



Tim Mawson
OXFORD UNIVERSITY, UK

Divine Motivation Theory

By Linda Trinkaus Zagzebski

Cambridge: Cambridge University Press, 2004; 428 pp.; hb. £ 40.00,
pb. £ 18.99; ISBN: 0-521-82880-5/0-521-53576-X

[1] In *Divine Motivation Theory*, Linda Zagzebski articulates and defends a motivation-based virtue ethic, one which makes emotion primary and then goes on to make the virtues and emotions of God, specifically the Trinitarian God of Christianity, central. Thus the resultant theory 'might be called a divine virtue theory' (1) but in fact ends up being called 'Divine Motivation Theory', the name of the theory thus giving the book its title. Towards the latter half of the work, she provides interlocking accounts of how DM theory, as it has by then been abbreviated, enables us to assess the importance of *imitatio Dei* and the incarnation as well as to solve the seeming paradoxes of perfect goodness; the Euthyphro dilemma; and the Problem of Evil. Finally, there is a discussion of the implications of DM theory for ethical pluralism.

[2] Zagzebski first advances arguments for the conclusion that the moral properties held by persons, their acts, and the outcomes of their acts are 'derivative from a good motive, the most basic component of a virtue, where what I mean by a motive is an emotion that initiates and directs action' (1). She then proposes 'a more substantial theory in theistic metaphysics according to which the motives of God are the ontological basis for the value of everything outside of God' (6). 'