pereyra, Gonzalo. 1997. 'There Might Be Nothing: The Subtraction Argument Improved', *Analysis*, 57 (3): 159–211

Rodriguez-Pereyra, Gonzalo. 1997. 2000. 'Lowe's Argument Against Nihilism', *Analysis*, 60 (4) (October 1): 335–340

Rodriguez-Pereyra, Gonzalo. 1997. 2002. 'Metaphysical Nihilism Defended: Reply to Lowe and Paseau', *Analysis*, 62 (2) (April 1): 172–180

Rodriguez-Pereyra, Gonzalo. 1997. 2013. 'The Subtraction Argument for Metaphysical Nihilism: Compared and Defended', in Tyron Goldschmidt (ed.), *The Puzzle of Existence* (New York: Routledge), pp. 197-214

Plantinga, A. 2003. 'Why Propositions Cannot be Concrete', in Matthew Davidson (ed.), *Essays in the Metaphysics of Modality* (Oxford: Oxford University Press), pp. 229-34

Pruss, A. & Joshua Rasmussen. Forthcoming. *Necessary Existence* (Oxford: Oxford University Press)

Pruss, A. 2015. 'Possibility is not Consistency'. *Philosophical Studies* 172: 2341-2348.

Rasmussen, J. *Defending the Correspondence Theory of Truth* (Cambridge: Cambridge University Press)

Rasmussen, J. 2011. 'A New Argument for a Necessary Being', *Australasian Journal of Philosophy*, 89: 351–356

Rasmussen, J. 2010. 'From States of Affairs to a Necessary Being', *Philosophical Studies*, 148: 183–200

Swinburne, R. 2012. 'What Kind of Necessary Being Could God Be?', *European Journal for Philosophy of Religion*, 4: 1-8

Turri, John. 2011. 'A New and Improved Argument for a Necessary Being', *Australasian Journal of Philosophy*, 89: 357–359

Weaver, G. 2013. 'A Church-Fitch proof for the universality of causation', *Synthese*, 190: 2749-2772

TERTULLIAN THE UNITARIAN

DALE TUGGY

State University of New York at Fredonia

Abstract. Tertulian is often celebrated as an early trinitarian, or at least a near-trinitarian, proto-trinitarian, or trinitarian with unfortunate 'subordinationist' tendencies. In this paper I shall show that Tertullian was a unitarian, and not at all a trinitarian.

I. INTRODUCTION

The pugnacious African controversialist Tertullian was a late second and early third century public champion of catholic Christianity. Despite his reputation as anti-intellectual, based on a popular 'sound bite' sized quotation from his works,[1] he is an enormously erudite writer, a philosophical theologian who knows how to argue, and perhaps enjoys it a little more than he should. In his works he asserts Father, Son, and Spirit to be *tres personae* but *una substantia* (three persons and one substance), and he is the first known writer to use the word *trinitas*.[2] For these reasons, as well as his much read polemic against contemporary 'monarchians', he is often celebrated as an early trinitarian, or at least a near-trinitarian, proto-trinitarian, or trinitarian with unfortunate

---

[1] '... the Son of God died; it is by all means to be believed, because it is absurd.' (*On the Flesh of Christ* trans. by Holmes in *The Ante-Nicene Fathers Volume III: Latin Christianity*, ed. by Alexander Roberts, James Donaldson, and Arthur Coxe (1885), chapter 5, p. 525). I shall hereafter abbreviate this volume as *ANF III*.

[2] For what may be Tertullian's earliest extant use of *trinitas*, and some relevant translation difficulties, see my "trinitas" in Tertullian's *On Modesty* (*De Pudicitia*), available at: http://trinities.org/blog/trinitas-in-tertullians-on-modesty-de-pudicitia/ [accessed 18/8/2016].

'subordinationist' tendencies.[3] Against these confusions, I shall show that Tertullian was a unitarian, and not at all a trinitarian.

II. DEFINITIONS

In order to read Tertullian carefully, we must start with clear and uncontroversial definitions. A 'trinitarian' Christian theology says that (1) there is one God (2) which or who in some sense contains or consists of three 'persons', namely, the Father, the Son, and the Holy Spirit, (3) who are equally divine, and (4) (1)-(3) are eternally the case.

In contrast, a 'unitarian' Christian theology asserts that the (1) there is one God, (2) who is numerically identical with the one Jesus called 'Father', (3) and is not numerically identical with anyone else, (4) and (1)-(3) are eternally the case. On this sort of theology, God is 'unipersonal', which is to say that God just is a certain great self. As they are logical contraries, a theologian can't consistently hold both views, although one may have a theology which is neither.

It is a mistake to think that 'unitarian Christian theology' is an oxymoron because Christian theology is *by definition* trinitarian. If that were so, Christian theology wouldn't have existed until some time in the latter half of the fourth century. Needless to say, there were Christians before then, and a few of them were theologians.

A unitarian Christian theology need not be a modern or post-Reformation theory. It is true that when Christians went 'back to the sources' in the 16th century, many of them found no Trinity doctrine there.

---

[3] Partisan interests intrude even on the title page of the most-read English translation of Tertullian's *Against Praxeas*. The translator supplies the subtitle, 'In which he defends, in all essential points, the doctrine of the Holy Trinity.' (Tertullian, *Against Praxeas* trans. by Holmes in *ANF III*, p. 597.) To my knowledge, the case that Tertullian's theology is unitarian and not trinitarian was first cogently argued by John Biddle (1615-1662). (*The Testimonies of Irenaeus, Justin Martyr, Tertullian, Novatianus, Theophilus, Origen ...* [1648] in *The Faith of One God* (London, 1691). While one may cite many recent treatments of Tertullian as (near-) trinitarian, a main source of that confusion has surely been George Bull's *Defensio Fidei Nicaenae* (1688) (English Translation: *Defensio Fidei Nicaenae. A Defense of the Nicene Creed, out of the Extant Writings of the Catholick Doctors, Who Flourished During the First Three Centuries of the Christian Church* (Oxford: John Henry Parker, vol. I (1851), vol. II (1852)) Bull's work was stimulated in large part by the work of the Jesuit scholar Denis Pétau (1583-1652), a.k.a. Dionysius Petavius, who argued that for Tertullian, the Son was younger and less great than the Father. (Bull, *Defense Vol. I*, pp. 9-10; Bull, *Defense Vol. II*, p. 534)

But they were hardly the first to see the one true God as the Father alone. Nor should a unitarian be understood as an anti-trinitarian. While many unitarians nowadays do define their view in opposition to ecumenical-creed-compliant theologies, a unitarian needn't have ever even heard of any Trinity theory. Nor need a unitarian be a 'rationalist', a deist, or a denier of miracles. Nor need a unitarian belong to any specific group or denomination which names itself by that term.[4]

Is my usage of the term 'unitarian' idiosyncratic? Not at all. Christian thinkers often describe Jewish and Islamic theologies as 'unitarian', though they need not be modern, explicitly anti-trinitarian, or 'rationalistic' (whatever that means). Again, this label has been applied to many Christian thinkers of the last three centuries by both friend and foe. This usage is neither historical nor polemical, but simply descriptive. The terms 'unitarian' and 'trinitarian' are convenient contrary terms; a coherent theology, if it is one, can't also be the other.

III. THE WORD 'TRINITY'

Famously, Tertullian uses the word 'trinity', the Latin *trinitas*. Isn't this a dead giveaway that he's trinitarian? No. The Word 'Trinity' has come to mean the tripersonal God, consisting of the eternal, equally divine Father, Son, and Holy Spirit. It is used as a singular referring term for the one God, assumed to be tripersonal.

But both now and then, the word 'trinity' can simply refer to these three: Father, Son, and Holy Spirit. Here it's used as a plural referring term, and use of it does not imply that the items mentioned are parts of a whole, or that they are in any way equal, or even that they are things of the same kind or category. It refers to a triad, a triple, a group of three. For example, trinitarian theologians sometimes say that 'the Bible from start to finish is all about the trinity.' This is true even though the biblical authors evidently had no concept of a triune God, so long as we understand 'trinity' to be a plural referring term, indicating God, God's Son, and God's spirit, whatever those are, and however they're related.

---

[4] Many unitarians have belonged to mainstream, catholic Christian groups. It was only in the late eighteenth century that there began to be churches calling themselves 'Unitarian', though the term had been coined about a hundred years before, as an alternative to the less descriptive, polemical terms 'Socinian' and 'Arian'. I don't capitalize 'unitarian' because here it is used as a description, not as the name of any denomination or group.

Tertullian uses the word in this latter way; for him 'the trinity' is a triad, a group, a plurality, consisting of those three selves. This plurality is not a god, although it is intimately related to, and includes the one god (i.e. the Father). Confusion would be greatly reduced if all Christians did what some translators of Tertullian do: they use the all lower case 'trinity' for the plural referring term, reserving the capitalized 'Trinity' for the triune God of later catholic orthodoxy. If this is our rule, we should not use the capitalized word 'Trinity' anywhere in translating Tertullian.

Does his use of *trinitas* as a plural referring term imply that he's *not* a trinitarian? No, for trinitarians can and do use the word in that way. So, of course, do unitarians.[5] Does this cause confusion? Yes. Capital letters are a good and useful invention. Used here, they help one to see that Tertullian's use of *trinitas* is not, by itself, evidence that he's a trinitarian. I'll now argue that we know from carefully reading all his relevant extant works that he is not.

IV. TERTULLIAN'S THEOLOGY

Although Tertullian is a unitarian, he's an aggressively speculative one who, unlike some latter-day unitarians, does not hesitate to add to the simple creeds of his day a fairly developed metaphysical theory about how God is related to his Son and his Holy Spirit.

For Tertullian, the one eternal God isn't the Trinity, but rather the Father himself. He never describes, mentions, or implies the existence of a tripersonal God. Throughout his works Tertullian assumes the identity of the one God and the Father. In his *An Answer to the Jews*, he says that Christians have 'been converted to the same God from whom Israel ... had departed'.[6] In his *Apology* he says that Christians are those 'who

---

[5] This can be seen, for example, in the book titles of some famous early modern unitarians: Samuel Clarke's magnum opus is entitled *The Scripture-Doctrine of the Trinity* (1st. ed. 1712), and John Biddle published *A Confession of Faith Touching the Holy Trinity, According to the Scripture* (1648).

[6] *An Answer to the Jews* in *ANF III,* ch. 1, p. 152. (See also ch. 14, p. 173.) In his *Prescriptions Against the Heretics*, he presents us with a rule of faith, seemingly a commonly used creed in the catholic churches he knew. It says, in part, '...we believe that there is but one God, who is none other than the Creator of the world, who produced everything from nothing through his Word, sent forth before all things ... this Word is called his Son, and in the Name of God was seen in divers ways by the patriarchs ... and finally was brought down by the Spirit and Power of God the Father into the Virgin Mary, was made flesh in her womb, was born of her and lived as Jesus Christ ... ' *(Prescriptions*

have been led to the knowledge of God as their common Father'.[7] He exhorts the Romans to 'give up all worship of, and belief in, any other being as divine', sticking with the 'one true Lord, the God omnipotent and eternal'.[8] As we'll see momentarily, this can only be the Father.

But don't Christians worship the man Jesus? Tertullian clarifies that 'We worship God through Christ. Count Christ a man, if you please; by Him and in Him God would be known and adored'.[9] Like many a later unitarian, Tertullian argues that the ultimate, *indirect* object of Christian worship of Jesus is the Father; he is worshiped by means of our *directly* worshiping Christ.[10]

Not that Tertullian grants that Jesus was a 'mere man'. He argues that if a mere man like Moses could reveal God, all the more so 'Christ ... had a right to reveal Deity, which was in fact His own essential possession'.[11] We'll see what he means by this shortly.

Tertullian, like other second and third century logos theologians, teaches that Christ is rightly called 'God', and that he is in some sense 'divine'. But so do many later unitarians, such as Samuel Clarke, or John Biddle. In that same chapter of his *Apology*, Tertullian lays out his christology:

> ... nor do we differ from the Jews concerning God. We must make, therefore, a remark or two as to Christ's divinity.
>
> ... He [Christ] proceeds forth from God, and in that procession He is generated; so that He is the Son of God, and is called God from unity of substance with God. For God, too, is a Spirit. Even when the ray is shot from the sun, it is still part of the parent mass; the sun will still be in the ray, because it is a ray of the sun – there is no division of substance,

---

Against the Heretics, trans. by S. L. Greenslade, in *Early Latin Theology,* ed. by S. L. Greenslade (Philadelphia: Westminster Press, 1956), pp. 31-64 (p. 40).) For another such short rule see *On the Veiling of Virgins in ANF III*, ch. 1, p. 27, and *Against Praxeas* translated by Holmes, *ANF III*, ch. 2, p. 598a-b. On the point that the god preached by Jesus and Paul was none other than the god of the Old Testament, see *Against Marcion* in *ANF III,* I.21, p. 286, III.1, p. 321.

[7] *Apology*, translated by S. Thelwall in *ANF III*, ch. 39, p. 46.

[8] *Apology* ch. 34, p. 43.

[9] *Apology* ch. 21, p. 36. The idea of worshiping God through Christ is also mentioned in *An Answer to the Jews*, ch. 4, p. 155.

[10] For this idea and discussion of New Testament passages that presuppose it, see my 'Who Should Christians Worship?', available at: <http://trinities.org/blog/who-should-christians-worship/> [accessed 18/8/2016].

[11] *Apology* ch. 21, p. 36.

but merely an extension. Thus Christ is Spirit of Spirit, and God of God, as light of light is kindled. The material matrix remains entire and unimpaired, though you derive from it any number of shoots possessed of its qualities; so, too, that which has come forth out of God is at once God and the Son of God, and the two are one. In this way He is made a second in manner of existence – in position, not in nature; and He did not withdraw from the original source, but went forth. This ray of God... descending into a certain virgin, and made flesh in her womb, is in His birth God and man united.[12]

For Tertullian, Jesus is in divine in the sense that his matter is also God's matter. But does he share *all* of God's matter, or only some of it? And is this derivation of Jesus from God eternal, or did it happen a finite time ago?

In his important *Against Praxeas*, against the 'monarchian' claim that Father and Son are one being and one person, Tertullian asserts many differences between the two. One of these is that '... the Father is all being, but the Son is a tributary of the whole and a portion, as He Himself declares: "Because the Father is greater than I."'[13] Tertullian is thinking in terms of what metaphysicians now call the material constitution relation. His point is that the Father is constituted by all of the divine matter (that is, spirit), whereas the Son is constituted by only a portion of that same batch of matter.

In his *Against Hermogenes*, Tertullian makes clear that like all the other logos theologians before Origen, he believes the Son to have been caused to exist by God a finite time ago. He tells us that God was not always a Father, for there was a time when the Son did not exist.[14]

---

[12] *Apology* ch. 21, pp. 34-5.

[13] *Tertullian Against Praxeas*, translated by A. Souter (London: Society for Promoting Christian Knowledge, 1920), ch. 9, pp. 46-7. Further citations from *Against Praxeas*, unless otherwise noted, are from this translation. On 'portion' also see ch. 26. Elsewhere he says that 'we should not suppose that there is any other being than God [i.e. the Father] alone who is unbegotten and uncreated... His wisdom [=the logos of John 1 and the Wisdom of Proverbs 8] ...was then born and created, when in the thought of God It began to assume motion for the arrangement of his created works... how can it be that anything, except the Father, should be older, and on this account indeed nobler, than the Son of God, the only-begotten and first-begotten Word?' (*Against Hermogenes*, trans. by Peter Holmes, in *ANF III*, pp. 477-502, ch. 18, p. 487.)

[14] ch. 3, p. 478 (compare: ch. 45, p. 502). Insisting that only 'one God the Father' is eternal (ch. 4), he argues against Hermogenes's theory that matter too is eternal. Citing

These are not mis-statements or slips; they represent Tertullian's settled view, and all these points appear in *Against Praxeas*. There he adds that

> ... even 'before the foundation of the universe' God was not alone, having in Himself alike reason and word in reason, which (word) He had made second to Himself by exercising it within Himself.[15]

This word 'under the name "wisdom"' was 'also created as the second person'.[16] After quoting Proverbs 8:22-5, Tertullian expounds this.

> Then, as soon as God had willed to put forth into His own matter and form that which He had in company with the reason and word of wisdom arranged within Himself, he first brought forth the word itself, having in itself its own inseparable reason and wisdom, that everything might be made through the very (word) by which all had been planned and arranged, or rather already made, so far as God's thought was concerned. For this they [things in the created cosmos] still lacked: they had yet to become known and remembered before the eyes of each person in their appearances and substances.
>
> It is then, therefore, that even the word itself takes its own appearance and vesture, namely sound and expression; when God says: 'Let there be made light.' This is the complete birth of the word, since it proceeds out of God. Having been first created by Him as far as thought is concerned, under the name wisdom – 'the Lord created me as a beginning of ways', [Proverbs 8:22] – then begotten to actuality – 'when he was preparing heaven I was with Him', [Pr. 8:27] – thereafter, making as Father for Himself Him from whom He proceeds and thus becomes His Son, He was made 'first-begotten', as having been begotten before everything, and 'only-begotten', as having been alone begotten from God, in a real sense

---

Proverbs 8, he scolds, 'Let Hermogenes then confess that the very Wisdom of God [i.e. the Son of God] is declared to be born and created, for the especial reason that we should not suppose that there is any other being than God alone who is unbegotten and uncreated. For if that, which from its being inherent in the Lord was of Him and in Him, was yet not without a beginning, - I mean His wisdom, which was then born and created, when in the thought of God It began to assume motion for the arrangement of His creative works, - how much more impossible is it that anything should have been without a beginning which was extrinsic to the Lord! [such as an eternal Matter]... how can it be that anything, except the Father, should be older, and on this account indeed nobler, than the Son of God, the only-begotten and first-begotten Word?' (ch. 18, p. 487)

[15] *Against Praxeas,* ch. 5, p. 38.

[16] ch, 6, p. 38.

> from the womb of his own mind ... Rejoicing, He thereupon addresses
> Him ... 'Thou art my Son, this day I have begotten thee'... [Psalm 2:7][17]

It is easy to miss the significance of this passage. Just as all created things first exist in God's mind, so too does the Word, the pre-human Jesus. It has, Tertullian seems to assume, some kind or degree of reality, but it can't yet, any more than the still merely mental objects of the later cosmos, such as you and me, enter into causal relations with the things in the created realm. But then, as it were, this merely mentally existing thing puts on its proper garment. In *our* language, this is when the Word comes to exist – not as a mere idea, intention, or representation in God's mind, but rather as an intelligent being, like God, 'having in itself its own inseparable reason and wisdom'.[18] Like God himself, the Word is now a body (a substance constituted by matter, namely spirit).[19] In short, the Father 'made His word a son for Himself'.[20]

And this Son is, he says, appropriating Gnostic terminology, a 'projection' from the Father, though he is neither separated from him nor composed of different stuff. The Son is a shrub to the Father's root, a river to the Father's source, and a ray to the Father's sun.[21] Tertullian emphatically embraces the implication that they are two things (beings, entities) not one.[22] He emphasizes, though, that these two are not

---

[17] ch. 6-7, pp. 39-40.

[18] ch. 6, p. 39. 'For whatsoever the being of the Word was, I call it a person and I claim the name 'Son' for Him, and in recognising Him as Son, I claim that He is second to the Father.' (ch. 7, p. 42)

[19] Tertullian's view that God is composed of a special sort of matter (spirit) is not unique to him. See David L. Paulsen, 'Divine Embodiment: The Earliest Christian Understanding of God', available at: http://publications.mi.byu.edu/fullscreen/?pub=1100&index=10 [accessed 18/8/2016].

[20] *Against Praxeas,* ch. 11, p. 50. Compare: *Against Marcion* II.27, p. 318. The idea that there was a time before the pre-human Jesus existed was not invented by fourth century 'Arians', as some suppose. Rather, the view was common to many early logos theorists, those holding to what Wolfson calls 'two-stage' logos theories. On these see Harry Austryn Wolfson, *The Philosophy of the Church Fathers: Faith, Trinity, Incarnation, Third Edition, Revised* (Cambridge, Massachusetts: Harvard University Press, 1970), ch. 12.

[21] ch. 8, p. 44.

[22] 'Everything that proceeds from something, must be second to that from which it proceeds, but it is not therefore separated. Where, however, there is a second, there are two, and where there is a third, there are three. The Spirit is third with respect to God and the Son, even as the fruit from the shrub is third from the root, and the channel from the river is third from the source, and the point where the ray strikes something is third from the sun. ... Thus the Trinity running down from the Father through stages linked and

'separated', in that the portion of matter constituting the Son is still a portion of the Father's matter; that portion didn't break off when God extended it out to bring his Son into existence.

In sum, the Son of God is literally younger than and lesser than God. Also, in contrast to many later catholic interpreters, Tertullian holds that that the Son is 'ignorant of the last day and hour, which are known to the Father only'.[23] Thus, the Son of God knows less than God.

Here is a graphic illustration of Tertullian's trinity.
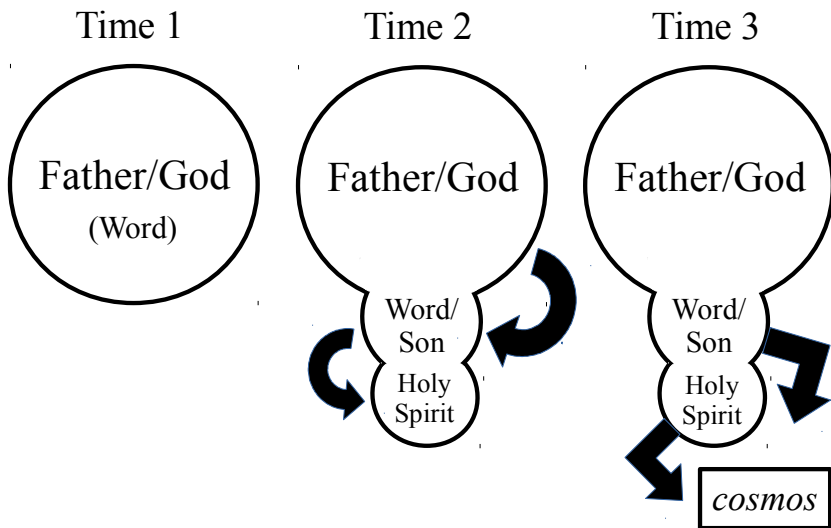
There is one eternal God, but three temporal stages in the career of his divine matter. At first (Time 1), God alone is composed of that matter. At this time 'Word' (*logos*) can only refer to his attribute of wisdom or reason. When it is time to create (Time 2), God extends or stretches himself out so that some of his spiritual matter comes to serve also as the matter of two other beings. In this there is no separation – seemingly, no *spatial* separation of any portion of God's matter from the rest – it remains a contiguous mass. He uses, but does not lose a portion of it, in bringing into existence two other beings. These are literally younger

---

united together, offers no obstacle to monarchy and conserves the established position of the economy.' (ch. 8, p. 45)

[23] *Against Praxeas* 26, p. 103; Matthew 24:36; contrast with *Against Marcion* II.24, p. 315.

and less great than God. Still, both may, like God, be called 'God'.[24] At Time 3, God creates and governs the cosmos by means of these second and third beings. From this time one, we have 'the Trinity of the one Divinity', which is to say, the trinity of beings sharing (various portions of) the divine matter.[25]

This scheme is what Tertullian calls God's 'economy', his way of creating and managing the cosmos (i.e. all that resulted from the Genesis creation). God is the founding member, as it were, of the group he calls the 'trinity', following, it would seem, the second and third century philosophical fashion for positing transcendent triads.[26]

My chart might make you think this whole portion of matter is one thing with three parts at Time 3; but no, it is not a thing, but merely a quantity of matter. The Father is one entity, the Son is a second, and the Spirit is a third. Nor are they parts of any whole; they simply share some of the divine stuff. The three together are not one God. They are three beings constituted by different portions of the divine matter: all (Father), less (Son), and still less (Spirit).

Tertullian's discussion of the Genesis creation in *Against Praxeas* ch. 12 is instructive. There are, he argues, only four possible explanations for why God says 'Let us make man in our image'.[27] First, perhaps God is really alone but is 'either deceiving or making fun of us'.[28] Tertullian doesn't take this seriously. Second, perhaps God (the Father) is speaking to his angels. This Tertullian derides as Jewish. Both of these explanations are less than Christian, he thinks. The third explanation is: 'Or was it because He was Himself Father, Son, Spirit, for that reason, showing Himself to be plural, He spoke in a plural way to Himself?'[29] The trinitarian reader may jump at this as the best answer, but to Tertullian it can only be a heretical, 'monarchian' mis-reading. (In ch. 10-11 he's at length refuted the 'monarchian' idea that God made himself his

---

[24] The reader must also be aware that Tertullian sometimes calls the pre-incarnate Son God's 'Spirit', 'Substance', or 'Reason'. (*Against Marcion* III.6, 326; *On Prayer*, translated by Thelwall in *ANF*, ch. 1, p. 681.)

[25] *On Modesty*, translated by S. Thelwall in *ANF*, ch. 21, p. 99.

[26] See the discussion and sources cited in my 'History of Trinity Doctrines (Supplement to 'Trinity')' in *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), ed. by Edward N. Zalta, available at: http://plato.stanford.edu/entries/trinity/trinity-history.html [accessed 18/8/2016].

[27] Genesis 1:26.

[28] ch. 12, p. 54.

[29] ch. 12, p. 55.

own Son.) He rejects this third explanation in favour of a fourth: that God is at this time working together with two others to create the human race, and so he speaks to these two 'as with servants and eyewitnesses'.[30]

V. WHAT WAS TERTULLIAN'S ANSWER
TO THE 'MONARCHIAN' CATHOLICS?

Tertullian tells us that many ordinary Christians in his day worried that this theory implied a denial of monotheism. They objected that if both Father and Son are divine, doesn't this mean there are at least two gods, something contrary to Christian tradition?[31] This candid admission by Tertullian is evidence that the sort of logos theory which was popular among some of the educated Gentile catholic elite in the late 100s and early 200s, was controversial among common Christians, who wondered if Jesus was being misunderstood as a second god, equal to his Father. This sort of concern about logos theologies was probably a main source of the 'monarchian' theologies of that era. It is exceedingly difficult to characterize these theologies; for present purposes we assume with Tertullian that at least some of these 'monarchians' thought the Father and Son to be numerically identical, and so incapable of differing, with the consequence that the Father was crucified and died.[32] Despite its evidently false implications, this theory does prevent the Father and Son from being two gods, since if 'they' are really numerically one being, 'they' can't be different *anythings*. Having secured monotheism in this way, they demand an answer of the logos theologian.

What is Tertullian's answer to his 'monarchian' critics? Does he argue that the Father, Son, and Spirit are really one god? No. It is here that

---

[30] *Ibid.*

[31] In *Against Praxeas* he observes that 'All simple people, not to say the unwise and unprofessional (who always constitute the majority of believers), since even the rule of faith itself removes them from the plurality of 'the gods' of this world to 'the one true God', become greatly terrified through their failure to understand that, while He must be believed to be one, it is along with His economy, because they judge that economy, implying a number and arrangement of trinity, is really a division of unity, whereas unity, deriving trinity from itself, is not destroyed by it, but made serviceable. Therefore they [i.e. the simple] circulate the statement that two and three are preached by us, while they judge that they are worshippers of one God... 'We hold to monarchy', they say...' (ch. 3, p. 31) But Tertullian too is committed to monotheism; he argues that it is impossible that there be more than one god in his *Against Marcion* I.3-8, pp. 273-6.

[32] *Against Praxeas* ch. 1-2, pp. 28-30.

later interests and commitments obtrude into our reading of Tertullian. We expect him to argue that because the three are 'one substance' they are one god. But that's not what he says. Does he argue, like some latter-day trinitarians, that the term 'monotheism' should be redefined, so as to make trinitarianism 'monotheistic'?[33] No. In *Against Praxeas* ch. 2, Tertullian *agrees* with the 'monarchians' about there being exactly one god. What did they believe in one of? They believed, it seems, in exactly one mighty self, then added, incoherently, that this one self was both the Father and the Son of the New Testament. So too does Tertullian believe in one self, the Father, as the one god. But he makes it abundantly clear that the Son and Spirit are not on the Father's level – they're not divine in precisely the way that the Father is divine.

Instead of making trinitarian moves, Tertullian replies that he too upholds the monarchy, the single rule of the one god. Just as a ruler administers his single kingdom through many lower subjects, so God, all admit, works through angels. But it's still God's unified domain of rule. And if multitudes of angels leave the Father's monarchy intact, all the more so when it comes to a Son who is composed of some of his own spiritual matter.[34]

Tertullian is well aware that this won't satisfy his 'monarchian' critics. The monarchians can retort that it is no renunciation of polytheism to claim that the gods cooperate under a chief god, to carry out the chief god's rule. Conceivably, a pantheon may be orderly, and ruled by a single boss. Tertullian records how they pressed their polytheism objection:

> ... [the monarchians object that] if God spoke and God acted, if God spoke and another acted, you are proclaiming two gods.[35]

His reply, citing several biblical texts, is that there is nothing wrong, so far as the Old Testament is concerned, with calling beings other than God 'God' or 'Lord', in particular Christ. This is an important and often forgotten point.[36] To it, Tertullian adds that those in the catholic tradition,

---

[33] For this sort of move in recent trinitarian theorizing, see my 'Constitution Trinitarianism: An Appraisal', *Philosophy and Theology*, 25:1 (2013), 129-62 (pp. 138-41).

[34] *Against Praxeas*, ch. 4.

[35] ch. 13, p. 57.

[36] Jesus makes this point in John 10. (Dale Tuggy, 'Jesus's argument in John 10', available at: <http://trinities.org/blog/jesuss-argument-in-john-10/> [accessed 18/8/2016]) In another context Tertullian forcefully makes the point that the fact of beings sharing a name doesn't imply that they are the same sort of being. 'If an identity of

although they hold three beings to be 'God', refuse to *say* 'two Gods or two Lords'. In the past, Tertullian thinks,

> ... two gods and two lords were preached simply in order that when Christ had come, he might be recognized as God and also called Lord, because he was the Son of God and the Lord.[37]

But now that the pagans are leaving their false gods to worship one God, it is important that Christians should leave off speaking of many gods and lords, even though the words 'God' and 'Lord' may be applied to Father, Son, or Spirit. (In part, this prevents Christians from having an easy escape from martyrdom, by swearing 'by gods and lords'.)[38] Therefore, continues Tertullian,

> ... I will not use at all the expressions 'gods' or 'lords', but I will follow the Apostle [Paul], and if I have to name the Father and Son together, I will call the Father 'God' and name Jesus Christ 'the Lord'.[39]

But when mentioning Jesus *alone*, he'll follow the same Apostle in calling him 'God'.[40]

> For I shall also name a ray of the sun by itself 'sun'; but in naming the sun whose ray it is, I shall not straightway call a ray 'the sun'. For I am not going to make out that there are two suns. Nevertheless, I will just as much count the sun and its ray two things and two aspects[41] of one indivisible material, as I do God and his Word, as I do Father and Son.[42]

Tertullian's point here is easy to understand. Imagine that a husband calls his wife, his daughter, and his cat 'Honey', but when his wife is near, he only calls her 'Honey'. Or imagine a workplace in which the employees

---

names affords a presumption in support of equality of condition, how often do worthless menials strut insolently in the name of kings - your Alexanders, Caesars, and Pompeys! This fact, however, does not detract from the real attributes of the royal persons. Nay more, the very idols of the Gentiles are called gods. Yet not one of them is divine because he is called a god. It is not, therefore, for the name of god, for its sound or its written form, that I am claiming the supremacy in the Creator, but for the essence to which the name belongs ... that essence alone is unbegotten and unmade - alone eternal, and the maker of all things ... (*Against Marcion* I.7, pp. 275-6)

[37] ch. 13, p. 60.
[38] ch. 13, p. 61.
[39] *Ibid*. Compare: Romans 1:7; 1 Corinthians 8:6.
[40] Romans 9:5.
[41] The Latin word here is *species*. Various translators here use 'forms', or 'manifestations'.
[42] ch. 13, p. 61.

call their supervisor 'Boss' and their manager 'Boss', but when the owner of the business comes for a visit, they only address the owner as 'Boss', reverting to 'Mr. Smith' and 'Mr. Jones' for the manager and supervisor.

VI. IS TERTULLIAN A CONSISTENT MONOTHEIST?

In my view, Tertullian's theology, whatever its problems, is, *pace* his 'monarchian' critics, monotheistic. Monotheism is the thesis that there is one god, not the thesis that there is only one being who may properly be described or addressed as 'god'.[43] Tertullian, like nearly all monotheists, also believes in beings who are greater than humans and 'divine', though not gods in the way that Yahweh, the one true God, is a god.[44] Above and before all these, there is one being, the Father, who is divine in a way that no other being is; he alone is eternally divine, and he alone possesses *all* the divine matter and all the divine attributes. Thus, Tertullian has strong grounds for claiming to be a self-consistent monotheist.

But in arguing with the 'monarchians', Tertullian muddies the waters. In *Against Praxeas* ch. 18 he argues that when, e.g. in Isaiah 45 God/ the Father/Yahweh asserts himself to be the only god, he has in view the idols, the pseudo-gods worshiped by the nations surrounding Israel (and by unfaithful Israelites). As Tertullian reads the passages, the Father, though the prophet, denies the existence of any other gods *not sharing his nature/substance*. (Or is it only 'gods', other beings who can be called 'god'? Lacking quotation marks, his Latin is ambiguous.) But his own Son, 'being in the Father' and sharing part of his substance, would not be ruled out. The right kind of monotheism, Tertullian is insisting, allows multiple gods (that is, multiple beings composed of divine matter), so long as they share a nature (some portion of matter) with the god of Israel.[45]

This would only further rile his 'monarchian' critics. They would object that now the theory has two gods (two beings with divine matter). And as Tertullian knows, they will object that in Isaiah God says that

---

[43] There is no good name for this latter thesis; I can only suggest the unlovely 'mono-"theos"-ism'. Note that it neither implies nor is implied by monotheism.

[44] On monotheism and its compatibility with belief in multiple lesser deities, see my 'On Counting Gods', *TheoLogica* (2016), 1-26. Available at http://revistatheologica.com/index.php/rtl/article/view/9/11 [accessed 20/12/2016].

[45] In my view Tertullian misreads the book of Isaiah here. For my take on its monotheistic passages see my 'Divine Deception and Monotheism', *Journal of Analytic Theology*, 2 (2014), 186-209.

'I am Yahweh, who made all things, who alone stretched out the heavens, who by myself spread out the earth ...'[46] Doesn't this rule out the logos theologians' contention that God created by means of his Son and Spirit – so, not alone, but through two helpers?

Tertullian insists that the prophet's statements rule out only 'other powers', evidently, beings not composed of divine stuff. He appeals to Lady Wisdom in Proverbs 8, with God at the time of creation, as obviously the pre-human Jesus, and as the one called *logos* in John 1, and the 'Word' by which God made the heavens.[47] In sum, the Father 'alone stretched out the heavens ... namely, alone with his Son, even as He is one with the Son'.[48] This can convince only those already committed to the logos theologians' readings of those passages.

But more importantly, Tertullian hasn't made clear why, even though there are two who can be addressed or described as 'god', and two who are composed of divine stuff, there is only one god, and not two. Exasperated, he resorts to linguistic fiat. As these two share a portion of matter, we should not *say* 'two Gods' or 'two Lords'.[49]

But this is a *non sequitur*; if two beings share a portion of divine matter, it doesn't seem to follow that we shouldn't call them 'two gods', because it is unclear why two gods couldn't share a portion of matter. Conjoined twins, for instance, share a portion of their matter, yet we consider them two human beings. Their parent will truly say, 'These are my daughters.' We can understand why the 'monarchians' challenged the logos theologians to actually preach 'two gods' and 'two lords' in accordance with their theory.[50]

But Tertullian here fails to draw on his own theological resources employed elsewhere. In his *Against Marcion*, he defines what he means by a 'god' (in the highest sense of the word).

> In order ... that you [i.e. Marcion] may know that God is one, ask what God is, and you will find Him to be not otherwise than one. So far as a human being can form a definition of God, I adduce one which the conscience of all men will also acknowledge, – that God is the great Supreme, existing in eternity, unbegotten, unmade, without beginning,

---

[46] Isaiah 44:24; *Against Praxeas* ch. 18-19.
[47] Psalm 33:6.
[48] ch. 19, p. 78.
[49] ch. 19, p. 79.
[50] ch. 13, p. 57.

without end. ... nothing is equal to Him ... if He had an equal, He would be no longer the great Supreme. ... That Being... must needs be unique ... God is not, if He is not one. ... Whatever other god, then, you may introduce, you will at least be unable to maintain his divinity under any other guise, than by ascribing to him too the property of Godhead – both eternity and supremacy over all. How, therefore, can two great Supremes co-exist, when this is the attribute of the Supreme Being, to have no equal, – an attribute which belongs to One alone, and can by no means exist in two?[51]

Like all monotheists, Tertullian defines the one God as necessarily unique; there is a contradiction in supposing two of them. Note that on his own views, as we explored them above, the Son and Spirit are *not* gods. They are neither eternal (without beginning or end in time), nor unbegotten, nor (arguably) unmade, nor without beginning. Of the divine attributes Tertullian lists here, the one they would presumably have is being *without end*, that is, never ceasing to exist. In contrast, in his view, the Father has all these attributes, because the Father just is the one God himself.

VII. OBJECTIONS AND REPLIES

Those committed to trinitarian theology have always been reluctant to admit Tertullian to be a unitarian, as defined above. Generally, they ignore at least one aspect of his overall theory, and seize on something he explicitly says, arguing that it implies a more Nicene-compliant view. In this section, I'll briefly answer four such objections.

Objection 1: Tertullian says that in eternity God was not alone. But 'how frivolous, how unmeaning ... is this mode of proof: God was not alone before the creation ... because even at that time He was rational!'[52] Surely, then, his point is that his *logos* was even then, in eternity past, a divine self.

Reply: Tertullian's point is that before the projection of the Son, God was literally alone, but *figuratively* he was not alone, since he was accompanied, as it were, by his own reason or thought. Here's the crucial passage in *Against Praxeas*:

---

[51] *Against Marcion* I.3, p. 273; compare I.5. On divine eternity see I.8.
[52] Bull, *Defense Vol. II*, p. 519.

... at first God was alone, He was to Himself both universe and place and everything, alone, moreover, because there was nothing outside but Himself. But even at that time He was not alone; for He had with Him what He had in Himself, namely, His reason. ... This reason is his own thought; this is what the Greeks call 'Logos', which word we translate also by 'speech', and therefore it is now our (Latin) custom by a simple translation to declare that 'the Word was in the beginning with God', although it is more fitting that reason should be regarded as the older, because a God rational even before the beginning is not from the beginning given to speech, and because even speech itself, since it depends on reason, shows that the latter is earlier, as being its foundation. Yet for all that there is no difference. For although God 'had' not yet 'uttered his Word', [Psalm 107:20] all the same He had it both with and in reason itself within Himself, while silently meditating and arranging with Himself what He was afterwards to state in word. For meditating and arranging in company with His reason, He made that into word which He was dealing with by word.[53]

Tertullian then makes an analogy with human thought; this is how we're to understand what he says above.

See, when you silently meet with yourself in the process of thinking ... By means of reason you think in company with word, and speak, and when you speak through a word, you are thinking. So somehow there is in you a second word, through which you speak when meditating and through which you meditate when speaking: the word itself is different.[54]

We can't read Tertullian here as holding that your thought, your 'word' is literally a second being, a second self, with whom you converse. Rather, you're meeting *with yourself* – this 'second' is only your thought, and potentially something you say. Now, his application:

With how much more completeness, then, does this take place in God, whose 'image and likeness' [Genesis 1:26] you are deemed to be! Since He has reason in Himself even when silent, and in having reason has word also ... even then 'before the foundation of the universe' God was not alone, having in Himself alike reason and word in reason, which (word) He had then made a second to Himself by exercising it within himself.[55]

---

[53] ch. 5, pp. 36-7.
[54] ch. 5, p. 37.
[55] ch. 5, pp. 37-8.

Tertullian then proceeds to explain this pre-cosmos creation or emanation of the Son. Is this view of God's eternal reason and thought (word), as the objector says, 'frivolous'? Even if it were, it's clearly what Tertullian is asserting.[56]

Objection 2: Tertullian says that the Word was 'always in the Father' and was 'always with God'. Thus, the Word (the personal, divine being, the pre-Incarnate Jesus) existed from eternity past.[57]

Reply: No, Tertullian's point in that passage is that even though the Son is projected out of the Father (as we've seen, a finite time ago), he never separates from the Father.[58] The point concerns stages 2 and 3 of the career of the divine matter, and it doesn't follow that the Son/Word exists as a self at stage 1.

Objection 3: Tertullian expressly says that God is eternal, and that the Son is God. He must think, then, that the Son is eternal, and so clearly your interpretation has gone wrong somewhere or other.

Reply: No; as we've seen, Tertullian repeatedly and expressly, in multiple works, says that the Father is older than the Son, who proceeded out from him when it was time to create. Like many a latter Christian apologist, when addressing a general, non-Jewish audience, Tertullian is inclined to leave Jesus out of it. He wants to urge belief in one God, as against the many gods of Roman tradition, and instinctively knows that mentioning the Son as 'God' or 'a god' who distinct from the Father will raise unwelcome objections. (If *you* can have more than one god, why can't *we*?) In an apologetic public letter to a Roman official, he mentions

> ... the human being next to God who from God has received all his power, and is less than God alone ... as less only than the true God – he is greater than all besides ... [59]

Tertullian is not talking about the Son of God here, but rather the Roman emperor! Another striking example of leaving Jesus out of it is his lengthy *Ad Nationes*, an apology to the citizens of Rome.[60] Yes, Tertullian clearly

---

[56] See also the passage from *Against Hermogenes* in note 14 above.

[57] Bull, *Defense Vol. II*, pp. 527-8.

[58] ch 8, p. 44.

[59] *To Scapula*, translated by S. Thelwall in *ANF III*, p. 106.

[60] In *ANF III*, pp. 109-47. See also *The Soul's Testimony* in *ANF III,* pp. 175-80. An exception is his *Apology*, addressed to Roman leaders (ch. 21).

asserts that the one God is without either beginning or end in time.[61] In such contexts, he makes various unqualified statements about what a *deus* must be, but he clearly has in mind the Father, and not the others who may be addressed or described using other meanings of *deus*. In such contexts he also says that *deus* is 'unbegotten', the great and unique Supreme without equal, and invisible – things which he denies about the 'God' who is the Son.[62]

Objection 4: Look at the capstone of *Against Praxeas*; it is clear that like Nicene-era catholics, Tertullian has moved beyond the mere monotheism of Judaism to a properly trinitarian theology, or nearly so. He says,

> But this attitude of yours [you 'monarchians',] belongs to the Jewish faith, I mean the belief in one God in such a way as to refuse to count the Son along with Him, and after the Son the Holy Spirit. ...What need is there of the Gospel ... if the Father and Son and Holy Spirit, 'three' objects of belief, do not thereafter establish one God? God wished to make the mystery new in such a manner that He should be believed to be One in a new way through the Son and Spirit, that he should now come to be known as God face to face in His own special names and persons, who though preached in the past also through the Son and Spirit, was not understood.[63]

Reply: The passage poses no problem for the interpretation of Tertullian argued for here. The new way to believe in God is believing in God's (that is, the Father's) 'economy', by which (Tertullian thinks) he made and has always governed the cosmos. These two 'special names and persons' are two other beings, composed of portions of God's stuff, divine but lesser and younger agents in addition to God who work on behalf of God, even appearing and speaking in God's place throughout, Tertullian thinks, all of the Old Testament. The problem with the Jews isn't that they fail to believe in a tripersonal god, stubbornly sticking with a unipersonal one.

---

[61] *Against Marcion* I.3; *Apology* ch. 30; *Ad Nationes* II.3; *The Soul's Testimony* ch. 1; *Against Hermogenes* ch. 3-4, 6.

[62] *Against Marcion* I.3; *Ad Nationes* II.4.

[63] ch. 31, pp. 116-7. Some trinitarian translators can't resist anachronistic mischief here. Thus, Evans, 'what is the confidence of the New Testament ... unless thereafter Father Son and Spirit, believed in as three, constitute one God?' (*Against Praxeas*, trans. E. Evans (1948), available at: <http://tertullian.org/articles/evans_praxeas_eng.html> [accessed 18/8/2016]) Another translator has: 'both believed in as Three, and as making One Only God.' (*Against Praxeas*, trans. Holmes (1870) in ANF, 597-632, ch. 31, p. 627.)

Rather, they reject God's work through his Son, and in so doing they reject God himself.[64]

VII. CONCLUSION

Let's revisit our definitions in section II above, to be absolutely clear about where Tertullian fits in. Tertullian's theory clearly meets the first condition of both definitions, as there is for him one true God at any time, though at stages 2 and 3 there exist lesser beings which are divine in lesser ways, each of whom may be called 'God'. Just as clearly, his theory fails both the third and fourth conditions for being trinitarian.

Does it meet the second condition for being trinitarian? Arguably it does, at least at stage 3, if 'contains' means that various portions of God's spiritual matter constitute or compose three selves. Here is the kernel of truth in the common misreading of Tertullian as a trinitarian. He does have an idea in common with at least *some* trinitarians, those who hold that the 'persons' of the Trinity share a common stuff or matter, or something analogous to this.[65] Of course, it puts him at odds with many other trinitarians, particularly those who deny that God is a material object, those who affirm a 'classical' doctrine of divine simplicity, and those who hold to a one-self understanding of the Trinity.[66]

In contrast, Tertullian uncontroversially satisfies the all the criteria for having a unitarian Christian theology. He's a unitarian just as much as a number of well-known early modern philosophical theologians, such as Isaac Newton, Samuel Clarke, Thomas Emlyn, and Noah Worcester, as well as most leading pre-Nicene theologians, such as Justin, Irenaeus, Origen, and Novatian.[67]

---

[64] See the end of chapter 31, the very end of *Against Praxeas*.

[65] Jeffrey E. Brower and Michael C. Rea, 'Material Constitution and the Trinity', *Faith and Philosophy* 22:1 (2005), 57-76. Unlike Tertullian, they suggest that the 'nature' or 'essence' shared by the Persons of the Trinity is only *somewhat like* matter, playing a role in the Trinity similar to the role matter plays in a form-matter compound such as a marble statue. Also unlike Tertullian, they suggest this *all* of this nature is shared by the Persons, and that this sharing occurs eternally. They clearly face the objection that their theology isn't monotheistic, but are arguably less successful than Tertullian in answering it. (Tuggy, 'Constitution', pp. 138-41)

[66] For the classification of some recent Trinity theories as 'one self' theories, see my 'Trinity', *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), ed. by Edward N. Zalta, available at: http://plato.stanford.edu/entries/trinity/ [accessed 18/8/2016], section 1.

It is not true that mainstream or catholic Christianity has always taught a trinitarian theology. These second and third century theologians were leading catholic intellectuals in their day, teaching unitarian theologies. It is an interesting question how and why, starting in the latter half of the fourth century, this began to change.[68]

THE PROBLEM OF EVIL AND A CRITIQUE
OF RELIGIOUS REASON

VLADIMIR K. SHOKHIN

Russian Academy of Sciences, Moscow

Abstract. The author's goal is to weigh capabilities of theistic reason in regard to the problem of evil, and two formats of reasoning in this regard are strictly differed, i.e. attempts at building theodicies (as universal, generally valid and transparent for all reasonable persons, both believers and nonbelievers, models of explanation of causes, dimensions and distributions of evils and sufferings in the world of the Divine origin and government) and defenses (counterarguments to atheistic inference of the non-existence of God from the abundance of sufferings in the world). The upshot is that while there is no doubt that the great multitude of evils and sufferings in the world are surely beyond reach of any theodicies, it is similarly doubtless that many sound reasons are suitable for countering atheist "evidential refutations". Some new arguments are offered to counterbalance Rowe's "friendly atheism", Draper's "hypothesis of indifference", and Schellenberg's "argument from hiddenness", along with analysis of wishful thinking underlying all of them.

The problem of evil, one of the most discussed in philosophical theology[1] from its very beginnings, is being solved nowadays in a fashion different from the prevailing manner of coping with it during the former ages. Instead of resolving the issue by means of rational argumentation in the

---

[1] I emphasize that the question is about *philosophical theology* and not *philosophy of religion* in spite of very widespread identification of these disciplines, especially in analytic philosophy. Here I allude to my view that philosophy of religion should be *philosophy-on-religion* having Religiöse (Johann Gottlieb Fichte) and not *philosophy-in-religion* for its topic. For detailed arguments see, inter alia: V.K. Shokhin, "Methodological Pluralism and the Subject Matter of Philosophy of Religion", in *Knowledge, Action, Pluralism*. Ed. by S.T. Kolodziejczyk and J. Salamon. Frankfurt am Main: Peter Lang GmbH, 2014, pp. 321-324.

framework of the established religious world-outlook, contemporary religious philosophers try with increasing frequency to modify the very concept of God to fit it unto current ways of thinking. The case with open theism, where God is being relieved from responsibility for the world overburdened with moral evil and physical suffering by delivering Him from the burden of "meticulous providence"(Michael Peterson) along with "ultimate attributes" of theistic Personal Absolute has been only first gambit here to be followed by such devices as what I would call "theology of a weak Divinity" (in vogue with many continental "theological avant-gardists") , i.e. of a low-powered God from whom abuses for bad state of affairs in the world could be safely called away. In my opinion, the price of such endeavors to "help God" is too high, approximately as counteracting hair loss by using a guillotine.

In a slightly similar way opposition to the atheistic argument from of evil at the cost of essential characteristics of classical theism is on hand also in the latest issues of EJPR. To give one example, in his controversy with the renowned argument from unbelief one author admits that one can have personal relations with another without being aware who exactly that other is, to the upshot that "one might have a relation to Truth or Justice, while not realizing that those absolute values are God"[2]. What is overlooked there is that, first, personal relations have nothing to do with respect for abstract values and, second, that Truth and Justice might be understood so differently and widely that even an avowed atheist can be regarded a theist with such expanded understanding of the communion with God. The main thing, however, with such an argument is that Personal Absolute is evened to "absolute values", and this is a farewell to theism. In another article the endorsement of the former verdict of the author on the issue of evil (along with on the issue of religious diversity) is made[3] which resolves the problem by means of "theodicy of justice as fairness" (methodologically in line with Rawls' theory of justice). According to this theodicy based on "our newly acquired egalitarian insights" it is more plausible to assume that good and just God never intervenes to change the natural course of events to prevent evil from

---

[2] Cyrille Michon, "On Schellenberg's Argument from Ignorance", *European Journal for Philosophy of Religion*, 2015, Vol. 7/1, pp. 87-88.

[3] Janusz Salamon, "Atheism and Agatheism in the Global Ethical Discourse: Reply to Millican and Thornhill-Miller", *European Journal for Philosophy of Religion*, 2015, Vol. 7/4, p. 212.

happening than to admit that He does it sometimes in an "outrageously selective way" (and has therefore some "favourites")[4]. But such a God is closer to deistic than to theistic divinity inasmuch as deism insisted just on the Divine non-intervention into the affairs of the world, as was the case with "aristocratic" Epicurean gods dwelling in the intermundia, the space between the cosmoi without any concern for our concerns[5]. On the other hand, we have here a typical deistic "religion within the limits of reason alone", because theistic God says of himself that *my thoughts are not your thoughts, neither are your ways my ways ... For as the heavens are higher than the earth, so are my ways higher than your ways, and my thoughts than your thoughts* (Isaiah 55: 8-9) and it is highly doubtful therefore that He should be governed by our philosophical theories of justice[6].

In contrast, I suppose that classical theism has not as yet exhausted all its potentialities for self-defence against atheistic argument from evil in such a degree that it should be substituted by some other, more "up-to-date" religious doctrine, provided that it should be able for critical

---

[4] J. Salamon, 'Theodicy of Justice as Fairness and Sceptical Pluralism', in: *Knowledge, Action, Pluralism*'. Ed. S.T. Kolodziejczyk, Frankfurt, 2014, pp. 250-251; for a related argument in favour of "mystical inclusivism", cf. J. Salamon, 'Light Out of Plenitude', *European Journal for Philosophy of Religion* 2(2010), pp. 141ff..

[5] It was not by chance that the first critics of Deism in the 17th century nicknamed its champions Epicureans. Deists of that age were by no means atomists, and neither is Janusz Salamon, being much closer to Plato and Kant. The point is that deistic *attitude* to religious issues is by no means historical heritage alone and is in a sense a theological archetype. For example, the whole movement of "religious pluralism" culminating with John Hick has many parallels with deistic attitude to religious diversity in the fashion of Wollf's follower Georg F. Meier (1718-1777), Hermann S. Reimarus (1694-1768) and his famous pupil Gotthold Ephraim Lessing (1729-1781).

[6] The Old Testament is corroborated by the Gospels. One can refer only to the Parable of the Workers in the Vineyard where those who worked the whole day and *have borne the burden and heat of the day* were made by the householder equal to those who *have wrought but one hour* and listened that it was his right to do what he wills with his own (Matthew 20: 12-15). There is no doubt that such reasoning is in the way of implementing any versions of distributive and even moral justice theories from Aristotle to Rawls' followers to "Divine economy". Other transgressions of the principle of human justice are at hand also in the Parables of the Prodigal Son (Luke 15: 29:30) and of the Unjust Stewart (Luke 16: 18-19), while some elements of "injustice" are present also in that of the Lost Sheep (15: 3-4). The reason thereof from the classical theistic point of view is ontological: our reason is not an appropriate instrument for measuring all the reasons of its Creator.

attitude not only to its opponents but also some its own historical clichés. And I express profound gratitude to the editorial board of EJPR for giving me an opportunity to defend this thesis.

When one hears the term *the problem of evil,* one does not need many comments. The argument under discussion has always been the most vexing for religious mind, and not only for classical theism, but also for other religious world-views. Two and half millennia ago Plato had to take into account the views of those who were skeptical about gods' participation in the affairs of this world because of prosperity of impious people and their descendants, as well as misfortunes of the virtuous ones[7]. At the same time some early Brahmanists, who were in a sense "open theists"[8], had to deal with Buddhist taunt concerning doubtful compatibility between the existence of a good creator and sufferings of living beings[9]. But classical theism is indeed highly sensitive to the issue of evil. In truth, the argument against it from allegedly more rational materialist alternative world explanation can be counteracted without much toil by unmasking its essential fideism, where faith in God is only replaced by the irrational faith in the designer omnipotence of quite blind and accidentally acting forces[10]. A more modern argument from

---

[7] Plat. Leg. X.899e-900b. But similar reproach of Zeus was expressed still earlier, e.g., in pessimistic verses of Theognis of Megara (approximately the six century BC) who complained that it was against justice to distribute equal fates to righteous and impious people.

[8] Like God of open theists, Īśvara of Indian theists was not viewed as the Omnipotent Being, but in this case not because of desire to deliver him from the burden of responsibility for everything in the world and to bring him nearer (also in the ontological sense) to created beings, but in accordance with the all-powerful faith in the law of karma/samsāra which was thought of as the general mechanism of causality without beginning and independent in itself.

[9] We can find sarcasms toward the pretended "unrighteous lord of the world" who might be incapable to prevent sufferings of living beings or even satisfied with them as early as in the collection of the Jātakas from the Pali canon (VI.208). The same manner of discourse on Īśvara is on hand in Vasubandhu's commentary on his own famous *Abhidharmakośa* (II.63-64) and in Samantabhadra's commentary on the same text (the fourth century A.D.). Related to the topic under discussion is such still not outdated contribution as *Hayes R.P.* Atheism in Buddhist Tradition // Journal of Indian Philosophy, 1988, Vol.16, p. 5-28.

[10] A historical survey of the reductio ad absurdum argument against this naturalistic world-view (from the Stoics up to Fred Hoyle), where what can be designated as short-cut narrative argumentation was implemented is presented inter alia in my recent paper: V.K. Shokhin, "Natural Theology, Philosophical Theology and Illustrative Argumentation",

the very existence of alternative religious traditions allegedly of equal (if not excelling) value is less irrational, but it can also be countered by comparison between the attributes of the theistic God and other "ultimate realities". In contrary, the argument from evil encroaches just on these very attributes by contrasting their ideal perfectness with the robust quantum of negative realities of the empirical world which cannot be explained away by a supposition (popular in Antiquity and Middle Ages) that evil has no rights for existence.

Meanwhile, the other word-combination used in the title of this paper, *critique of religious reason,* calls for amplification, since, in the first place, I use it in the Kantian sense. The question is about the resources of this reason in relation to the subject under discussion. My view is that this reason has always struggled to fulfil the task beyond its capacities but often simultaneously, fails to do its best with one quite commensurate with its capabilities. Time constraints prevent me from delving into details, so I have to confine myself with only a few contours[11].

An unfeasible task which no one has imposed on this reason and in which it has ever been very interested is building theodicies, i.e. universal, generally valid and transparent for all reasonable persons (both believers and nonbelievers) models of explanation of causes, dimensions and distributions of evils and sufferings in the world of the Divine origin and government. From the Late Antiquity up to the Late Modernity four such main models were not only acknowledged but also highly praised: 1) evil as something ontologically unlawful or a kind of non-being in the final analysis, 2) evil as, contrary to this, a necessary element of the world system and its harmony and order, 3) evil as a product of a free and false will of limited conscious beings; 4) evil as an instrument of Divine dispensation aiming at the improvement of human beings.

None of these explanations managed to be sufficient in and of itself, inasmuch as all the most distinguished "theodicists" combined them in their texts without, usually without being aware of it. To begin with, Plotinus, whose influence on developing theodicies in the later, in the first place Christian, natural theology can hardly be overestimated, presents all the four models in a clear-cut fashion. His interpretation of

[11] The ideas contained in this paper were first presented at the 21th Conference of the European Society for Philosophy of Religion in Uppsala, 25-28 August 2016.

evil as either a product of matter which is almost the same as non-being, or only a diminuation of good (Enn. I.8, cf. III.2.5) corresponds to model (1); snickering at those who would like to get rid of evil in the world as similar to ignorant critics rebuking an artist for not using one color, a producer for including also negative characters into his play, citizens for establishing the service of an executioner (III.2.12, 17) is model (2); his view of evil as a result of free choice of the outward world at the expense of divine contemplation and, in addition, of arrogance and self-isolation of a soul (III.5.1, III 9. 3; V 1.1) is (3), while consideration of evil as a means to help humans in increasing vigilance, waking sound reason, withstanding obstacles and realizing how beneficient virtue is in comparison with the disasters which befall wrongdoers (III.2.5) definitely corresponds to (4). With Origen we have pattern (1) when he insists that evil is something nonexistent because everything existent must have been created by God who could not create what contradicted to his nature (In Ioan. II.3.93,99), pattern (3) in explanation of the origin of evil by bad upbringing of human beings, free self-corruption and vicious environment (De princ. II.6.4-6; III.1.3; Contra Cels. III.69; IV.12, 20, 21, 64), but he was also sure (pattern 4) that sufferings were similar to bitter drugs used by doctors for recovering patients (Contra Cels. VI.56; Philok. 27,7). A very similar distribution of explanatory models we find also with Tertullian. For St.Augustine evil was a lawful condition for the proper world-government (pattern 2) and he supplemented Plotinus' similarities by the necessity to combine high-pitched sounds with low-pitched to produce a song and long syllables with short ones to make a poem (De ord. I 1, 7; II 19), but it was he who emphasized more expressively than all his predecessors pattern (3), i.e. that perverted and sinful free will is the main (if not singular) cause of evil, the will that (again in accordance with Plotinus) prefers transient goods to eternal ones (De lib. arb. II 19; Enchrid. 23- 24, De ver. rel. 14; Contr. Fortunat. 15 etc.), while acknowledging (pattern 4) also an "educational" dimension of evil which makes one understand attractiveness of good (Enchrid. 11 etc.). All the four patterns are at hand with Pseudo-Dionisius and Aquinas. And the same is with Leibniz whose *Theodicy* (1710) displays the renown medieval "privative conception" of evil (which is *privatio boni*), corresponding to model (1), still more renown his own conception

of the best of all possible worlds (2), along with detailed reference to (3) and resolute acknowledgement of (4)[12].

In principle, such a "multi-explanation" is not contradictory to reason, inasmuch as evil is a very profound and multidimensional reality, but these explanations have been not too well compatible with each other. Some of them contradicted to all others, i.e. model (1), while model (3) to model (2). One contradicted even to itself, i.e. model (2) which treated evil as a lesser but, in reality, indispensable good. And one of them, i.e. model (3), being in itself very cogent for spiritual reason, led inquisitive minds to dead-end disentanglement of puzzles, e.g., how Lucifer, created as good and lightful, could have fallen, if the very "place" wherein he could have "entered" was just the result of his very fall. Models (1) and (2), which I prefer to call metaphysical, are not in real use now[13], but models (3) and (4) retain their places, along with model (5) that treats suffering in this life as a means to acquire bliss in the afterlife. They are evidently less vulnerable to criticisms than counterintuitive model 1 (we can regard evil as non-existent only when we inflict it on others and by no means when they inflict it on us) and model 2 (evil as a good), but they also do not work as universal explanations. It would be very presumptuous, for example, to interpret all environmental disasters as resulting from by moral evils in the world of reasonable beings, and still more arrogant to explain ruin of persons *A, B* and *C* in a plane crash by their better preparation for the heaven than in case with persons *E, F* or *G* who had not managed to get tickets for the same flight.

All said above does not contradict to the fact that presumptions behind models (3), (4) and (5) conform well to religious reason, and, what is still more important, religious vision. More important inasmuch as inner feelings, intuitions and contemplations (cf. the basic meaning of the term *theōría*, one of the fundamental notions of the European culture[14]) are first-rate sources of knowledge in epistemology of

---

[12] On some non-Christian counterparts of these models (in Islam, Judaism and Hinduism) see, e.g.: V.K. Shokhin, "Philosophical Theology and Indian Versions of Theodicy", *European Journal for Philosophy of Religion*, 2010, Vol. 2/2, pp. 177-200.

[13] Among most notorious expositions of model (2) one could mark: N. Pike, "Hume on Evil", *Philosophical Review* 1963, Vol. 72, pp. 180-197 and R. Chisholm, "The Defeat of Good and Evil", *Proceedings of the American Philosophical Association*, 1968, Vol. 42., pp. 21-38.

[14] One of the best investigation of both the origin and essence of *theōría* (with very valuable remarks also concerning pre-Platonic contexts of the notion) is doubtlessly:

religious beliefs and have more access to the topic under discussion than syllogisms. A man of even average spirituality can discern with his inner eye that many evils in his life are allowed for his profit and that there are some connections between some of his sufferings and sins, while very rare persons of highest spiritual gifts can behold some future rewards for their patience in this life. But any attempts to erect generalized explanatory patterns split against Revelation which asserts definitely *for now we see through a glass, darkly, but then face to face, now I know in part, but then I shall know even as also I am known* (1 Corinthians 13:12), and any self-confident "theodicist" should be, one can believe, so scarcely agreeable to God as were Job's friends rejected by Him (who were also "theodicists" with ready clear-cut explanations of the sufferings of the most righteous man in the world). In reality, Revelation cannot be anything but uniquely authoritative for any "man of the Book", and it does not encourage any explanatory universalism. In some cases, and even numerous, severe diseases and other sufferings are said to be definitely connected with sins, as, e.g., in the narrative about healing of the man sick of the palsy according to the Synoptic Gospels (Matthew 9:1-7, Mark 2:3-12, Luke 5:18-25) and the fourth one (John 5:14), in other cases – as with a woman who had a spirit of infirmity for eighteen years – with Satan's actions (Luke 13:10-16), in other ones – as was the case with the man blind from his birth – with their use for manifestation of the works of God (John 9: 1-3). And God, who nurses a man more diligently than a mother her babe in arms (according to Isaiah 49:15 at least), allows, as in the last two cases, vast volumes of human sufferings for realization of the highest goals. In addition, what is of great significance for our topic, He prohibits any measurement of correlations between human actions and their retribution, as we see from Jesus's discourse on the Galileans whose blood Pilate mingled with their sacrifices and those slain by the tower of Siloam (Luke 13:1-5)[15].

More feasible for religious reason could be the task of counteracting attempts to repudiate the existence of God by evidence from negative

---

G. Picht, "Der Sinn der Unterscheidung von Theorie und Praxis in der griechischen Philosophie", *Idem.. Wahrheit, Vernunft, Verantwortung. Philosophische Studien.* Stuttgart: Ernst Klett, 1969, pp. 108-135.

[15] Cf. very resolute emphasis in Jesus' answer (after resurrection) to the Apostles' question about the future kingdom of Israel: *It is not for you to know the times or the season, which the Father hath put in his own power* (Acts 1:7).

facts (and even their abundance) in the world, which is not a theodicy but a defense. But here theistic achievements in the contemporary controversies are not too impressive. It is true that Alvin Plantinga has managed to coerce into silence the so called famous logical argument from evil offered by John Mackie[16] (in reality very similar to one offered by Pierre Bayle more than three hundred years ago), i.e. from incapability of the Omnipotent Being to create such perfectly free wills which would not be able to carry out unrighteous choices. Plantinga convincingly demonstrated that even Being "than which a greater cannot be thought" (to use Anselmean idiom) cannot produce anything self-contradictory and, therefore, absurd, because it would contradict to its perfection[17]. But after this unsuccessful frontal attack the "friendly", "indifferent", "modest" and other atheists learned how to make more skillful mine holes – from what is called evidential arguments. Many theists have not found anything better than to reply to the renown William Rowe's argument from gratuitous disasters exemplified in the basic version of it by an imaginable example of a fawn perishing in a forest fire[18] by a logical trick known as G.E. Moore's shift of premises, ironically suggested to them by Rowe himself, instead of answering him to the point. Counterarguments from the position of skeptical theism presented by Stephen Wykstra and then elaborated by William Alston[19] were much better, but analogies they adduce are more suitable in cases of more as it were "calm evils" than with "horrendous ones", to use Marylin Adams' words[20]. In addition, some atheists managed to use very theistic reasoning on the Divine Hiddenness to their profit, as we see with John Schellenberg's argument from cases of the so- called human inculpable

---

[16] J.L. Mackie, "Evil and Omnipotence", *Philosophy of Religion: An Anthology of Contemporary Views*. Ed. by M.Stewart. Boston, 1996, pp. 333-344.

[17] A. Plantinga, *The Nature of Necessity*. Oxford, 1974, pp. 164-196.

[18] W. Rowe, "The Problem of Evil and Some Varieties of Atheism", *American Philosophical Quarterly*, 1979, Vol. 16, p. 337; *Idem*. "Evil and Theodicy", *Philosophical Topics*, 1988, Vol. 16, p. 119.

[19] See, e.g.., S. Wykstra, "The Humean Obstacle to Evidential Arguments for Suffering. On Avoiding the Evils of Appearance", *International Journal for Philosophy of Religion*, 1984, Vol.16, pp.73-93; W. Alston, "Some (Temporarily) Final Thoughts on the Evidential Arguments from Evil", *The Evidential Argument from Evil*. Ed. by D. Howard-Snyders. Bloomington, IN, 1996, pp. 311-332.

[20] See already her early manifesto: M.M. Adams, "Horrendous Evils and the Goodness of God", *Proceedings of the Aristotelian Society*, 1989, Vol. 63 (Supplementary), pp. 297-310.

unbelief (which God as a loving father, had He existed, could have removed without problem for the welfare of His children)[21], to which theists usually object using traditional patterns of theodicies mentioned above which are scarcely useful here[22].

Meanwhile, it is not a secret that any defense can be effective only when escorted also by counterattacks on the territory of the opponent, and they are here mostly lacking. To fill some gaps in this regard, I would ask Row's followers to ponder a little bit on why he summoned God to "a Hague tribunal" for sufferings in the animal world[23] and not for horrors of the Gulag, Sino-Japanese war, Holocaust, massacres in Kampuchea, Rwanda or Sudan (I follow the chronological order) not to mention other genocides. May be because innumerable human victims have not been "gratuitous" and, correspondingly, have been justifiable? Or, in other words, by the reason that animal lives are regarded more valuable nowadays than human ones, as was the case with some medieval Jains who cherished bacteria and regarded some Hindus deserving violence? To Paul Draper, the author of the so-called hypothesis of indifference[24], I would put another question, i.e. whether his outright equating of evil with physical sufferings and good, correspondingly, with bodily pleasures does not remind one of extremely oversimplified ethics thrown away on the level of reasoning already in the works of Aristotle? And those for whom his reasonings sound persuasive, I would ask to

---

[21] See: J. L. Schellenberg, Divine Hiddenness and Human Reason, Ithaca. 1993, p. 83; *Idem.* "The Hiddenness Argument Revisited (I)", *Religious Studies*, 2005, Vol. 41 (2), pp. 201–215; *Idem.* "The Epistemology of Modest Atheism", *European Journal for Philosophy of Religion*, 2015, Vol.7/1, pp. 51-69 and numerous other papers of the same author, as well as his responses to opponents.

[22] Schellenberg himself (as well as his followers like Theodore Drange) does not acknowledge his argument as a variety of the argument from evil (because he aspires to opening of the new era of atheism which has to use new weapons instead such old ones), but in reality it doubtlessly is, because he stresses that God's admission of unbelief of some people deprives them of consolation in their hardships and enlightenment and deprivation of a very important good is, certainly, evil.

[23] It was only much later that he decided to use a report from a Detroit newspaper about rape followed by cruel murder of a five-year-old girl "Sue" as an additional piece of evidence against God after the primary "Bambi" case – W. Rowe, "The Evidential Argument from Evil: A Second Look", *The Evidential Argument from Evil*. Ed. by D. Howard-Snyder. Bloomington and Indianapolis: Indiana University Press, 1996, p. 264.

[24] See: P. Draper, "Pain and Pleasure: An Evidential Problem for Theists", *Philosophy of Religion: The Big Questions*. Ed. by E. Stump and M.J. Murray. Malden, 1999, pp. 164-175.

meditate on why it was still in the first decades of the twentieth century (to say nothing about previous ages) that the moral evil (injustice in the first place) predominated attention of those who discussed the topic and it is again only nowadays that physical evil occupies almost the whole horizon of those dealing with it. Is it not again a mark of retrogression of the culture? Schellenberg's attention could be attracted to such a nuance that his calculations of the inculpabilities in human unbelief (and here we have the pivotal member of his syllogisms) are accessible only to the Omniscient Being that knows the hearts of all humans along with their inner histories and whose existence he endeavors to refute just on the ground of these calculations[25]. And does his attempt to deprive God of his parental rights (for admitting unbelief of His pretended children) really disclose problems with the existence of God and not (I again attract attention to the change of "historical seasons") with today's sociocultural patterns of thinking in accordance with which all the required for the good of a child should be required from a parent alone? But even at the level of elementary logic, if unbelief of some people is regarded a sufficient reason for negation of the existence of God, should not belief of others (who are more numerous) be regarded a good argumentum a fortiori for affirmation of His existence?

But here we come to the most profound layer of atheistic mentality which lies at the bottom of all concrete arguments. This is typical wishful thinking dressed up in the objective (if not even scientific) investigation of the correlations between Divine attributes and states of affairs in our world. I mean that we observe manifested double standards or, in other words, false scale in all the discussed solutions of the problem of evil. However serious and dramatic this problem is in itself, adversaries of theism prefer to ignore completely that profusion of gratuitous benefactions (let us use Rowe's language), which is being poured out every hour of the universe's history and still oftener on all sentient beings including themselves and which (if one embarks on calculations like

---

[25] E.g., I'm not sure that Schellenberg himself is capable to determine how far his own "fair unbelief" was caused by his "innocent ignorance" of the Divine existence and how far by his ambitious desire to open a new era in the history of atheism and then promulgate a new "experiential religion" better accommodated to the present stage of the evolution of the mankind (whereas theism and "classical atheism" do correspond to a lower stage of it in his opinion). See: J. Schellenberg, *The Will To Imagine: A Justification of Skeptical Religion,* Ithaca: Cornell University Press, 2009.

those offered by Draper) surpasses volumes of all evils a lot. Not to notice it (as does, for example, Rowe's eminent disciple Alex Trakakis who wonders wherefrom he could infer the goodness of God[26]) is to become like a chief character of one old Russian satirical poem who asserted that the sun is useless because it lights only when there is full day light even without it. Or, to take another analogy, atheistic investigation of evil is mostly similar to investigation of a hated husband's way of life by a wife yearning to divorce from him or of work of a bank-employee by a top manager wishing to dismiss him at any rate[27]. Without realizing that they encounter motivations, theists will continue be sure that they do only with propositions.

But why they themselves do not appeal to abundant gratuitous benefactions, is not too clear for me. One could surmise they may be sure that it is the sun's natural obligation to shine (and carrying out natural obligations does not deserve gratitude). But, surely, it is a topic of quite a separate paper.

---

[26] N. Trakakis, "The Evidential Problem of Evil" : http://www.iep.utm.edu/evil-evi/ (28.09.2016).

[27] To give only one example, Drange, a veteran atheist and faithful follower of Schellenberg acknowledged quite frankly that the argument from unbelief is more effective weapon than earlier ones inasmuch as it strikes the Christian God better, because "there is Biblical evidence that if the God of Christianity were to exist then he would have a great concern about humanity's widespread lack of belief": Th. Drange Th. "The Arguments From Evil and Nonbelief (2006): http://infidels.org/library/modern/theodore_drange/aeanb.html (30.02.2016). Such a hatred and animosity toward a presumably non-existent object (as God for those who deny His existence) is one of the most profound paradoxes of atheistic mentality which in reality is moved much more by feelings than by rationality.

A DEFENSE OF FIRST AND SECOND-ORDER THEISM: THE LIMITS OF EMPIRICAL INQUIRY AND THE RATIONALITY OF RELIGIOUS BELIEF

CHARLES TALIAFERRO & CHRISTOPHE POROT

St. Olaf College   *Harvard University*

Abstract. We argue that the use of the term "supernatural" is problematic in philosophy of religion in general, and in the contribution by Thornhill-Miller and Millican (TMM) in particular. We address the disturbing parallel between Hume's case against the rationality of belief in miracles and his dismissal of reports of racial equality. We do not argue that because Hume was a racist (or an advocate of white superiority) therefore his view against miracles is faulty, but we draw attention to how Hume sets up a framework that, for similar reasons, discounts evidence of black intelligence and divine intelligence (or evidence of acts of God). We go on to argue against TMM's revision of Hume on miracles. We then argue that empirical testing on the veracity of petitionary belief is impossible for there is no control case (everyone is prayed for) and that empirical testing can no more evaluate the evidential merits of most religious experiences than it can assess the merits of any robust philosophical position in epistemology, metaphysics, value theory, logic and mathematics. We express doubts about the integrity and scope of how one might enjoy the good of religion without belief. In a final section we offer a defense of the rationality of believing in specific religious traditions based on religious experience along with what we refer to as *sufficient philosophical reasoning*.

Branden Thornhill-Miller and Peter Millican advance an important argument against what they characterize as first-order supernatual beliefs but they defend a modest form of second-order supernatural beliefs. This allows educated, reasonable persons to adopt a generally teleological (perhaps even theistic) view of the cosmos and to enjoy some of the benefits of religious belief, but without justifying any more specific theistic or other religious tradition. Janusz Salamon offers a fascinating rejoinder that is no mere rejoinder, but an independent

'agatheistic' conception of the nature and structure of religious belief that is fundamentally oriented toward values and what he terms the agathological imagination.

In our contribution to this exchange, we focus on challenging the arguments against first-order supernatural beliefs advanced by Branden Thornill-Miller and Peter Millican (henceforth TMM). Janusz Salamon grants that some of TMM's points are correct, but we challenge these points. Salamon suggests that rejecting some of these points would be more of a rhetorical rather than philosophical move, and we hope to advance philosophical rather than rhetorical arguments in turn.[1] We leave to another occasion our positive response to Salamon's constructive alternative philosophy of religion.

There are six sections that follow. In the first, we offer reasons against TMM's use of the concept and term "supernaturalism." Section two takes issue with Hume's own case against miracles, while section three offers criticism of TMM's revision of Hume. Section four offers reasons why empirical testing of petitionary prayer and religious experience is of very limited use or philosophical significance, just as, in our view, empirical testing cannot settle robust philosophical positions in epistemology, metaphysics, value theory, logic and mathematics. In section five, we express doubts about the integrity and scope of how one might enjoy the good of religion without belief. Section six ends with a defense of the rationality of believing in specific religious traditions based on religious experience when this belief is paired with what we will refer to as *sufficient philosophical reasoning*.

I. A PLEA FOR NOT EMPLOYING THE TERM "SUPERNATURAL" IN ASSESSING RELIGIOUS BELIEFS

Given the importance of TMM's notion of the supernatural – they use the term "supernatural" and "supernaturalism" 118 times – we begin

---

[1] See "Atheism and Agatheism in the Global Ethical Dialogue" by Janusz Salamon, *European Journal for Philosophy of Religion*, Vol. 7, no. 4 (2015), 226, p. 200: "Their restrictive approach to the justificatory grounds of religious belief may disappoint some defenders of the rationality of religious belief, but the vision of advocates of various metaphysical outlooks engaging in intellectually honest and genuinely philosophical – rather than merely rhetorical – debate, involving readiness to admit that there are objective limits to the strengths of the arguments supporting one's own position, may provide a model of a constructive atheism/religion debate."

with observations about the concept of the supernatural.[2]

"Supernaturalism" (and its cognates) is not the standard philosophical term in referencing God or the divine or sacred in philosophical literature. Nor is it common in theology. Perhaps using the term "supernaturalism" might be welcome if the term appears in religious belief and practice, but the term does not appear in much, if any, practicing religions that we know of. For example, in the largest religion in the world, Christianity, the term "supernatural" does not occur in any of its scripture, creeds or rituals. This alone does not provide good reason for not using the term, just as philosophers use the term "dualist" to describe a host of thinkers (including Pythagoras, Plato, Aristotle, Augustine, and Descartes) who never used the term.[3] [4] There is a clear distinction between retroactively applying a term to thinkers who are no longer active and attributing the term to extant thinkers, but we still concede that even if almost none of the religions TMM discuss use the term "supernatural" this alone is not sufficient to find the term suspect. We propose that the use of the term "supernatural" has three drawbacks in philosophical reflection on the credibility of religious belief.

First, the use of the term "supernatural" does not entail, but it suggests that we have a clear understanding of what is natural, and implies that the "supernatural" is in some way beyond nature, not natural or unnatural. How could we know what is clearly supernatural if we lack a clear consensus about the definition of natural? For something to be unnatural, we most likely need to know what constitutes the natural. Many philosophers in classical Judaism, Christianity, Islam, and theistic Hinduism have claimed natural features to their beliefs: these include, but are not limited to, notions of the Divine nature or to the nature of God or Allah or Brahman, as well as natural law or purpose.

There is limited clarity about what counts as "natural." Many philosophers recognize today, we currently lack a clear consensus on

---

[2] TMM use the term "supernaturalism" 19 times, and "supernatural" 99 times: compare this with Salamon's reply paper where "supernaturalism" is used twice and "supernatural" is used 21 times. See: Branden Thornhill-Miller and Peter Millican, "The Common-Core/Diversity Dilemma: Revisions of Humean Thought, New Empirical Research, and the Limits of Rational Religious Belief", *European Journal for Philosophy of Religion*, Vol. 7, no. 1 (2015), pp. 1-49. Subsequent references to this work are footnoted as "TMM" followed by page number.

[3] For an example, consider Nussbaum, M. C., 1984, 'Aristotelian Dualism', *Oxford Studies in Ancient Philosophy*, 2: 197–207.

[4] Incidentally, we think the term "dualism" in these contexts is misleading (also).

what counts as nature or the natural. This is especially borne out by those who observe the impact of current physics on our ordinary concept of the natural or what counts as physical or material. Anthony Kenny, probably the greatest living historian of philosophy, observes:

> At one time it seemed as if a robust and substantive naturalism could be easily stated. This was a conception that thought of the world as being made up of solid, inert, impenetrable and conserved matter – a matter that interacts deterministically and through contact. But twentieth-century physics posited entities and interactions that did not fit the materialist characterization of reality, and which took science far away from a world of solid, inert, massy material atoms.[5]

Adding to Kenny's observation, consider Michel Bitbol's claim: "Material bodies are no longer the basic objects of physics... Ironically, the notion of material body motivated the very research that eventually dissolved it."[6] Tim Crane and D.H. Mellor write: "The 'matter' of modern physics is not at all solid, or inert, or impenetrable, or conserved, and it interacts indeterministically and arguably some times as a distance. Faced with these discoveries, modern materialism's descendants have understandably lost their metaphysical nerve."[7] Galen Strawson and Noam Chomsky similarly lament the lack of a clear understanding of what is material or natural. Chomsky observes that "The notion of the 'physical world' is open and evolving."[8] We are not suggesting that there is a clear consensus on the meaning of God as conceived of in theism, but there seems to be more confidence and consensus in understanding divine attributes in theistic tradition than there is in extant developments of the attributes of the natural world.[9] Bertrand Russell went so far as to propose that our concept of the paranormal is, in some respects, no less bizarre than

---

[5] Anthony Kenny, "Faith in Lions," *The Times Literary Supplement,* Number 5751, June 21st, 2013, 3-4.

[6] Bitbol, Michel. 2007. "A More Radical Critique of Materialism: A Dialogue with Bas van Fraassen About Matter, Empiricism and Transcendentalism." *Images of Empiricism, Essays on Science and Stances, with a Reply from Bas van Fraassen*, Oxford University Press.

[7] Tim Crane and D.H. Mellor, "There is no question of physicalism," in *Contemporary Materialism: A Reader,* ed. Paul K. Moser and J.D Trout (London: Routledge, 1995), 66.

[8] Chomsky, Noam (1980). *Rules and Representations*: New York: Columbia University Press.

[9] See, for example, *The Routledge Companion to Theism* ed by C. Taliaferro, V. Harrison, and S. Goetz (London: Routledge, 2012).

the concept of the physical world as revealed in the physical sciences. "Matter has become as ghostly as anything in a spiritual séance."[10]

Second, related to the first point, the term "supernaturalism" is employed by critics of theism to construe *the supernatural* as in some way a distorted, derived concept of that which is natural and material. Wesley Wildman is representative of this (in our view) hostile description of theism. In *Science and Religious Anthropology*, Wildman refers to theistic (or "supernatural") worldviews as centered on "discarnate intentional beings," "disembodied forms of intentionality," and "disembodied causal powers."[11] Terms like "discarnate" or "disembodied" suggest that some thing (a being or power) lacks some quality and is (possibly) damaged or disfigured. Why not refer to theism as simply upholding the reality of a divine intentional, purposive power or, employing terms that distinguish God from (what we assume to be) material objects, to claim that this intentional, purposive power (God) is *transcendent* or *incorporeal*? "Incorporeal" does not, in our judgment, come with the same pejorative connotation ("baggage") as "disembodied" for, according to classical forms of Christianity (in the Platonic Christian tradition) human persons are themselves incorporeal, albeit fully embodied as a functional unity in this life.[12]

Third, the way in which the "supernatural" as a category is currently defined in English, includes not just God as conceived of in Abrahamic faiths. It also includes ghosts, spooks, vampires, telepathy, astro-projection, witches, Delphic oracles, dead ancestral spirits, poltergeists, and so on. This large set of possibly supernatural entities ultimately is not related to the research referenced by TMM. They write about the benefits of religious belief, but they do not write about the benefits of belief in ghosts and telepathy etc. In order for the deployment of *supernaturalism* to be constructive in their argument, we believe they would need to include evidence that not only theistic beliefs but also more peculiar, supernatural aspirations, provide individual and social benefits (or not). Since neither they, nor we, intend to comment on the full set of supernatural claims, we believe it is imprecise and unhelpful to deploy the term. Further, because the set of what is supernatural includes such

---

[10] Russell, Bertrand (1927). *An Outline of Philosophy*. London: Routledge.

[11] See his *Science and Religious Anthropology* (London: Routledge, 2009).

[12] For an articulation of different conceptions of human and divine nature see *The Asghate Companion to Theological Anthropology* ed by J. Farris and C. Taliaferro (Ashgate 2014).

oddities, there is a tendency to associate what is supernatural with that
which is superstitious.[13] There are several culturally significant examples
of the term "supernatural" being employed as a synonym for "beyond
one's current understanding" or "incomprehensible" until a naturalistic
explanation is provided.[14] Therefore using the term "supernatural" seems
very close to being what in philosophy is called a *persuasive definition,*
a term that employs a (in this case controversial) value judgment.[15]

   We believe that in philosophy of religion, especially, it is important to
use a golden rule of treating the other person's philosophy in the same
way you would like your own to be treated. While TMM express a keen
respect for appreciating what may be the benefits of and (on a certain
level) the rationality of religious belief, we believe that a further step that
might be taken is to use the terms that have been introduced to describe
religious traditions since the very start of philosophy being carried out
in English, terms that do not carry pejorative baggage. The first extensive
writing in English by philosophers on religion was carried out by the
Cambridge Platonists, from whom we get terms like *theism*, *philosophy
of religion*, and even *consciousness*.[16] We commend this heritage.

   As an aside, while we take issue with TMM's use of the term
"supernatural," we complement them for not using terms that Wildman
uses in his naturalistic critique of theism like "a personal deity" and
"anthropomorphic conception of ultimate reality." No theists we know of
(at least theists who are philosophically active) historically or today refer
to the God of Abrahamic tradition as (using the lower case) "a personal
God," albeit classical Christians have and do refer to the Godhead as
constituted by three Persons. And, of course, no theist believes the concept
of God in classical tradition is "anthropomorphic." It is classically held

---

[13] Consider the popular TV show titled "Supernatural" It's 5.7 million viewers see
the title as descriptive of the content: wherein two brothers battle demons, ghouls, and
vampires.

[14] Consider, as an example, Arthur Conan Doyle's "The Hound of Baskerville" within
which the hound is perceived as a supernatural being before Sherlock Holmes conquers
the superstition by solving the case.

[15] Examples of "persuasive definitions include "naïve realism" and "shallow ecology"
or "anthropocentrism" Few persons who defend direct realism in epistemology or
humanism in environmental ethics or value theory would want to refer to themselves
as naïve or shallow.

[16] See Hutton, Sarah, "The Cambridge Platonists", *The Stanford Encyclopedia of
Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.
edu/archives/win2013/entries/cambridge-platonists/>.

that humans are made in the image of God, but not that God is made in the image of humans, as the Feuerbachian tradition claims.[17]

II. WEIGHING THE EVIDENCE WITH DAVID HUME

We commend TMM on weeding through some interpretations of Hume on miracles. One of us has argued in multiple places that there is a significant parallel between Hume's disparaging of reports of miracles and reports of intelligent black Africans.[18] As is widely known, Hume believed in the supremacy of whites over blacks and, despite the widespread availability of evidence of black equality with whites, he preferred reductive explanations of positive reports about blacks' capacities.[19] Here is a famous claim by Hume about what he takes to be the natural inferiority of blacks to whites.

> I am apt to suspect the Negroes and, in general, all of the other species of men (for there are four or five different kinds) to be naturally inferior to the whites. There never was a civilized nation of any other complexion than white, nor even any individual eminent either in action or speculation. No ingenious manufactures amongst them, no arts, no sciences. On the other hand, the most rude and barbarous of the whites, such as the ancient Germans, the present Tartars, have all still something eminent about them, in their valor, form of government, or some other particular. Such a uniform and constant difference could not happen, in so many countries and ages, if nature had not made an original distinction betwixt these breeds of men. Not to mention our colonies, there are Negro slaves dispersed all over Europe, of which none ever discovered any symptoms

---

[17] See Solomon, Robert C. *What Nietzsche Really Said* (Random House, Inc., New York, 2000) pp 87 "Attributing human traits to God, human beings have disowned their own powers and accordingly have lost awareness of how to use them. Human beings have become estranged from themselves. Fauerbach urges that we rediscover our own capacities and reinternalize our projected powers. Until we do so, we will continue to be victims of our own conviction that we ourselves are powerless and utterly dependent."

[18] See "Hume's Racism and his case against the Miraculous" by C. Taliaferro and A. Hendrickson, *Philosophia Christi* 4:2 (2002), 427-441.

[19] See "Hume's Racism and his case against the Miraculous" for details about the evidence Hume had to ignore when it came to the availability of cases of black intelligence. Hume's racism has been widely commented on in multiple places by Richard Popkin. See, for example, Popkin's "Hume's Racism" in *The High Road to Pyrrhonism*, ed. by J.F. Fare (San Diego, 1980). See also "Race and Racism in the works of David Hume" by Eric Morton, *Journal on African Philosophy,* 1999.

of ingenuity, tho' low people without education will start up amongst us, and distinguish themselves in every profession. In Jamaica indeed they talk of one Negro as a man of parts and learning; but 'tis likely he is admired for very slender accomplishments like a parrot, who speaks a few words plainly.[20]

The Jamaican Hume refers to was Francis Williams who graduated from Cambridge University and was widely known as a poet of Latin verse.[21] Note how ready Hume is to provide an alternative account of Williams' apparent skill as an intelligent, thoughtful person that grants him no more skills than a parrot. Hume seems just as ready to dismiss reports of intelligent blacks as to dismiss miracle stories due to an errant imagination or any number of non-congitive factors (e.g. misplaced hopes and fears, cognitive biases, lack of critical judgement ...).

Hume might be accused of merely making a hasty generalization about all blacks based on a small sampling, but we suggest that Hume was so entrenched in his negative view of black lives that he felt free to ignore the massive amount of evidence available to him if he had ventured outside his circle of friends and acquaintances. There were well over ten thousand blacks (most of whom were free) living in London during Hume's lifetime who were engaged in many intellectual activities, including writing (some of whom were employed by Samuel Johnson) and studying.[22] Yet this evidence seemed completely beyond Hume. For example, Phillis Wheatley, a black American poet, was sent to England to write poems in public and otherwise give public demonstrations of black intelligence and feeling. Popkin writes about Hume on Wheatley "Hume was either oblivious or unimpressed. He never changed his view about blacks. Hume's views were quoted over and over again in America by defenders of slavery and opponents of abolition. They would say, "As the eminent philosopher, David Hume ..."[23] We may have here what Popkin himself thinks is an empiricist who is not sufficiently empirical. We believe this supports the possibility that Hume was willing to reject

---

[20] David Hume, "Of National Characters" in T.H Green and T.H Grose (eds) *The Philosophical Works of David Hume, Vol. 3 (*London: Longmans, 1886), p. 252.

[21] MacDermot, T.H. 'From a Jamaica Portfolio – Francis Williams'. Journal of Negro History (April 1917), 147-59.

[22] G. Gerzina, *Blacks in London: Life Before Emancipation* (New Brunswick, NJ: Rutgers University Press, 1995), p. 5. (predictions of black population in London run from 10,000-30,000).

[23] Popkin, "Eighteenth-Century Racism," p. 512.

substantial evidence on thin theoretical grounds instead of giving a full account of that evidence, as it is. We do not claim to be able to read Hume's motives, but Hume's level of confidence that it is unreasonable to recognize black intelligence seems to us to be as high as Hume's level of confidence that it will always be more reasonable to distrust stories of divine intelligence (miracles) and Kant's own assertions that blacks are categorically inferior to whites on a transcendental level. (Kant was open, in principle, to the rationality of believing in miracles, but he had such an entrenched view of the inferior nature of black persons, he had an antipathy to evidence of black equality very similar to Hume's methodological disposition to doubt both reports of black positive capacities and reports of miracles.)[24]

In our view, there is a profound parallel between Hume's lack of sympathy for reports of black intelligence and his lack of sympathy for reported miracles. Thus, in the following narrative, Hume's case of an ostensible miracle that may have sufficient justification lacks any hint at teleology or purpose. The so-called "miracle" seems more like a freak event. There is not the slightest effort to make his case of world-wide total darkness even remotely akin to any miracle narrative.

> For I own, that otherwise, there may possibly be miracles, or violations of the usual course of nature, of such a kind as to admit of proof from human testimony; though, perhaps, it will be impossible to find any such in all the records of history. Thus, suppose, all authors, in all languages, agree, that, from the first of January, 1600, there was a total darkness over the whole earth for eight days: suppose that the tradition of this extraordinary event is still strong and lively among the people: that all travellers, who return from foreign countries, bring us accounts of the same tradition, without the least variation or contradiction: it is evident, that our present philosophers, instead of doubting the fact, ought to receive it as certain, and ought to search for the causes whence it might be derived. The decay, corruption, and dissolution of nature, is an event rendered probable by so many analogies, that any phenomenon, which seems to have a tendency towards that catastrophe, comes within the reach of human testimony, if that testimony be very extensive and uniform.[25]

---

[24] See "The Colour of Reason" by Emmanuel Chukwudi Eze in *Post Colonial African Philosophy* (Blackwells, 1996) pp. 50-83.

[25] David Hume (1711–76). An Enquiry Concerning Human Understanding. in Steven M. Cahn (ed.) "Seven Masterpieces of Philosophy," (London: Routledge 2016) pg. 255.

These considerations give us reason to resort again to a Golden Rule in philosophy. We submit that the above narrative does not have any proper analogy to any religiously significant miracle reports.[26] A more challenging case for Hume or for a Humean would be to imagine the following: Imagine we have some positive reasons for thinking there is an all-powerful, all-good Creator. There are many reports, historically and today, that this Creator God has been revealed experientially across many times and cultures. In one lifetime, a figure who claimed to speak for or as this God incarnate has been believed to be experienced as a resurrected presence, and so on. We will not labor readers with more details, only taking note of how easily narratives can be constructed to offer analogies that invoke ostensible divine teleology rather than rely on Hume's freaky, ostensibly pointless events.[27]

III. DOES A REVISED ACCOUNT OF HUME'S CRITIQUE OF MIRACLES WORK?

TMM accomodate the purpose of miracles, to an extent, by seeing them as confirmations of religious beliefs. They argue that, in a context of multiple belief systems claiming different miracles, it is the incompatibility of the belief systems that generates a good reason to not believe in miracles: "the contradictions between different religious belief systems ... persuasively challenge the rationality of most kinds of supernatural belief."[28] We believe that the incompatibility between belief systems

---

[26] This may be partly due to the fact that the way Hume defines what is a miracle makes no reference to the purpose of a miracle. Hume identifies miracles as brought about by God as a violation of nature. No reference is made to a miracle's function as a revelation or as some sign of divine goodness. Further, it seems any global event that lasted several days could be attested to with more confidence than any localized event in history.

[27] The problem with Hume on miracles is that his very definition of what it is to be a miracle (as a violation of a law of nature) makes no reference to teleology. All the classical cases of miracle stories we have in Biblical and Christian tradition are each constructed to advance some meaningful, intentional, purposive end, such as the display of love and mercy in a healing narrative or they consist of events that authenticate Christ's teaching or they augment a prophecy calling persons to repent and so on. Given Hume's formal definition of a miracle, a miracle could occur if God were to alter the position of an electron when this was not in accord with the relevant laws of physics on a distant planet with no observers and no consequences of any ethical or religious or aesthetic significance. From a religious point of view, to call such an event a miracle would be absurd.

[28] TMM 1.

should not be an overwhelming concern and this can be demonstrated by applying the same dilemma to naturalistic belief systems as well. There are an astonishing number of various naturalist claims that are incompatible – not only the bizarre, such as the loch ness monster, bigfoot, UFOs, what people see in Acid trips, the paranoid claims of one suffering from mental delusion etc – but also more commonplace considerations about, for instance, the different naturalistic accounts of human persons. We make note of this below in observing how competing theories of mind in naturalism can create a problem, if we use TMM standards. If theism must be evaluated in light of all "supernatural claims" and the consequent ambiguities and incoherences between competing claims then why can't naturalism be evaluated in the same light? Our comment about naturalism and the divergence problem resembles Campbell's critique of Hume, as we believe using the divergence dillema to disprove miracles would generate objectionable results when used on disputes between naturalists.[29]

It's important to remember that TMM's revision of Hume's critique of miracles relies not on distinctions between miracles themselves but on incompatibilities between the propositional claims of various religions that miracles are supposed to confirm. As they phrased it, if M1 (Miracle 1) is supposed to confirm R1 (Religion 1) and M2 is supposed to confirm R2, then the conflict between R1 and R2 leads them to undermine each other and, consequently, undermine the truth of both M1 and M2.[30] However, the conflict between the interpretations was not itself explained in great detail. For example, should we distinguish between miracles that support a religious claim and miracles that are *the only* confirmation of a religious claim? Do TMM assume that miracles are the only evidence for these competing religious claims? We believe this problem of interpreting events could apply equally to both naturalism and "supernaturalism."

First, if a miracle is disproven by a conflict in interpretations then the same miracle could undermine itself so long as two religions (or more) both see it as confirming their theories. Imagine that members of three religions all witnessed multiple children, from multiple religions, be saved from a fire that miraculously stopped burning, with no naturalistic explanation available. The members of those religions might each see

---

[29] TMM 11.
[30] TMM 16.

this miracle as evidence that the God they believe in exists and that their religion is true. But it may not be the case that this miracle is evidence that all three religions are true; however, the competing intellectual content and propositional claims of the religions does not undermine the fact that this event happened. Even if each of these three believers wants their religion to be true, and thus has reason to deny the miracle claim of the others, it nonetheless remains possible that two of them are wrong or that all three of them are wrong. Thus, the miracle may be undeniable to multiple faiths while there is a conflict in the interpretation of the miracle. To say that incoherence between doctrines establishes that the miracle did not happen, even when all three faiths could (in principle) agree that it did, is misguided.[31]

Second, if incompatibilities between theories surrounding an event ought to undermine the existence of that specific event, then this would present a very serious problem for naturalists. For instance, there are many naturalists who disagree about what constitutes the human mind, or if the mind even exists. This is especially true when we consider the different naturalistic interpretations of human minds at any given time in history, as TMM do with miracles. They follow Hume in assuming the divergence problem accounts for not only divergent religions during the same time period but, instead, all religons accross time. If a behaviorist, an eliminativist and a naturalistic dualist (like C. J. Ducasse who was a dualist atheist) see a person wince in pain, they (by their lights) have confirmation for very different theories.[32] Both the behaviorist and naturalistic dualist want their theories to be true as well, and thus take the wincing in pain as proof of their theories.[33] If the conflict between T1 (Theory one) and T2 (Theory two) undermines E1 (Event 1) and E2 (Event 2) then the proper conclusion is that no one winced in pain (or appeared to wince in pain) in front of either the behaviorilist or the naturalistic dualist.

It would seem reasonable to have faith in the prospect of determining whether the behaviorlist or naturalistic dualist is correct about their

---

[31] It is true that, for the sake of argument. TMM consider direct experience of miracles to be a plausible source of knowledge. While our example assumes direct experience, the point is not about the encounter with the miracle, it is about how we are to assess competing, incompatible claims.

[32] Ducasse, C. J.: "In Defense of Dualism" in *Dimensions of Mind*, Sydney Hook, ed. (Macmillan, NY 1961).

[33] See Skinner, B. F, 1974. *About Behaviorism*, New York: Vintage.

understanding of the person. The mere incompatibility of their views does not indicate that no one winced in pain in front of either of them. So the diversity dillema, in and of itself, should not be treated as a reason for religious believers to forfeit their faith in miracles.

One reply might be that TMM are discussing miracles that the believer in question has not seen him or herself. Therefore, this distance between the believer and the miracle makes the conflict very different from the conflict between people who share the same data but merely disagree about what it reveals or means. We believe that this reply, rather than support TMMs argument, would succesfully return the question back to whether or not testimonies of miracles ought to be trusted, and under which conditions they can be. This question must remain distinct, we believe, from the incompatible content of religions themselves.

Finally, TMM argue, it is the believer who wants to claim that the other religions are false and that the miracles used to prove it are false – but, since the multiple religions have the same epistemic grounding for their miracles (testimony), the believer would also have to conclude that the miracles related to their own religion are false if they reject the miracles of other religions.[34] As they phrased it, "The point here is not that Christian believers are logically compelled to deny the miracles of rival religions (as the contrary religions argument would suggest), but rather, that these believers will in fact want to deny them."[35] The problem, structurally, is that the conflict that matters is between the non-miraculous content of religions. So it seems perfectly coherent for a believer to say that they don't believe the testimony of a miracle for a religion that generates other false claims. For instance, if one believed (for religious reasons) that all men are created equal they could legitimately reject, as false, the testimony of a miracle for another religion that proposed that the world should be divided into a hierarchy between men (perhaps racially or class divided). It would precisely be the false content, as they understand it, of the other religion that increases their skepticism about the testimony of miracles for the other religion. It is not the testimony itself, but the testimony paired with false conclusions on other matters. Since testimony is based on trust, it is reasonable not to trust the testimony within a tradition that you believe to have propogated other false claims. Thus, it would precisely be the divergence between the religions that makes it coherent

---

[34] TMM 19.
[35] TMM 19.

to deny the miracle claim of another faith while not denying the miracle claim of one's own.[36] We are not advocating that a believer thinks this way, instead we are assuming the condition explained by TMM – that a believer does not want other religions to be true. Since the problem TMM present involves the perspective of the believer, we think it is important to note that, should one be concerned with the falsehood of other religions, then they could rationally reject the miracles of other faiths without rejecting the miracles of their own faith.

Since TMM are keen to rehabilitate and revise the diologue between "supernaturalists" and naturalists, we believe this endeavor will be more fruitfully undertaken if the reasons given for rejecting supernaturalism (again, we prefer to use the term *theism*) are ones that can be universally applied.

IV. CAN ONE EMPIRICALLY ASSESS THE VERACITY OF PETITIONARY PRAYER AND RELIGIOUS EXPERIENCE?

We think it deeply problematic to suppose that empirical inquiry can establish whether or not petitionary prayer is effective and that is, in part, because there are no control groups. There is no one on earth who is not prayed for. In the *Book of Common Prayer*, prayers are offered for all people in their daily life and work, all who are (or have or will) suffer, and so on.[37] Empirical evidence about persons on distant planets would not be a control group, insofar as they are persons and all persons are prayed for.

What about religious experience? According to theism, God is not a biological or social object or chemical or force field that could, in principle, be detected in natural or social sciences. God is believed to be omnipresent; there is no place where God is not. In an authentic case in which God is experienced, would there have to be something that is empirically identifiable when God is not experienced? There might be some empirical work that is relevant. Imagine we have good independent philosophical reasons for believing that God is all good

---

[36] We are aware that this point fails to comment on TMM's point against accepting all miracles as true, but we believe they are right to claim that this would indeed be a peculiar and unlikely reply to the dilemma. (see TMM 19-21)

[37] see Church of England (1662), *The Book of Common Prayer*, London: Everyman's Library (published 1999).

and we have empirical evidence that someone claims to experience God commanding what is evidently vile and ethically abhorrent. In that case, empirical evidence with the help of a thological framework may have some evidential bearing. But in base-line judgements about whether God exists or not, it is hard to see how empirical inquiry can be successful. If the empiricist or scientist claims to have evidence that persons would believe there is a God (or have apparent experiences of God) even if there is no God, this counter-factual would be unintelligible, given theism. According to perfect being theology (Anselmian theism), God exists necessarily; there is no possible world in which God does not exist. Empirical inquiry typically works in the context of contingent matters (when there can be control groups, etc), but God's every being is non-contingent such that if it is granted that it is even possible (metaphysically) that God does not exist, it follows (for most, but not all theists) God necessarily does not exist.[38]

The problem facing empirical inquiry into authenticating the likelihood of God's existence is similar to the problem for psychological explanations of necessary entailment relations (logic or mathematics). The truth of the claims "Necessarily, if proposition P entails Q and P is true, then Q is true" or "Necessarily, 6 is the smallest perfect number" are manifestly not evidentially resting on pschological testing of how actual human beings reason, for even if it were found that the vast majority human of beings thought these were not the case, any person capable of understanding logic and mathematics would have abundant, self-evident reasons for knowing the majority of human beings are wrong. Necessarily 6 is the smallest perfect number because it is the smallest number equal to the sum of its divisors, including 1, but not including itself (6=1+2+3), not because the majority of human beings happen to think this is the case. Note the parallel absurdity of these two claims: "Persons would believe and ostensibly experience God even if there is no God" and "Persons would believe and ostensibly think it necessary that 6 is the smallest perfect number, even if 6 is not the smallest perfect number."[39]

We suggest that the limited use of empirical inquiry in settling main issues in philosophy of religion is not unique to that domain of

[38] See *Contermporary Philosophical Theology* by C. Taliaferro and C. Meister. (London: Routledge, 2016).

[39] See Gottlob Frege, *The Foundations of Arithmetic*. Oxford: Basil Blackwell, 1950).

philosophy, but also concerns deep matters in metaphysics, epistemology, value theory, philosophy of language, and other sub-fields. We suggest that empirical inquiry could not settle the question of the existence of abstract objects, an A and B theory of time, moral realism, the merits of external versus internal epistemology, and so on. Back to philosophy of religion and to summarize: we do not see how empirical study can confirm when a person's sense of God is generated from HADD versus the result of a largely reliable, mature awareness of the reality of a transcendent, divine reality.[40] [41]In general, we note, TMM's research simultaneously claims to comment on issues "beyond nature" (evidenced by employing the term "supernatural") yet treats them as objects which must be found "within nature." If God exists beyond or above nature, as they appear to claim, then it would be difficult to assess how to find God within nature as one among many natural items in this world.

V. ENJOYING THE GOODS OF RELIGION WITHOUT BELIEF

We are skeptical of the idea that one can reasonably appreciate the individual and social benefits of a belief system without seeing any legitimate merit to the content of those beliefs. Doing so sounds eerily similar to the promotion of a fool's paradise or Brave New World. If the reason for taking faith seriously is primarily for the individual/social benefits (despite likely falsehood) then it would seem that faith shouldn't be taken too seriously in the end. Many dystopian works (*The Giver*, *Brave New World* etc.) rely on a common intuition that sacrificing the truth for individual and social gain is objectionable.

We don't assume that TMM necessarily share this intuition, but we would like to know whether or not they do. We believe there is a notable parallel between their interpretation of supernatural claims and Marx' interpretation, insofar as they claim religion, for the most part, is generated out of material (or natural) conditions that are not connected with the truth of the religious claims.[42] These conditions, as they understand it, are more likely to be pyschological and biological

---

[40] For articulation of TMM's use of HADD see TMM 19-21

[41] Our critical stance on empirical inquiry should not be seen as a crticism of phenomenology. See *Contemporary Philosophical Theology*. by C. Taliaferro and C. Meister. (London: Routledge, 2016).

[42] Marx and Engels, *Collected Works*, vol. 5, *The German Ideology* (Progress Publishers, Moscow) 1975, p. 36.

rather than socio-economic.[43] Nonetheless, if supernatural claims are epiphenomenal to natural conditions instead of reflective of possible truths, then sponsoring the promotion of religion (even in an altered version) amounts to preserving something that is most likely an illusion on the grounds of its benefits.

We believe the value placed on the individual and social benefits of religious belief could influence their work in three ways. First, TMM conclude that it is worth finding a sufficiently rational form of religion because of the individual and social benefits.[44,45] We wonder if this is related to a legitimate possibility of religion being true or if it is motivated merely by the individual and social benefits?[46] If it is motivated primarily by the outlined benefits, then we are curious to know whether or not

---

[43] TMM (24) "Further investigations into the biological basis of these experiences has revealed that appropriately identified religious experiences appear to activate a family of neurobiological systems that are also involved in non-religious functions.50 Ingesting entheogens like psilocybin under appropriate conditions can also produce experiences qualitatively indistinguishable from spontaneously occurring religious experiences,51 or from those induced by meditation and prayer.52 And psychometric studies, making extensive use of the standard measurement of religious or mystical experience,53 have shown factor structures supporting the existence of a common core to religious experience in samples of Christian, Muslim, Hindu, and Buddhist populations, from different continents and with varying understandings of the origin and nature of their experiences, from within their different cultural, religious, and linguistic traditions."

[44] TMM (46-47) "From a rational, empirically-informed point of view, there seem to be two plausible ways forward: walking a path either of scepticism or of reconciliation ... But in the meantime, it is not obviously *unreasonable* to base one's religious commitments on this optimistic second-order theistic view, as long as it remains unrefuted and seems to bring substantial psychological and social benefits."

[45] See David M. Wulff, *Psychology of Religion: Classic and Contemporary*, 2nd edn (Oxford, England: John Wiley and Sons, 1997), p. 135 referenced in TMM (45): "Religion may be necessary for ordinary people ... [because some aspects of it] could be replaced only through an extraordinary management effort".

[46] It is worth noting that TMM's apparent understanding of second-order theism, based on Skinner's own near-retirement view as effectively managerial resembles a common misconception of Marx's own account of religion as an opiate. For Marx, religion serves as an essential relief from the psychological alienation of wage-labour and not as a system of social control. But assuming that Skinner's position is adopted by TMM, the second-order theism advocated in their paper descends into a fool's paradise in which the less educated, less rational are herded into cohesivity. When viewed through this particular lense, wiith societal and intellectual leaders seemingly accepting faith as a simple expedience to such cohesion, it is difficult to avoid seeing the dystopic scenarios realised should the societal elite be aware that second-order theism is the only rational option while encouraging the less-educated to believe in the irrational, for managerial reasons.

there is substantial evidence that second order theism would produce the same benefits as first-order theism? If the research referenced predominantly discusses first-order supernatural belief systems, then it is possible that something within them (the tradition, miracles, or immediacy of the divine in experiences that confirm that tradition for the experient) could uniquely contribute to the outlined benefits while second-order theism (lacking those conditions) can not. We are not committed to this being the case, but we feel that it would be difficult to assume that second-order theism will generate the same results solely on the basis that it shares some (but not all) of the qualities of first-order theism. If evidence were to demonstrate that second-order theism does not produce the same results as first-order theism, then we would like to know whether they would encourage first-order beliefs (despite its irrationality in their view) or promote second-order theism (despite lack of social benefits)? Could there be, for instance, a form of first-order beliefs that is more rational than others in their view which would then offer a middle-ground?

Second, we believe the common core interpretation of religion (as TMM present it) runs the risk of overlooking the possibility of living in a "fool's paradise" along with the possibility of one form of first-order supernaturalism being more rational than another. It may be the case that there are multiple religious experiences that share a common feature, but we believe this would not give a grounding for rejecting first order religious claims as much as it would ground a concern for whether or not some are living in a fool's paradise while others are not.

The assumption behind the common core interpretation is that demonstrating a common condition involved, perhaps necessary for, specific kinds of religious experiences would demonstrate that those experiences are primarily caused by that common feature rather than revealing that feature to be a part of a larger causal nexus.[47] This assumption overestimates the role of common core by positing it as the *only or fundamental* source of the experience rather than as a part of the causal chain related to the experience. Before explaining this as it relates to "supernaturalism", I would like to apply the same assumption to a set

---

[47] TMM (28). Here the diversity in detail of afterlife beliefs tends to undermine them all as revelatory of metaphysical truth, while their similarities point, not towards a genuinely supernatural basis, but rather, towards a common natural cause: the human experience of NDEs across all cultures and epochs."

of non-supernatural narratives. For example, feeling in a trusting state towards others (perhaps including the common features of (A) Trust and (B) Oxytocin.[48]

Imagine the following: three different people (Persons A, B and C) are experiencing the sensation of trust. Person A had lost an item that was then returned to her by a total stranger who she then hugged: both the returning of the wallet and the hug itself triggered the release of oxytocin which thus made her feel trusting. Person B, on the other hand, recently learned that his wife of 20 years was NOT having an affair, and then enjoyed an intimate evening with her. The news and the intimacy triggered a release of oxytocin and thus a state of trust. Person C, unlike the others, had no interpersonal reason for trust. He was actually in a room where a chemical triggering the release of oxytocin was sprayed and then he felt more trusting towards those around him. The three narratives undeniably share the common connection of having the release of oxytocin promote a trusting state but they do not have the same basis for why the oxytocin was released: to say they did would require denying the clear distinctions between the narratives responsible for their trusting state.

We believe this thought experiment illustrates that one can espouse a common core to a plurality of experiences, but accuracy requires that we attend to the narrative of the experience that an individual is going through. This is not merely an exension of the point that context (cultural or otherwise) need to be taken into account or even that a proximate cause is compatible with an ultimate cause. Instead, the various claims and religious experiences themselves need to be evaluated. For instance, imagine that Person B was actually deceived and his wife was in fact having an affair on him. We could say that his state of trust was legitimately false, since the thing that induced it was a lie. Therefore, when discussing the lowered activity of the parietal lobe (a common feature of both the meditative state of Tibetan monks and the prayer state of Fransican nuns) the distinction in their theory (one group is praying to a God they believe exists, the other is not) should not be too hastily overlooked.[49] Just as the false trust of person B may mean he is

---

[48] M. Kosfeld, M. Heinrichs, P.J. Zak, U. Fischbacher, E. Fehr, "Oxytocin Increases Trust in Humans", *Nature* 435 (7042), 2005, pp. 673-676.

[49] TMM (25) "One example is that of 'introvertive mystical experience'. Identified as unity devoid of content, or as 'pure consciousness', it arguably must represent a tradition-

living in a fool's paradise, so too a falsely induced religious experience might aptly be called a "fool's paradise." The primary question should not be "is there a common neurological state?" but, rather, "how likely are the claims of the specific religious believers to be true?" By turning our focus away from neurological states in themselves, and towards the propositional content of those going through religious experiences, we could have a more precise diolague about religion that can comment meaningfully on the distinctions between first-order belief systems.

In order for the narrative and content of the experience to be considered in the way we suggest, we must allow for two essential possibilities. First, a common core could point either to (A) a common "supernatural" origin that is interpreted differently by different experients or (B) common conditions involved in very different religious narratives, each of which is capable of reflecting varying degrees of truth. We believe that there is no prima facie reason why consistency of a common feature or condition reveals a natural origin or undermines these two possibilities. Finally, we run the risk of encouraging a fool's paradise when the narrative and content attached to the experient's claims are overlooked.

VI. A DEFENSE OF FIRST ORDER RELIGIOUS EXPERIENCES

Let's imagine TMM having a first-order religious experience. Imagine the following circumstances:

TMM finished presenting their paper on Hume and petitionary prayer in the Senior Common Room at Oriel College. In the course of walking back to Hertford College, both of them started having a novel, excited sense that some great, loving power was becoming apparent to them. Neither of them reported this at first to the other, but when they got back to the college, one of them said: "You know, I might be losing my mind, but I had the oddest feeling returning to College that I want to pray. Why would I feel this way?" "I have no idea," the other replied.

---

transcending 'common core' since it is an emptying experience – wordless, thoughtless, and not constructed by language. 57 Although more work needs to be done on this topic, when Andrew Newberg and colleagues conducted a comparative neuroimaging study on Franciscan nuns praying and Tibetan monks meditating to achieve this state, both groups showed decreased activity in the orientation association area of the parietal lobe. 58 This part of the brain's right hemisphere provides the sense of body and spatial orientation, so when it shuts down, the body seems no longer aware of its boundaries or of space and time, making the self appear to merge with all things."

"But I am feeling that I need to be open to something larger than myself, something great, transcendent, something powerfully real. It's none of that rubbish that Wildman writes about how the God of theism is some kind of disembodied spook. Frankly, it is more like the sense of the numinous that Rudolf Otto went on about." His friend swore: "Jesus Christ!" but then he laughed: "Maybe I actually should take this Jesus stuff seriously. My aunt sent me C.S. Lewis's *Surprised by Joy*."

At that point, a chaplain passed by who said "It is an enchanting evening." TMM both asked at once "Did you say it was an evening to be chanting?" "Why, no," said the chaplain, but he added, "I am on my way to hear a choir chanting the Magnificat. Care to join?" "Why not?" As the words were chanted, TMM found themselves experiencing the words as not merely sonorous but as communicative or revelatory. One reported to the other "I am not sure of a lot of things like the Virgin birth or the incarnation, but I am starting to feel addressed through this chanting by a divine power."

Let's imagine the above took place. We think that a sustained experience of what appears to subjects to be the divine would be reasonable to trust this ostensible divine disclosure if accompanied by *sufficient philosophical reflection*. What would be sufficient philosophical reflection? We think the following would suffice: Imagine the experients believe that theism or belief in a transcendent divine reality is a live philosophical option. In other words, they have considered theism and believe that there are no compelling reasons for thinking it false. They also believe that there is some (not compelling or overwhelming) but some evidence for thinking theism is true. They, finally, do not believe that an alternative, non-theistic worldview is more reasonable than theism. We can add the assumption that the experients are reasonably proficient in philosophical thinking, and thus had cautiously considered the questions above. In this case, we believe that these subjects would be justified in believing and practicing a specific religious tradition. Assuming the experients are philosophically inclined enough to cautiously consider theism and its alternatives before the experience persuades them to believe in theism, we see no objection to their accepting this experience as a form of evidence.

What about the likely existence of other experients who are drawn to different religious traditions based on similar experiences? Perhaps the experients might be led to make appropriate adjustments of their level of confidence in their emerging religious convictions. That is, they may be reluctant to think they *know* their religion is true versus claiming

that they have a justified belief in its truth.[50] But in light of their apparent experiential awareness of the divine, we believe it would not be rational for them to reject these appearances as incredible.

Compare the above case with two undergraduates debating the philosophy of mind. They have both read a book that argues, powerfully, that the self does not exist except as a narrative point of gravity. To believe that they exist is to believe in a fiction. They consider whether to believe that their experience as agents in the world stems from an evolutionarily adaptive, but false belief. Their professor tells them they should not rely on what seems like an experience of themselves as agents, for that could be due to their "vivid imagination, driven by hopes and fears, cognitive biases, a lack of critical judgement, and a delight in" narratives about selves. And yet the students persist in their seeming to be directly aware of themselves as agentive subjects. Would it be reasonable for them (under those cicumstances) to believe that there has to be something wrong with the professor and the claims that they do not exist? We think it would be abundantly reasonable for them to retain the belief in their own realities as selves.

VII. CONCLUSION

The net conclusion we are arguing for – the rationality of first-order theistic beliefs – is based on a negative line of reasoning to the effect that TMM have not given us good reasons for doubting first-order religious experiences. We don't believe that either the divergence dillema, on its own, or the common core dillema, on its own, amount to a persuasive and compelling philosophical case against the rationality of first-order theistic beliefs. Therefore, we don't believe that the pairing of the two problems constitutes a philosophically strong reason to reject first-order theism. The positive proposal in this last section has taken the shape of claiming that if TMM were to have certain experiences and have some prior, philosophically justified beliefs, they would be rational in being observent in a specific religious tradition. We could have constructed a different thought experiment that would (in our view) show TMM to be justified on the basis of relevant experiences to be warranted in

---

[50] We believe this adjustment is not vastly different from the appropriate stance one takes on many other matters: for example, feeling rationally justified in a belief about a purely scientific discovery but being aware that one could be wrong.

becoming practicing Hindus or Buddhists. In evaluating our position, we note our own conviction that in philosophy it is highly rare to have proofs or arguments or positions that are irrefutable or are widely recognized as compelling. Instead, many philosophical positions and arguments might be reasonable and yet it be the case that there are strong reasons for not accepting their conclusions. We commend Gary Gutting's important book, *What Philosphers Know; Case Studies in Recent Analytic Philosophy,* in which nine cases are analyzed when philosophers in the last century thought their positions were indominable and obvious, only to show that they were very far from that.[51] We conclude our own work in this essay, not claiming to know we are right, but in claiming that we believe TMM have not shown that first-order theistic beliefs and experiences are not rationally trustworthy. We further claim, with greater confidence, however, that use of the term "supernatural" in philosophy of religion is neither charitable nor helpful.[52] Finally, we would like to thank TMM for a substantial contribution to Philosophy of Religion.

---

[51] Gary Gutting, *What Philosophers Know; Case Studies in Recent Analytic Philosophy*. Cambridge: Cambridge University Press, 2009.

DIVERSE RELIGIOUS EXPERIENCES
AND FIRST ORDER RELIGIOUS BELIEFS:
A RESPONSE TO BRANDEN THORNHILL-MILLER,
PETER MILLICAN AND JANUSZ SALAMON

S. MARK HEIM

Yale Divinity School

The distinctive element in Branden Thornhill-Miller and Peter Millican's article[1] is their contention that empirical research in cognitive science can reshape the philosophical argument over supernaturalist religious beliefs. With reference to such research, they restate Hume's critique in terms they believe persuasively challenge the rationality of most forms of supernatural belief – a "common-core diversity dilemma." Similar research however leads them also to recognize that supernatural beliefs have significant empirical personal and social benefits, rooted in normal cognitive processes, and that these benefits are unlikely to be widely available in the foreseeable future apart from such beliefs. This leads to a "normal/objective dilemma" which asks how much weight should be accorded these benefits in a rational assessment of the validity of religious beliefs. They propose a path of reconciliation in which the rational acceptance of a "second order" religiosity is combined with a frank rejection of the supernatural claims of first order religions. This alternative is preferable to a strict skepticism because of its presumed ability to maintain the benefits of religious beliefs (while minimizing their costs) and because it is presumed to provide an additional benefit, that of reducing the virulence of both naturalist-supernaturalist debate and inter-religious conflict. This would free energy for a more

---

[1] Branden Thornhill-Miller and Peter Millican, "The Common-Core/Diversity Dilemma: Revisions of Humean Thought, New Empirical Research, and the Limits of Rational Religious Belief", *European Journal for Philosophy of Religion*, Vol. 7, no. 1 (2015), 1-49.

"cooperation-and-humility-enhancing understanding of religious diversity in a tense and precarious globalized age."

My paper responds to this proposal under two main heads. In regard to the common-core diversity dilemma, I argue that consideration of religious experience (as opposed to a particular understanding of miracle claims alone) undermines the skeptical effect of the dilemma. Diversity of religious and soteriological experiences need not disqualify first order religious claims if they are understood not as conflicting causal claims but logically compatible empirical outcomes, i.e. as combining a "common core" supernaturalist belief with religious outcomes in part constituted by evaluative choices. In regard to the normal/objective dilemma, I argue that empirical cognitive research suggests the benefits attribute to religious belief are most closely associated with first order religion, and their proposal faces not only the practical problem of being unacceptable to most religious people but an internal contradiction. Cognitive research is not yet at a point to play the conclusive role TMM assume, and even in its current state it suggests a more complex relation between first order and second order religion than TMM's "model of the moon" allows.

Branden Thornhill-Miller and Peter Millican (hereafter TMM) model a philosophical framework that promises "progress in interreligious dialogue and in the naturalism/supernaturalism debate," a progress to be judged at least in part by the practical standard that it will provide a "more cooperation-and humility-enhancing understanding of religious diversity in a tense and precarious global age."[2] The attainment of that end is a good that could rationally justify countenancing a certain number of questionably rational religious beliefs, should those beliefs themselves be of net practical benefit for individual and social life. Their argument turns on the premise that recent empirical research in the cognitive sciences decisively shifts the ground in traditional debates.

That empirical research buttresses the naturalist's argument in the form of a "common-core/diversity dilemma." Hume's maxim on miracles states that one can be accepted only if its occurrence is more probable than false testimony on the part of those that assert it. TMM say Hume's maxim can be restated with even greater force by reference to the realm of empirical psychology.[3] Given research that indicates a predilection

---

[2] ibid., 2.
[3] ibid., 12-13.

in human cognition toward identifying agents and intentionality, to the point of too-readily crediting invisible agents, the probability of generating false reports specifically about the action of invisible agents is so great that it sets an extremely high bar for miracle verification.[4] Though in theory the cumulative total of miraculous reports from all religious (and non-religious) sources could be weighed against this standard, the apologist for religion faces a difficult dilemma. If a witness contends that a miracle supports the distinctive beliefs of one religion as opposed to others, then cumulative reports become warring data that cancels itself out. If one wishes to appeal to the cumulative data as support for supernatural possibilities generically over against pure naturalism, then one must face the first argument regarding a cognitive predilection biased toward the production of just such phenomena as a class. This is the "Common-core/diversity dilemma" (hereafter CCDD), newly sharpened by cognitive research, that suggests strict skepticism toward religion.

TMM recognize that the empirical research already referenced demonstrates that particularly religion-friendly cognitive faculties are thoroughly normal features of our nature.[5] As such, it is neither realistic nor wise to contemplate their near term eradication. Though naturalists argue the religious expression fostered by these faculties exceeds their epistemologically appropriate domains and exacts real costs in religious conflicts or obscuration, they need to acknowledge that there are significant empirical individual and social benefits of religious adherence that must also be weighed. Thus there may be a rational argument for preserving the benefits of (even irrational) religious belief while avoiding as far as possible its negative effects. This is the "Normal/Objective Dilemma" (hereafter NOD) that is posed to the naturalist: if the psychological causes of religious belief are part of normal mental function and produce various positive outcomes, "should these rationally weigh more heavily with us than objective epistemological considerations would allow?"[6] The force of this dilemma can be made even more pressing if it is advanced on behalf of a "second order" religion, one that abandons the "competing dogmatisms of first-order supernaturalism" and instead falls back on an "undogmatic version of its second-order cousin."[7]

---

[4] ibid., 12-13.
[5] ibid., 37-39.
[6] ibid., 40.
[7] ibid., 46.

Together, these two dilemmas pose a meta-dilemma, requiring some kind of trade off. TMM contend that the CCDD is even more devastating of first order supernatural beliefs than Hume's original arguments. But this is less true when applied to "second order" supernaturalism, a belief directed to an ultimate, guiding power behind the general structures of the world (a source "distant and unknowable") whose existence is functionally irrelevant for the causal understanding of any question science might address. Such a supernaturalism would be least affected by the CCDD and so best situated to enjoy the support of the NOD.[8] That rational case for the human psychological benefits of religion can be supplemented with a "thin" evidential case for the existence of a religious object, based on the fine tuning of the universe argument.[9]

Balancing these elements, TMM see a path toward reconciliation of religion and reason. This path has the benefit over the strictly skeptical one that it holds promise for mitigating both the conflict among religions and the conflict between naturalism and religion. It would do the first because second order religion recognizes all first order religious beliefs are cultural and relative and provide no basis for inter-group contestation. It would do the second because its particular formulation of NOD would carry convincing weight with rationalists. Insofar as acceptance of this path of reconciliation has a reasonable prospect to actually diminish conflict among religions or between rationalists and believers, that specific benefit would be one more rational reason to accept the reconciliation and its approval of religion on a cost/benefit analysis of its psychological impact.

TMM make a thoughtful and engaging case. They are surely correct that cognitive research will be increasingly important to philosophical discussion, prompting rational assessment of specific topics in light of empirical information about how our minds are constrained to think about them. I am less certain than TMM that cognitive research as yet can change the discussion as much as they suggest.[10] On balance, I think that research actually tells against key aspects of their proposal. My comments will focus first on the question of the diversity of religious

---

[8] For description of second order religion, see for instance ibid. 46.

[9] ibid., 47.

[10] See for instance the wide-ranging assessment of work in this area and its implications by Wesley Wildman in Wesley Wildman, "The Significance of the Evolution of Religious Belief and Behavior for Religious Studies and Theology" in *Where God and Science Meet* ed. Patrick McNamara (Praeger, 2006).

experience in relation to the CCDD and second on the relation between first and second order supernaturalist beliefs in light of cognitive research.

RELIGIOUS EXPERIENCES AND THE DIVERSITY DILEMMA

My contribution in many ways supplements Janusz Salamon's response. Despite his appreciation for their dialogical approach and support for their aim, he gives a crisp summary of what he finds lacking in TMM's proposal:

> ... unless we want to engage in an implausible argument which starts with an admission that religion may after all be 'natural' because it does not seem to go away, and end with a recommendation that it should be replaced with a second-order religion which lacks nearly all the relevant characteristics of first-order religions to which billions of people adhere, we have to accept that religious belief has to, above all, shed light on the question of the ultimate meaning of human existence, and this by reference to human values, not merely facts about the physical universe. For this reason, religious belief cannot lack soteriological/eschatological, metanoeitic/transformational, relational/inter-subjective and other existentially relevant aspects, or else it is unlikely to appeal to adherents of first-order religions. [11]

Because their proposal "misconstrues the nature and ground of religious belief," Salamon expects it to find little resonance among religious believers and so to do little to mitigate conflict, a relevant point since TMM base the rationality of the proposal in part on benefits of this sort. [12] He accepts their critique of an evidentialist approach to grounding religion, but suggests a different epistemological basis, one that he believes can achieve what they want while "being a great deal less revisionist than their second order religion and showing that abandoning fundamental beliefs that are central to one's first order religious tradition is not a prerequisite of holding a rational religious belief under the condition of religious pluralism."[13]

---

[11] Janusz Salamon, "Atheism and Agatheism in the Global Ethical Discourse: Reply to Millican and Thornhill-Miller", *European Journal for Philosophy of Religion*, Vol. 7, no. 4 (2015), 226.

[12] ibid., 197.

[13] ibid.

Salamon proposes an alternative epistemological defense of first order religious beliefs, an axiological grounding. He views religion as rooted in a sense of supreme good (*Agatheos*). Religious faith ascribes to ultimate reality the function of being the ground and end of all that is good. This is a "supernatural" belief (by virtue of the transcendent character of its object), based on the teleological and value-laden nature of our self-consciousness, an empirical fact.[14] The phenomenal common core of religion is deployment of this category of good (wider than theism, say, which is one particular way of describing the good and its basis) and the belief in its instantiation in trans-mundane reality.[15] So religious belief is grounded in our value-laden consciousnesses, and our sensibility of a dramatic gap between our current state and the realization of this greater good. Since it is "not possible to derive values solely from the facts about the physical universe," religion cannot be adequately assessed on that basis alone, as TMM's approach tends to do.[16]

Evidential arguments for God's existence, like the fine tuning argument, do not necessarily imply an absolute with the religious qualities of a morally and teleologically supreme good. Salamon views this point as fatal to TMM's project, since it divorces their second level religion from the entire soteriological dimension at the heart of actual religion. He offers to rectify this problem, suggesting that the only way to ground religious belief in an ultimate endowed with agathological attributes is "by reasoning from human axiological consciousness to God as the ultimate good, towards which that consciousness is ultimately directed."[17] Given such grounding, there can be a "justificatory descent" so that (for theists, for instance) particular beliefs such as that God offers

---

[14] Salamon's view on the nature of religious experience can find support in some cognitive research of the type TMM commend, which suggests that it is just such axiological "meaning making" that can be key to certain psychological benefits of religion. See for instance Michael Inzlicht, Alexa M. Tullett, and Marie Good, "The Need to Believe: A Neuroscience Account of Religion as a Motivated Process," *Religion, Brain and Behavior* 1, no. 3 (2011).

[15] Salamon acknowledges there are non-religious ways to define the good and to constitute values, and there are views that contest the very possibility of an answer to the agathonic question. These may be non-religious in rejecting any ultimate source or definition of the good, but they likewise address the existential question of human values and action. In this respect, Salamon's axiological formulation offers a framework for common conversation about the good of the sort TMM desire, involving both religious and non-religious perspectives.

[16] Salamon, "Atheism and Agatheism in the Global Ethical Discourse ...", 201.

[17] ibid., 211.

revelation to rational creatures can be inferred (i.e. if God is source of ultimate good, it is reasonable to expect God might have these qualities). Salamon suggests that a "ladder "of justification descends from such ultimate agathological beliefs to these more particular beliefs in first order religion.[18]

Religious belief systems are expressions of different visions of what their adherents consider to be the optimal ways of conceiving human potentialities vis-a-vis the ultimate reality as the ultimate good toward which their existence is directed. [19] They are of a piece with a search for individual transformation (toward greater realization of the personal good), an extension of that good into social and historical relations (including those with nature), and a hope for collective and universal fulfillment of the good (eschatology). This existential import is essential to religion. Like moral beliefs, religious ones are formed in connection with thinking about human good, and in this sphere "nothing more than agathological certainty, plus coherence of one's worldview, may be expected and demanded."[20] Choosing a religious option that "identifies the ultimate good with the Absolute religiously conceived ... may be as rational a choice as any."[21] From Salamon's perspective, this means that the naturalist and the religious believer are placed epistemically "on par" when the appropriate frame of reference of each is considered. They are "on par" with regard to "the rationality of their worldviews to the extent all worldviews contain a central component that has an axiological and teleological nature, and as such gives rise to questions regarding subjectively relevant meaning and conduct of human life which cannot be settled by natural science." [22] Salamon argues that agatheism can provide many of the advantages TMM see in second order religion (because it is less invested in evidence of supernatural action in the physical realm), while attracting support from believers because it acknowledges their existential concerns.[23]

---

[18] What rational religion of this sort rules out as irrational would be beliefs not rooted in a notion of supreme good (so ungrounded in that way) or inconsistent with the type of supreme good axiologically assumed (so incoherent). It is possible to argue about which among competing versions of supreme good are most encompassing or consistent, but impossible to reach an absolute conclusion based on reason alone.

[19] Salamon, "Atheism and Agatheism in the Global Ethical Discourse...", 203.

[20] ibid., 207.

[21] ibid., 204.

[22] ibid., 205.

[23] ibid., 204.

Salamon defends the rationality of some first order religious beliefs, a class TMM would like to defuse in its entirety. These beliefs are particular (God speaks to us) rather than general (there is a power behind the universe). They also have an unconditioned or immediate character, whose legitimacy Salamon likewise defends. Such beliefs could neither be generated or widely received except by being recognized by their adherents as the "optimal way of conceiving the nature of the Absolute and its relation to the world and humanity," an approximation to a God's eye view, worthy of unconditional devotion.[24] Any believer's beliefs will likely reflect the existential conditions they have experienced, but this does not contradict "each believer's conviction that his belief is an optimal expression of truth about God *available to him,* because a believer has no other option but to rely on his *present* agathological intuitions regarding the nature of God and God's relation to the world and humanity."[25]

However, though the object of this belief is supernatural and the quality of the belief has an unconditioned character, the content of the particular beliefs is in fact variable, because the agathological imagination with which people construe or recognize their supreme good is itself shaped by on-going experiences. This can only result in an historical process in which the transformative experiences of believers may reframe the imaginations with which they perceive the ultimate good and infer support for particular beliefs. This accounts for the diversity among and within religious traditions. [26] The certainty specific to religious beliefs may be perceived subjectively and shared inter-subjectively, since it applies to the axiological and soteriological worlds and to the category of supreme good. But it need not be confused with an objective certainty in regard to the "middle range" realities of the physical and scientific worlds. If first order religious beliefs are understood in this way, Salamon suggests, they are not bound to be a breeding ground for irrationality or conflict.

We are now in a position to see Salamon's response to the common-core/diversity dilemma. As he sees it the common core of religions is the agathonic task. Insofar as religious experiences evidence "common core" characteristics, they can be taken as supporting the agathonic

---

[24] ibid., 232.
[25] ibid., 233.
[26] ibid., 231-32.

project and its presumption of an ultimate source of the good. Insofar as various religious or mystical experiences explicitly confirm or conform to different first order religious convictions, they do not cancel out their "common core" evidential value, but point to the importance of the characterizations of the supreme good brought to these experiences. Fundamental agathetic belief may "constitute the epistemic foundation of a number of different religious belief systems."[27] Religious diversity stems from the different ways the nature of supreme good can be framed and particular beliefs can be inferred or organized in relation to it. Agathological imagination will play a crucial role in choosing between religious and non-religious options as well as among different types of religious goods. Diversity of religious systems is a space of exercise of agathological imagination, a dimension of the faculty of practical reason directed towards the ultimate good (in a transcendental Kantian sense)[28] (202) There is a range of religious "landscapes" that have been "conceived throughout human history by geniuses of agathological imagination."[29] (204) This means that belonging to a first order religious tradition, and affirming its particular beliefs, insofar as they are derived from the sense of supreme good in that tradition, is consistent "with adhering to the fundamental agatheistic belief, despite there being a *plurality* of such *evolving* religious traditions."[30] In short, Salamon concludes, "we have a 'common core' and diversity, but no dilemma."[31]

I want to reinforce this point. TMM do not directly address the religious experiences Salamon views as central. Their paradigm case of is that of miracles, similar allegedly supernatural events that religious reports attribute to mutually exclusive supernatural causes or associate with incompatible beliefs. To the extent this miracle shows one supernatural cause is real, that miracle shows a competing cause or explanation is real. TMM show that, logically speaking, divergent miracle reports could still lead to a rational presumption in favor of supernaturalist possibilities over pure naturalism. But this move offers no support for one particularist religious view against another, and in fact could as well be seen as supporting polytheism. It requires believers to appeal

---

[27] ibid., 231.
[28] ibid., 202.
[29] ibid., 204.
[30] ibid., 234.
[31] ibid., 245.

to the support of "common core" evidence, which then places them in line for that horn of the CCDD.[32] Whether or not this is an adequate treatment of miracles, it is severely lacking in regard to the soteriological religious experiences Salamon stresses.[33] It is not clear that diversity here can be treated the same way. Different conceptions of a divine good agree on the categorical nature and the ultimacy of their object. They are *consistent* testimony in a way that Hume's characterization of miracle reports (as claims that attribute identical *causality* for identical events to mutually exclusive agents) denies they can be.

I argue that the diversity aspect of the CCDD is misconceived in its assumption that a variety of concrete religious experiences necessarily imply skepticism about the differing particulars. I suggest that religious experiences stand in relation with different types of actual religious fulfilment, states that may have both an eschatological, perfected form and anticipatory, historical expressions.[34] Varying states of religious fulfillment may have factors in common – absence of suffering, for instance. But they also have distinguishing characteristics. For a Christian these might be an experience of personal communion with God and Christ; for a native American these might be a permanent unity and harmony with ancestors and a specific landscape. For a Buddhist these might involve a realization of emptiness. There may be key elements in each of these that are fully incompatible with or unnecessary for the others. But there is no contradiction in affirming that several such fulfilments or ends are actually achievable, and experienced in their particularity by adherents. The "conflict" is not logical but existential,

---

[32] This dilemma is especially telling as directed against arguments that limit themselves exclusively to "common core" evidence. An example would be John Hick's argument in support of religious belief, since he regards only the common elements of religion as of epistemic value. Supporters of Hick's view would respond to the CCDD by contesting the force of the naturalist explanations of common elements, but would essentially concede that diversity yields contradiction. See John Hick, *An Interpretation of Religion: Human Responses to the Transcendent* (Basingstoke: Macmillan, 1989).

[33] TMM simply extend the miracle paradigm to all religious experience. Thornhill-Miller and Millican, "The Common-Core/Diversity Dilemma...", 19.

[34] I have set out this view at much greater length elsewhere. Here I want to stress only ways in which it runs strongly parallel with Salamon's argument with regard to religious diversity. My discussion of multiple religious fulfillments corresponds in large measure to his treatment of different agathonic goods, and my emphasis on the evaluative dimension of religious faith corresponds in large part to his description of agathonological imagination. See S. Mark Heim, *Salvations: Truth and Difference in Religion*, Faith Meets Faith (Maryknoll, N.Y.: Orbis Books, 1995).

the impossibility of some conditions being true at the same time for the same person. Even where the states or experiences reported in different religious cases are so distinct that they could not possibly be realized by the same people at the same time, they do not become contradictory – in Hume's evidential sense – unless we assume that only one kind of human religious fulfillment is possible in relation to the divine, which may interact with humanity only to one effect. That is itself a religious doctrine, open to debate, and neither a logical or empirical given. That is no indication that the particulars of distinct religious realizations are not real or that they have not been achieved in relation to the same object.[35]

In this respect, the claimed realization of concretely different religious fulfillments are mutually supportive in pointing to non-naturalist possibilities but much less vulnerable to the common core horn of CCDD. TMM dismiss consideration of such "polytheism" primarily by assuming that religious believers would refuse to compromise their exclusive claims with such an outlook. But this seems questionable. This approach is actually embodied wherever forms of multiple religious practice occur (Buddhist-Confucian, Muslim-Christian, Jewish-Buddhist), since the premise of such practice must be that there are features of fulfillment in one path not available in another. When a Buddhist expresses the belief that faithful Christians will go to heaven precisely as they expect, but that this heaven will be a period of reward in a continuing cycle of rebirths, the factual conflicts about empirical expectations are much smaller than the evaluative divergences.[36] Though I cannot elaborate further in this setting, the primary point is that diversity *per se* need not have the implications that Hume and TMM presume, and thus cannot play its appointed role in the CCDD.[37]

---

[35] In their own internal reflections, religions recognize this of different valid spiritual paths, not all of which can attain the same ends, and in many cases extend a similar kind of analysis to cases within other religions. This is the case for instance with Buddhist views of different vehicles or skillful means, and Christian ideas of natural knowledge of God or progressive revelation.

[36] Such is the stated view of the Dalai Lama, for instance. See Dalai Lama, "The Bodhgaya Interviews," in *Christianity through Non-Christian Eyes*, ed. Paul J. Griffiths (Maryknoll, NY: Orbis Books, 1990), 169.

[37] It is true that different religious fulfilments may presume metaphysical conditions incompatible with those presumed in another fulfilment (existence of God, illusion of self, etc.), and in these respects the religions pose alternative accounts. How to explain the others from the perspective of one of these alternatives is a theological question. Here the point is that that disagreement does not necessarily extend to the concrete realities

TMM's CCDD can be challenged on several accounts. Its claim that there is no rational justification for first order religious beliefs to compete with naturalistic critiques is countered by Salamon's axiological grounding. Its account of religious experience does not attend to the soteriological themes central to religion. Its formulation of the diversity horn of the dilemma overlooks the way in which distinctly different religious experiences can be credited and not cancelled. For these reasons, the CCDD is not able to carry the weight that TMM attribute to it. We turn now to the normal/objective dilemma.

THE LIMITATIONS OF SECOND ORDER RELIGION

As TMM observe, Hume is not clear on how to resolve the impasse between the strict rationality required for judgements of truth and the irrational instinct necessary for actual life in the absence of certainty, what we might call Hume's dilemma.[38] Their essay intends to be responsive to this concern. Salamon suggests that what Hume treats as instincts (and so as irrational) can be known in light of research to be "in fact facets of the proper functioning of our complex cognitive faculties that ultimately aim at truth and thus are not irrational, despite the fact that we are not able to establish in an internal fashion whether and to what degree the beliefs produced in such a way are warranted."[39]

TMM recognize the epistemic naturalness of our religion-forming capacities, primarily with reference to the difficulty this presents in implementing a strict rationalism. Of course skeptics have long noted, as Hume so pithily does, that humans have a ready love of 'surprize and wonder.'[40] But it has equally long been an assumption that there was or could be an inverse relation between the qualities of mind that incline in that direction (superstition, to Hume) and rational qualities of mind. TMM do make a major change in the naturalist's case by explicitly dropping the contention that religious belief is *per se* a mental pathology, either in genesis (caused by some kind of deformity in our mental equipment) or in practice (marking its adherents as necessarily deficient

---

of the experienced soteriological fulfillments, which can be viewed as valid from varying perspectives on the ultimate.

[38] Thornhill-Miller and Millican, "The Common-Core/Diversity Dilemma...", 5.

[39] Janusz Salamon, "Atheism and Agatheism in the Global Ethical Discourse...", 206.

[40] Branden Thornhill-Miller and Peter Millican, "The Common-Core/Diversity Dilemma..." ibid., 13.

in human accomplishments or satisfactions). Naturalists have charged believers with being "cognitively challenged" because of a mental bias toward supernatural ideas. The charge is sometimes reversed, and nonbelievers charged with mental deficits in the areas of theory of mind or empathy. TMM note that the data do not seem to support mental disability in either group, and report that they incline toward the view that religious belief can be partly explained by a "*preference* for an intuitive, as opposed to analytical cognitive *style*" rather than differences in ability.[41]

This is a suggestive observation, that bears on the assumption that rationality and its qualities of mind, as opposed to religious faith, and its qualities of mind, can vary in inverse proportion almost without limit. One of the striking things about the research in question is its increasingly detailed picture of the intertwined nature of the qualities of mind we are discussing. It was common, for instance, for some naturalists to argue for an inverse proportion in the exercise of emotion and reason. But it seems increasingly clear that emotion is an integral element in the way "higher" human reason works. As Anthony Damasio writes, though emotions can distort our reason, " the *absence* of emotion and feeling is no less damaging, no less capable of compromising the rationality that makes us distinctively human and allows us to decide in consonance with a sense of personal future, social convention and moral principle."[42] Damasio's work stressed particularly the importance of emotion as a shorthand for collapsing numerous steps of reasoning or an illuminator to focus rational processes in fruitful areas.

TMM do not focus primarily on emotion but on cognitive capacities such as the theory of mind (our ability to attribute mental states to other beings) and our hypersensitivity for pattern or agent detection. These qualities of mind are so essential and valuable for life in a social world that we are primed to deploy them in any and all circumstances, and so to posit invisible, intentional agents like ghosts or gods. They figure in TMM's discussion of religion entirely as a source of bias. In a footnote, TMM acknowledge that the Normal/Objective Dilemma could equally well be stated another way than they do. It could be stated as involving a "choice between being more humanly 'normal' (by being irrational or biased in some respect), and being more 'objective,' 'rational,' or

---

[41] ibid., 38.

[42] Antonio R. Damasio, *Descartes' Error : Emotion, Reason, and the Human Brain* (New York: G.P. Putnam, 1994), xii.

'unbiased' (and thus more humanly 'abnormal' in this same respect), in contexts where each choice is likely to entail some unrecognized costs and/or benefits."[43] They equate 'normal' with irrationality or bias because they are speaking particularly of religious beliefs, which they regard as irrational in evidential terms. But if we focus specifically on the cognitive capacities that are the actual subject of the research, the capacities through which religion arises or registers, the matter is somewhat different. In regard to these faculties themselves, the "dilemma" would be much more straightforward: whether to be humanly normal (in full possession of these faculties) or be humanly abnormal (lacking these faculties in some measure). To be lacking in emotion and/or the religion-susceptible cognitive faculties would be to be unable to function effectively or rationally in much of human life.

The cognitive components of a healthy mind, which constitute the distinctive human intelligence with its aptitude for living in social groups of conscious beings, are the same ones that are "religion forming" in the sense of producing first order religious beliefs. Unless scientific thinking is to be carried out by other kinds of minds, to be effectively rational is to be religion-susceptible.[44] Rational thinking itself makes use of these same components to some extent, as is the case with regard to emotion. For people with brains like ours, it is hard to see how we could arrive at dramatically counterintuitive pictures of the world *apart* from entanglement with cognitive capacities like agency attribution, pattern recognition and theory of mind. Surely considerations like this are relevant to Hume's dilemma and what TMM call our "epistemological duty."

In his book, *Why Religion is Natural and Science is Not,* Robert N. McCauley picks up our topic at exactly this point.[45] There is, he says, no "religion department" in our brains, but instead there is a suite of cognitive dispositions which together almost inevitably give rise to first order religious beliefs, though they each have other key functions as well. We will limit ourselves to two that TMM discuss: theory of mind and

---

[43] Thornhill-Miller and Millican, "The Common-Core/Diversity Dilemma ...", 40.

[44] Here the areas of machine intelligence and artificial intelligence are relevant. Is it possible to "outsource" fully rational thinking to entities that operate without some of these cognitive processes, or can artificial intelligence itself be built or learned only with some approximation of these same faculties?

[45] Robert N. McCauley, *Why Religion Is Natural and Science Is Not* (New York: Oxford University Press, 2011).

agency detection. McCauley says it is important from the beginning to distinguish different types of cognitive behavior. Two of these are what he calls "maturationally natural" for human beings. We automatically become adept at them, like speaking a language, rather than adept through training, like riding a bike. The first, which we might call common sense/nature, equips us with a naïve physics of the non-sentient world, immediate intuitions about gravity and the behavior of bodies, for instance. The second, common sense/social, equips us with similar rapid cognitive processes for life in a world of other sentient beings, including faculties such as theory of mind and agent detection.[46]

These faculties put our perceptions and attention in rapport with the way the world actually works to an extent that individual learning could not. In many respects they are simply highly condensed means to lead us to mental or behavioral conclusions that could equally but ineffectually be produced through laborious secondary reflection, given enough time and the accumulation of cultural background knowledge. Regarding this aspect of these cognitive faculties, we may say that the "dilemma" is in fact the simple one of being normally human and normally rational or being humanly abnormal and subrational.

However, these faculties do have innate bias, as simple optical illusions witness in the field of common sense/nature and non-existent "things that go bump in the night" witness in the field of common sense/social. Cognitive faculties like theory of mind and agency detection incline us to err on the side of false positives, and make the acceptance of invisible, intentional agents an inviting rather than a thorny path. Both in the areas of common sense and popular religion, our cognitive faculties inevitably carry us into convictions upon which reason cannot pronounce or which are found contrary to reflective reason. In these cases – which of course are not functionally or sensibly distinct to the minds involved – the dilemma looks more as TMM state it, a choice between normal but possibly irrational function and abnormal but perhaps more rational function, each with attendant practical costs and benefits.

To these two types of maturationally natural cognitive processing, McCauley adds two other comparatively unnatural ones, theology and science, which he calls "reflective" modes of cognitive processing.[47]

---

[46] ibid., 231 ff. The development of maturationally natural systems is discussed in Chapter Two.

[47] ibid.

Natural modes of cognitive processing arise without conscious effort, are easily triggered, operate with great speed and efficiency, and yield results we easily assimilate. Reflective ones, while they may in unusual cases become so well established in some individuals as to seem almost intuitive for them, are laboriously acquired, require support, operate more slowly and deliberately, and often yield counterintuitive results. They are in these respects humanly abnormal. These are rational in ways better suited to uncover truths in the counterintuitive phase space outside that covered by our natural cognitive modes of processing, which are rational in ways better suited to uncover truth in the intuitive middle world of ordinary life. McCauley concludes that first order supernaturalism is most strongly grounded in our cognitive structures, while theology (second order religion) is a much weaker and counterintuitive enterprise. The same relationship obtains between the fragile intellectual endeavor of science and its more cognitively robust partner common sense.

What McCauley's picture brings home to us is the fact that just as there is no religion department in the brain, there is no reason or science department. Like religion, these two use a suite of varied cognitive processing modes, including ones that religion also uses. Many argue that religion is not the primary "target" of any of these processing modes. McCauley reminds us that as an empirical matter, nothing could be further from a "target" of any of the modes than science.[48] The extraordinary achievements of science stem from its ability to selectively, strategically resist and circumvent these natural cognitive inclinations.[49]

McCauley's analysis suggests a strong asymmetry. No naturalist can dispense with the operative rationality in the cognitive modes relating to theory of mind and agent detection, not in their personal mental development for life in human society, not in a career within science as a social enterprise, nor in many branches of explicit scientific practice itself (archaeology, the search for extra-terrestrial life, psychology).[50] But people can, and do, readily dispense with second order reflection,

---

[48] ibid., 102.

[49] ibid., 116.

[50] We could view many arguments from advocates of intelligent design as failing as evidential arguments, but successfully illustrating this point. It is not that one can prove there is an intentional agent behind particular phenomena, so much as it is that scientific inquiry presumes an as yet undiscovered intelligibility. While numberless facts may lie in the natural world unknown to any human being, the only residence we can imagine for an existing but still invisible intelligibility of those facts is other minds. We cannot

whether theology or science. In fact, McCauley stresses, scientists require care to avoid falling "back" into terminology or behavior shaped by these natural cognitive modes of processing. And in fact, when not fully "on the job" they frequently do exactly that, as theologians fall from theological correctness into the language of first order beliefs.[51] This is observation does not minimize the special power of such hard-won scientific reflection and its critical edge in regard to religion. It is meant to be clear about what we can expect from each of these types of mental engagement.

McCauley's reading of empirical research suggests that the benefits of religion which figure so centrally in TMM's argument attach directly to the cognitively robust first order beliefs, and not to the much weaker, personally episodic secondary reflection they would like to see supersede them. Their proposal to "rationally remove all the overlapping fingers [pointing at the moon] associated with our different religions" would remove all the first order beliefs, and so also the benefits that were the premise for their particular approach to the discussion of religion.[52] This conclusion seems to follow from the very type of empirical research to which they appeal, suggesting a naturalist program to harvest religious benefits will need to give more attention to "harnessing" the types of beliefs associated with those benefits, rather than abolishing them.

Religion is an extremely complex phenomenon, a fact we can now correlate with the spectrum of different cognitive processes implicated in it. The level of sophistication needed in a thesis like TMM's can be suggested if we look at the complexity in a comparable but simpler issue, that of the placebo effect. Recent research in this area indicates that it is possible in some conditions to benefit from the knowing administration of a placebo. [53] For instance, in one experiment patients , using a blister pack of pills they know to be a random selection of placebos and drug doses, received equal benefit from both. In a yet more fascinating experiment, patients treated with a pain-killing drug and then moved

---

successfully investigate such intelligibility without activating cognitive equipment that can implicate those attributions.

[51] TMM briefly note research to this effect, Thornhill-Miller and Millican, "The Common-Core/Diversity Dilemma...", 31, note 79. For more extensive review of such research, see McCauley, *Why Religion Is Natural and Science Is Not*, 128-33.

[52] Thornhill-Miller and Millican, "The Common-Core/Diversity Dilemma ...", 49.

[53] Information in this paragraph is taken from Jo Marchant, "Placebos: Honest Fakery," *Nature* 535, no. 7611 (2016).

to a placebo experienced pain relief mediated through the particular physiological pathways activated by the prior drug. One consistent element I see in this research is that the placebo effect depends upon the first order belief that one is being treated with an effective agent. Once that belief is elicited, it can to some extent be rationally manipulated at a second order level. But success in that respect requires priming by and contact with that first order belief. If the criticisms I have raised in this article are sound, then discussion of the connection of religion and religion's benefits to first order beliefs needs to proceed with a similar sensitivity to that integral relation.

I have focused on the particular, first order, contrasting religious beliefs that TMM wish to abandon. I have argued (with Salamon) that some such beliefs can be rationally grounded "from above" on axiological grounds. I have argued that on the basis of cognitive research these supernaturalist beliefs are the most robust form of religion that arises "from below" in our mental process and so are integral to any benefits from religion that are to be rationally assessed. I have also argued that the diversity in concrete religious experience does not devalue soteriological testimony supporting multiple religious fulfilments. All of these points run counter to TMM's thesis as currently stated. My discussion of the specific example of the placebo effect suggests that TMM could strengthen their thesis with greater attention to the empirically objective importance of first order beliefs.

BIBLIOGRAPHY

Damasio, Antonio R. *Descartes' Error : Emotion, Reason, and the Human Brain*. New York: G.P. Putnam, 1994.

Heim, S. Mark. *Salvations: Truth and Difference in Religion*. Faith Meets Faith. Maryknoll, N.Y.: Orbis Books, 1995.

Hick, John. *An Interpretation of Religion: Human Responses to the Transcendent*. Basingstoke: Macmillan, 1989.

Inzlicht, Michael, Alexa M. Tullett, and Marie Good. "The Need to Believe: A Neuroscience Account of Religion as a Motivated Process." *Religion, Brain and Behavior* 1, no. 3 (October 1, 2011 2011): 192-212.

Lama, Dalai. "The Bodhgaya Interviews." In *Christianity through Non-Christian Eyes*, edited by Paul J. Griffiths. Maryknoll, NY: Orbis Books, 1990.

Marchant, Jo. "Placebos: Honest Fakery." *Nature* 535, no. 7611 (2016/07/14/ 2016): S14-S15.

McCauley, Robert N. *Why Religion Is Natural and Science Is Not*. New York: Oxford University Press, 2011.

Salamon, Janusz. "Atheism and Agatheism in the Global Ethical Discourse: Reply to Millican and Thornhill-Miller ". *European Journal for Philosophy of Religion*, Vol. 7, no. 4 (2015), 197-245.

Thornhill-Miller, Branden, and Peter Millican. "The Common-Core/Diversity Dilemma, Revisions of Humean Thought, New Empirical Research, and the Limits of Rational Religious Belief." *European Journal for Philosophy of Religion*, Vol. 7, no. 1 (2015), 1-49.

Wildman, Wesley. "The Significance of the Evolution of Religious Belief and Behavior for Religious Studies and Theology ". Chap. 11 In *Where God and Science Meet* edited by Patrick McNamara, 227-72: Praeger, 2006.

ON SECOND-ORDER RELIGION,
AGATHEISM AND NATURALISM

A REPLY TO BRANDEN THORNHILL-MILLER,
PETER MILLICAN AND JANUSZ SALAMON

GRAHAM OPPY

Monash University

These comments, on the paper by Branden Thornhill-Miller and Peter Millican[1] and on the critique of that paper by Janusz Salamon[2], divide into four sections. In the first two sections, I briefly sketch some of the major themes from the paper by Thornhill-Miller and Millican, and then from the critique by Salamon. In the final two sections, I provide some critical thoughts on Salamon's objections to Thornhill-Miller and Millican, and then on the leading claims made by Thornhill-Miller and Millican. I find much to commend, but also some things to dispute, in both papers. As is so often the way, I shall focus on areas of disagreement.

I.

Thornhill-Miller and Millican argue that rationality requires a retreat from 'first-order religion'. Their argument has two main prongs: (a) 'The Common Core/Diversity Dilemma'; and (b) 'The Normal/Objective Dilemma'. (2-5)

The Common Core/Diversity Dilemma has two horns: (A) in so far as religious phenomena point towards specific aspects of particular religions, their diversity and mutual opposition undermines their

---

evidential force; and (B) in so far as such phenomena involve a 'common core' of similarity, they point towards a proximate common cause for these phenomena that is natural rather than supernatural. (20)

Thornhill-Miller and Millican argue that (A) is supported by, for example, considerations about medical miracles and intercessory prayer (21-3); and that (B) is supported by psychometric studies and considerations concerning near death experiences, meditative and introvertive religious experiences, hypersensitive agency detection devices, and theories of mind (23-31). Moreover, Thornhill-Miller and Millican also argue that considerations about egocentric bias, confirmation bias, and needs for significance and social cohesion point to proximate naturalistic explanations of the fact that religions are so assertive and persistent in their claims to special authority in the face of obvious disagreement from so many competing faiths. (32-7)

The Normal/Objective Dilemma is really a question: if the psychological causes of religious belief are associated with normal, healthy mental functioning and various positive (individual and social) outcomes, should these rationally weigh with us more heavily than objective epistemological considerations would allow? (40)

Thornhill-Miller and Millican note that there are individual and social benefits of religious belief – enhanced happiness, increased longevity, improved recovery from addiction, deepened in-group trust, heightened in-group empathy, greater in-group cohesion, and so forth (37-41) – as well as individual and social costs of religious belief – intensified out-group conflict, increased insularity, greater xenophobia, heightened prejudices, and so on (41-3). Thornhill-Miller and Millican take the view that the in-group benefits are outweighed by the out-group damage; in their view, there may be no greater threat to humanity than intergroup conflict motivated by exclusivist and other-worldly religious thinking. (41) However, Thornhill-Miller and Millican also note that the very naturalness of religion makes it very doubtful that we can simply replace it with other things that deliver the same goods that it delivers: humanity is deeply immersed in well-established religious traditions whose rituals have evolved to fit human needs. (45)

Thornhill-Miller and Millican diverge in their preferences concerning the form that retreat from first-order religion should take. Millican favours walking the path of scepticism and learning to live in a godless world, something that many unbelievers have managed, and that might – with sensitive reshaping of social structures – be possible for all.

Thornhill-Miller opts for a kind of 'second-order religion' – deism – which finds intimations of divinity in the general structures of the world and in our religious instincts, but which is fully committed to the enterprise of natural science. (46) A major challenge for both approaches is whether they can deliver enough of the individual and social benefits that are currently delivered by religion while avoiding the individual and social costs that are associated with it.

<p style="text-align:center">II.</p>

Salamon argues against a retreat from first-order religion. In his view, Thornhill-Miller and Millican overlook or downplay the importance of a number of fundamental aspects of religious belief. In particular, he thinks that Thornhill-Miller and Millican undersell the significance of hopes for immortality, desires for moral transformation, and estimations of the value of love, worship, and freedom of assent. (216) Moreover, and perhaps more importantly, Salamon thinks that Thornhill-Miller and Millican misunderstand first-order religious traditions, which all explicitly or implicitly presuppose 'agatheistic' religious belief.

According to Salamon, agatheism 'identifies religiously conceived ultimate reality with the ultimate good which is postulated as a transcendental condition of our axiological consciousness through which we perceive and evaluate the goods at which our actions are aimed and towards which our hopes are directed' (201), and 'answers questions about the ultimate meaning of our finite existence as perceived through the lenses of our axiological consciousness which directs our thoughts and hopes towards some ultimate good which does not seem to be realisable in the physical universe' (201n7).

Salamon notes that he uses the terms 'ultimate reality' and 'the absolute' as 'synonyms capturing in the most inclusive way the meaning of the divine or highest reality that is the central focus of all religious traditions'. (202n9) He adds that 'all post-Axial religious traditions presuppose some possibility of ultimate fulfilment of human potential by way of transcending the limitations and contingency of our present condition; whether conceptualised in terms of salvation, redemption, liberation, or in some other way, this soteriological and eschatological promise is usually associated with the possibility of some kind of unity with the ultimate reality'. (222)

Against the claim of Thornhill-Miller and Millican that first-order religious belief is irrational – because there is no evidence or reason sufficient to sustain it – Salamon objects that rational first-order religious belief is grounded in 'reasoning from human axiological consciousness to God as the ultimate good towards which human axiological consciousness is ultimately directed'. (211) Moreover, against the claim of Thornhill-Miller and Millican that second-order religious belief might be rationally grounded in consideration of the fine-tuning of our universe for life, Salamon objects that those considerations are inadequate to ground a belief in God or ultimate reality that can satisfy human existential needs. (217) According to Salamon, the most that theistic arguments can do – and all that they were traditionally intended to do – is to show how presupposed belief in God can cohere with other beliefs about our world. (209) 'Nothing more can be done to establish the rationality of agatheistic beliefs ... than ... to point to the concept of the ultimate good as the transcendental condition of our axiological consciousness and to see agatheistic beliefs as objects of Kantian "rational faith", or, better, rational hope'. (242)

Salamon claims to favour 'mystical inclusivism'. In his assessment, his mystical inclusivism '(a) allows for the possibility of veridical experience of God or ultimate reality in a variety of religious traditions; but (b) avoids the radical revisionist postulates of Hickian pluralism, akin to the revisionism advocated by Thornhill-Miller and Millican; and (c) leaves open the question whether the creed of any specific tradition is a better approximation to the truth about ultimate reality than the creeds of other traditions, creating space for a kind of pan-inclusivism that acknowledges that everyone else is also an inclusivist'. (243)

III.

I think that it is not the case that first-order religious traditions presuppose agatheism. Certainly, there are some members of some religions who believe that there is an ultimate reality that is also the ultimate good; and it may even be there are some religions in which the belief that there is an ultimate reality that is also the ultimate good is widespread and, in some sense, mainstream. However, there are clear cases of first-order religious traditions in which agatheism is simply rejected.

Consider the family of Buddhist traditions. In one sense, these traditions reject the notion of ultimate reality: there is nothing in these

traditions that corresponds to the God of the Abrahamic religions (e.g., nothing that is permanent, unchanging, eternal, the causal foundation of everything else, and so forth). In another sense, there are 'ultimate realities' in Buddhist traditions: *dharma*, *sunyata*, *samsara*, etc. But there is no sense in which any of *these* 'ultimate realities' is the ultimate good.

Moreover, in Buddhist traditions, it is not true that the ultimate good is postulated as a transcendental condition of our axiological consciousness through which we perceive and evaluate the goods at which our actions are aimed and towards which our hopes are directed. Rather, according to Buddhist traditions, the 'ultimate good' is nirvana: release from samsara and consequent cessation of suffering. While, in many Christian and Muslim traditions, the ultimate meaning of our finite existence lies in some future good that is not realisable in the physical universe, in many Buddhist traditions, our ultimate goal lies in the pursuit of individual-annihilation-facilitating enlightenment.

While it is true that there is a sense in which Buddhist traditions maintain the possibility of ultimate fulfilment of human potential by way of transcending the limitations and contingency of our present condition, it is not the case that this is conceptualised as the possibility of some kind of unity with ultimate reality. According to many Buddhist traditions, we transcend the limitations and contingency of our present condition by coming to a complete understanding of those limitations and that contingency: when we become fully enlightened about the causes of our suffering and have done all that our karma requires, we are released from samsara and have achieved nirvana.

It is not only Buddhist traditions that reject agatheism. The same is true for Hinduism, Daoism, Confucianism, Jainism, and most – if not all – indigenous religions. Consider the family of Hindu traditions. Classical Hinduism teaches that there are four proper objectives of human life – dharma, artha, kāma, and mokṣa – of which, perhaps, the last can be taken to be the ultimate goal of human life. Among the diverse Hindu traditions, some take mokṣa to involve union – or realisation of union – with an ultimate being; but even those Hindu traditions typically do not take that ultimate being to be the ultimate good postulated by agatheism. And there are Hindu traditions in which mokṣa is not taken to involve union – or realisation of union – with any kind of ultimate reality.

When we survey the religions of the world, we do not find that first-order religious belief is grounded in reasoning from human axiological consciousness to God as the ultimate good towards which

human axiological consciousness is ultimately directed. If we say – with Salamon – that *rational* first-order religious belief is grounded in reasoning from human axiological consciousness to God as the ultimate good towards which human axiological consciousness is ultimately directed, then we commit ourselves to the claim that rational belief is very sparely and unevenly distributed across the world's religions. It is, I think, quite clear that most – if not all – Buddhists, Hindus, Daoists, Confucians, Jains and members of indigenous religions do *not* reason from human axiological consciousness to God as the ultimate good towards which human axiological consciousness is ultimately directed. Moreover, it is hardly any less clear that a great many Christians, Jews, and Muslims do not reason from human axiological consciousness to God as the ultimate good towards which human axiological consciousness is ultimately directed. Perhaps, in this latter case, it might be replied that the rational religious belief of these many Christians, Jews and Muslims is grounded in the reasoning of *other* Christians, Jews and Muslims who are held to be authoritative when it comes to the teachings of the Abrahamic religions. But, even if we suppose that there is this division of rational cognitive labour, it seems to me to be largely false that Christians, Jews and Muslims who are held to be authoritative when it comes to the teachings of the Abrahamic religions accept that *rational* first-order religious belief is grounded in reasoning from human axiological consciousness to God as the ultimate good towards which human axiological consciousness is ultimately directed. The kinds of considerations that are marshalled in – for example – Part I, Book II, Chapter 2 of Kant's *Critique of Practical Reason* simply do not have the support of large swathes of those Christians, Jews and Muslims who are held to be authoritative when it comes to the teachings of the Abrahamic religions.

It is, I think, a fair point against the position that Thornhill-Miller prefers that considerations about the fine-tuning of our universe for life are likely to prove inadequate to ground religious beliefs that answer to human existential needs. On its own, the claim that our world is the product of intelligent design seems powerless to minister to people with existential anxieties about death, loneliness, social status, sex, and so forth. However, it is not at all clear that reasoning from human axiological consciousness to the ultimate good towards which human axiological consciousness is ultimately directed is any better suited to the task at hand. Since Hume wrote his *Natural History of Religion*, it has

been a commonplace that 'vulgar superstitions' are much better suited to the relief of existential anxiety than are the abstruse deliverances of theologians. Belief in an Abrahamic afterlife might assuage – though it might also amplify – anxieties about death; reading Part I, Book II, Chapter 2 of *Kant's Critique of Pure Reason* is most unlikely to lead to any similar outcome. (Salamon makes it clear that he does not endorse Kant's 'moral argument'; he says that 'the fundamental intuition behind agatheism has more in common with Plato than with Kant' (201). But exactly the same kind of critical point could be made in connection with Plato's writings.)

I think that Salamon is insufficiently sceptical about what it is possible for arguments to do. In particular, I think that he is insufficiently sceptical about what it is possible for theistic arguments to do. A deductive argument can establish that a set of claims is logically inconsistent. Showing that a set of claims that includes the proposition that *God does not exist* is logically inconsistent is interesting to a theist only if all of the other claims in the inconsistent set are accepted by particular sufficiently well-credentialed atheists. In that case, the argument shows only that the atheists in question should reconsider the set of claims in question: something must give, but it is an entirely open question what that should be. While detecting – and, if necessary, demonstrating – logical inconsistency is an important part of the project of evaluating worldviews, it is arguably much more important to think about the evaluation of logically consistent worldviews. (Of course, deductive arguments establish logical consequences; one can learn about what one is currently committed to by one's beliefs by being presented with a good deductive argument. But atheists typically are not in need of arguments that show that they are committed to the claim that God does not exist.)

Salamon thinks that arguments can have a role in establishing the consistency of worldviews: 'theistic arguments can show how presupposed belief in God can cohere with other beliefs about our world'. It is hard to see how this could be the case. Certainly, one can construct arguments for *relative* consistency: we can show that a worldview W1 is logically consistent given that worldview W2 is logically consistent. But arguments for the existence of God are never of this form. And merely putting together an argument in which you show that the claim that God exists follows logically from other things that you believe cannot possibly suffice to show that your worldview is logically consistent: for all

that your demonstration shows, it may be that your premises are jointly logically inconsistent.

The interesting comparative task for worldviews considers their relative theoretical virtue: which worldview effects the best trade-off between minimisation of theoretical (ontological, ideological) commitment and maximisation of explanatory breadth and depth? If 'coherence' is not logical consistency, then it seems to me best to align it with theoretical virtue: more coherent theories make better trade-offs between minimisation of theoretical commitment and maximisation of explanatory breadth and depth. A worldview that is committed to the bare claim that our world is the product of intelligent design and to the bare claim that there is an afterlife in which we flourish scores better – on the count of theoretical commitment – than 'vulgar' worldviews that pack more into the nature of the intelligent designer and the afterlife, but much worse on the breadth and depth of the explanation why the intelligent design includes an afterlife of that kind.

I am sceptical that Salamon's 'mystical inclusivism' lives up to its advertising. Agatheism is no part of many first-order religious traditions. If we are interested in defending the rationality of adopting any among the world's first-order religious traditions, then we should not be following the trail that Salamon blazes. While I agree with Salamon that Thornhill-Miller's 'second-order religion' does not provide adequate ministration to human existential needs, I think that exactly the same complaint can be lodged against Salamon's 'mystical inclusivism'. If we are concerned to defend the rationality of 'vulgar' religious belief – i.e. the kind of religious belief that does provide adequate ministration to human existential needs – then we need to be adopting a very different kind of approach.

IV.

I am a metaphysical naturalist; I agree with Thornhill-Miller and Millican that one can happily, comfortably, and reasonably 'walk the path of scepticism'. However, I disagree with Thornhill-Miller and Millican that there is only one other plausible way forward: a second-order religion that is strongly supported by fine-tuning considerations.

Thornhill-Miller and Millican claim that 'the fine-tuning argument', in contrast with other traditional theistic arguments, has not been 'decisively refuted' (47). In their view, 'Ontological Arguments are logically refutable, Cosmological Arguments are vitiated by their

reliance on general principles that seem initially plausible but go hugely beyond the scope of our experience, and Moral Arguments are founded on meta-ethical views that are both dubious in themselves and hostage to naturalistic accounts of morality' (47n144). (They go on to add that these topics are 'obviously too big to discuss further here'.)

I agree with Thornhill-Miller and Millican that extant ontological, cosmological and moral *arguments* are unsuccessful: they are not such as ought to persuade metaphysical naturalists to become theists. Moreover, I think that the same is true for all of the other classes of extant theistic arguments: other teleological arguments (e.g. biological teleological arguments), arguments from consciousness, arguments from reason, arguments from revelation, arguments from scripture, arguments from expert testimony, arguments from miracles, arguments from religious experience, and so on. In my view, there are no extant successful theistic arguments, and there is no reason to suppose that there are hitherto undiscovered successful theistic arguments. In particular, I think that there is no reason to suppose that extant fine-tuning arguments are in better standing than other kinds of theistic arguments.

Given my scepticism about what it is possible for arguments to do, it will be better for me to frame the coming discussion in terms of the bearing of certain kinds of considerations on the comparative theoretical virtue of theism and naturalism. I claim that the fine-tuning considerations do not favour theism over naturalism because there is no difference in the depth and breadth of explanation of the fine-tuning considerations that is afforded by theism in comparison with naturalism. Indeed, if all else were equal, then the fact that theism postulates an intelligent designer to explain the fine-tuning considerations would entail that we should prefer naturalism to theism on the count of minimisation of theoretical commitments.

Suppose, for the sake of argument, that the causal order is fine-tuned for life. For simplicity, let's pretend that we know that the natural order is just our big-bang universe, and that there is an initial state of that big-bang universe. These simplifying assumptions clearly do not load the dice in favour of naturalism: if there is a creator God, then we can take the causal order to begin with God's creation of our big-bang universe; and if there is no creator God, then we can take the causal order to be our big-bang universe.

Now let's ask ourselves: where in the causal order are the values of the fine-tuned constants determined? Is the causal order everywhere

fine-tuned for life; or does the causal order become fine-tuned for life at some non-initial point?

On either view – theistic or naturalistic – if the causal order is everywhere fine-tuned for life, then, in particular, the initial causal state is fine-tuned for life. On the theistic view, this will be a matter of God's initial disposition to create a big-bang universe – our particular big bang universe – in which the fine-tuned constants take the values that they actually do; on the naturalistic view, this will be a matter of the fine-tuned constants taking the values that they actually do in the initial state of our universe. While I suspect that most theists will favour the view that God's initial creative disposition is brutely contingent – God could have had quite different initial creative dispositions and there is nothing that explains why God had the creative dispositions that he did rather than other creative dispositions that he might have had – and while I prefer the variant of the naturalistic view on which the initial state of our universe is brutely necessary, it is clear that theists and naturalists can jump either way on the modal status of the initial state. But, if that's right, then the views are simply on a par with respect to the virtues of the explanations that they give of the fine-tuning of the causal order for life.

Suppose, instead, that the causal order only becomes fine-tuned for life at some non-initial point. On either view – theistic or naturalistic – the only way in which there can be a transition from a state in which the causal order is not fine-tuned for life to a state in which the causal order is fine-tuned for life is if the transition of state is chancy with respect to the values that are taken by the constants. (Remember that the causal order is fine-tuned for life only if it is determined at that point in the causal order that our universe will have the fine-tuning of constants that it actually exhibits.) On the theistic view, God freely chooses from among a range of possible universes to create and there is nothing that explains why God freely chooses to create the particular universe that God chooses to create rather than any other possible universe that God might have chosen to create. On the naturalistic view, there is a range of possible transitions of state of the universe, and there is nothing that explains why we get the particular transition of state that we do, rather than any other possible transition of state that might have occurred. But, if that's right, then the views are simply on a par with respect to the virtues of the explanations that they give of the fine-tuning of the causal order for life: in each case, the key part of the explanation is that there was a tiny chance of getting what came to pass.

So, no matter where in the causal order the values of the fine-tuned constants are determined, there is no difference in the virtues of the naturalistic and theistic explanations of this fact.

Suppose that I am right to think that there is nothing special about fine-tuning arguments: they are in no better standing than any other arguments for the existence of God. Then it seems plausible to conclude that we should not think that the only plausible way forward – other than following the austere intellectual path of scepticism – is to pursue a second-order religion grounded in fine-tuning considerations. Should we then conclude that the only rational position to occupy is that of the naturalist? That would be very hasty. There are several relevant sets of considerations.

First, we need to be given some reason to think that there is something privileged about theistic religions. Even if we had a good argument that the only plausible way forward for friends of theistic religion is to pursue a second-order theistic religion grounded in fine-tuning considerations, that need not be a good argument that the only plausible way forward for friends of religion is to pursue a second-order religion grounded in fine-tuning considerations. I do not believe that there are compelling considerations that place Abrahamic religions in better rational standing than Eastern religions and indigenous religions.

Second, even if we are persuaded that fine-tuning considerations cannot carry a heavier load than other considerations advanced to support the existence of God, we might give a higher estimation of those other considerations than is provided by Thornton-Miller and Millican. While my own verdict is that, when we weigh all of the relevant considerations, my favoured naturalistic worldviews are more theoretically virtuous than any competing religious worldviews, I do not think that everyone else is rationally required to follow me in this judgment. There are many considerations that must be taken into account in weighing worldviews; it is not implausible, given the sheer complexity and scale of the task of weighing the theoretical virtues of worldviews, that there can be rational disagreement about the merits of competing worldviews.

Third, there is a lot that rests on the conceptions of rationality that are operative in this discussion. From the get go – in the first line of the abstract to their paper – Thornhill Miller and Millican say that they are interested in the 'possibilities and rational limits of supernatural religious belief'. But what we take to be 'the rational limits' of classes of beliefs depends crucially on what we take the requirements of rationality to be.

Suppose I ask you whether someone could rationally believe that Crystal Palace won the 2016 F. A. Cup Final. You might think that the answer to this question is obviously negative. Even if you have no knowledge about – and no interest in – soccer, you can readily discover that there is a Wikipedia page devoted to the 2016 F. A. Cup Final, and that what is says about the result of the game is confirmed by information on the official F. A. Cup website. Since accurate information about the result of the 2016 F. A. Cup Final is so easy to access, it would just be irrational for anyone who has anything riding on their belief about who won the 2016 F. A. Cup Final to believe that Crystal Palace won the 2016 F. A. Cup Final.

But you might also think that the answer to the question is really not so obvious. It is easy to imagine ways in which someone could come to have a rational belief that Crystal Palace won the 2016 F. A. Cup Final while also having something riding on the truth of that belief. Suppose that I am someone who loathes soccer; I have no interest in the game, and I do not go out of my way to acquire knowledge about it. I have just rung up a radio quiz, and I've been asked who won the 2016 F. A. Cup Final. I pretty firmly believe that it was Crystal Palace, because, a couple of weeks earlier, while travelling on the train to work, I overheard snatches of a conversation between two soccer fans who were discussing the 2016 F. A. Cup Final. Although I had no interest in what they were saying, and wasn't following what they said too closely, they were speaking sufficiently loudly that some of their words registered with me. On the basis of the words that I did hear, I was rationally justified in forming the belief the Crystal Palace won the 2016 F. A. Cup Final. (That is, the fragments of their conversation that I heard constituted strong misleading evidence for the claim that Crystal Palace won the 2016 F. A. Cup Final.) Moreover, the belief became more firmly entrenched because I dreamed about the train trip a couple of times, and woke up vividly recalling the parts of the conversation that made it reasonable for me to form the belief that Crystal Palace won the 2016 F. A. Cup Final.

Perhaps, given this example, you might think that we need to distinguish between *internalist* and *externalist* conceptions of rationality: while there is an internalist sense in which it is possible for one to rationally believe that Crystal Palace won the 2016 F. A. Cup Final, there is also an externalist sense in which it is impossible for anyone to rationally believe that Crystal Palace won the 2016 F. A. Cup Final. Or, perhaps, given our example, you might think that we need to distinguish

between *conditional* and *unconditional* attributions of rationality: while it is true that it is possible for one to rationally believe that Crystal Palace won the 2016 F. A. Cup Final, it is not possible for one to rationally believe that Crystal Palace won the 2016 F. A. Cup Final *given* that one is in possession of certain readily available information.

An interesting feature of my example is that it suggests that there are cases in which it is possible for there to be rational belief of claims that are uncontroversially false. When we are thinking about claims that are uncontroversially false, we find it unproblematic to suppose that there are externalist and conditional senses in which belief in those claims is irrational. No one who is sufficiently intelligent, sufficiently reflective, sufficiently interested, and sufficiently well-informed believes that Crystal Palace won the 2016 F. A. Cup Final. Anyone who might plausibly count as an *expert* when it comes to assessing the claim that Crystal Palace won the 2016 F. A. Cup Final believes that it is not the case that Crystal Palace won the 2016 F. A. Cup Final.

There are many domains in which there is convergence of expert opinion *because* expert opinion tracks what any sufficiently intelligent, sufficiently reflective, sufficiently interested, and sufficiently well-informed person ought to believe. In mathematics, natural science, human science, medicine, engineering, pharmacy, and a host of other theoretical and practical disciplines, there are vast domains where expert opinion converges because expert opinion tracks what any sufficiently intelligent, sufficiently reflective, sufficiently interested, and sufficiently well-informed person ought to believe. In those domains, it would be irrational for sufficiently intelligent, sufficiently reflective, sufficiently interested, and sufficiently well-informed people to dissent from expert opinion.

But not all domains of inquiry are like this. In some domains of inquiry, there is no convergence of expert opinion. Philosophy is one such domain: the opinions of even the most intelligent, most reflective, most interested and most well-informed philosophers fail to converge. Sure, there are times and places where there is significant local convergence of expert philosophical opinion; but history does not disclose any stable global convergence of expert philosophical opinion. Some, but not all, intelligent, reflective, interested, well-informed philosophers have been, and some, but not all, intelligent, reflective, interested and well-informed philosophers are, determinists, substance dualists, consequentialists, communitarians, virtue ethicists, logical pluralists, phenomenologists,

existentialists, physicalists, legal positivists, and so forth. What credence, then, can we give to claims that it is *irrational* to believe in determinism, or consequentialism, or communitarianism, or virtue ethics, or logical pluralism, or phenomenology, or existentialism, or physicalism, or legal positivism, etc.?

And theism? Some, but not all, intelligent, reflective, interested, well-informed philosophers have been, and some, but not all, intelligent, reflective, interested and well-informed philosophers are, theists. Is it really credible to suppose that the theistic beliefs of those philosophers who are theists are irrational in ways that the naturalistic beliefs of those philosophers who are naturalists are not? Certainly, there are *standards* against which the theistic beliefs of those philosophers who are theists are irrational: for example, it is more or less certain that the theistic beliefs of those philosophers who are theists are not beliefs that would be held by ideal Bayesian agents. But, against those standards, it is true in general that the philosophical beliefs of philosophers are irrational. In any case, what really matters is not the standards that we adopt, but the comparative case that is to be made: is there any compelling reason to think that the naturalistic beliefs of naturalistic philosophers are *more* rational than the theistic beliefs of theistic philosophers?

Thornhill-Miller and Millican invoke human cognitive failings – egocentric bias, confirmation bias, optimistic bias, and the like – in the case that they make for the Common Core/Diversity Dilemma. But, of course, these human cognitive failings are universal: we can run up against these failings anywhere that people are engaged in reasoning and argumentation. Moreover, we have good reason to suspect that these kinds of biases are in operation wherever there are strongly held beliefs despite a lack of convergence of expert opinion. So a question naturally arises about the extent to which the views of Thornhill-Miller and Millican on the question of the rationality of naturalistic and theistic beliefs are themselves affected by these universal cognitive failings. Perhaps you can imagine a slightly different paper on the limits of rational *philosophical* belief that works with the following pair of 'Dilemmas':

(1) In so far as philosophical data (e.g. philosophical intuition) point toward specific philosophical theories or explanations, their diversity and mutual opposition undermines their evidential force; and, in so far as such philosophical data involves a common core of similarity, they point towards theories and explanations that are scientific rather than philosophical

(2) If the psychological causes of philosophical belief are associated with normal, healthy mental functioning and positive individual and social outcomes, should these rationally weigh with more heavily than objective epistemological considerations would allow?

I note, in passing, that there is a considerable empirical literature on philosophy for children that supports the claim that philosophical belief *is* associated with normal, healthy, mental functioning and positive individual and social outcomes.

To strengthen the case for scepticism about the claim that theistic – and, more generally, religious – belief is, *ipso facto*, irrational in ways that naturalistic philosophical belief is not, we might also consider the ways in which we all rely upon testimonial information in forming and revising our beliefs. While at least those of us who are sufficiently intelligent, reflective, interested and well-informed are pretty good at identifying genuine experts in mathematics, natural science, human science, medicine, engineering, pharmacy, and the rest of the theoretical and practical disciplines adverted to earlier, it is not within the bounds of credibility that we are good at identifying genuine experts that we can then reasonably take to be testimonial authorities on philosophical – or religious, or normative political – questions. But, of course, we all acquire much of our philosophical – and religious, and normative political – belief from the testimony of those whom we suppose at the time to be authorities on the matters in question. And there is no prospect of making over our philosophical – and religious, and normative political – beliefs in ways that free them from dependence upon the testimony of those we once regarded as authorities on these matters.

There is much that might be added to the rather sketchy considerations that I have advanced against the claim, defended by Thornhill-Miller and Millican, that the only *rational* ways forward lie with naturalism and second-order religion grounded in the fine-tuning data. However, rather than try to develop these considerations more fully, I shall conclude with a comment on another controversial aspect of the position that they stake out.

Thornhill-Miller and Millican say that second order religion may be able to deliver the in-group goods that are supported by first-order religion without also delivering accompanying out-group damage. But why should we be more optimistic about this prospect than about the

prospect that first-order religion might be able to deliver the in-group goods without also delivering the out-group damage? Thornhill-Miller and Millican themselves note that that the well-springs of the out-group damage lie in an authoritarianism that is not the sole preserve of first-order religions. While Thornhill-Miller and Millican quite correctly note that it is a mistake to interpret the widespread decline of religiosity in prosperous democracies as evidence that first-order religion is on the way out, it remains open that that data supports the claim that the spread of appropriate social conditions can moderate, or even eliminate, the out-group damage that has often been associated with first-order religion. Given that there are many first-order religionists who get the benefits of first-order religion without causing the out-group damage, why not suggest working to bring about social conditions in which first-order religion everywhere delivers in-group benefits without causing out-group harms?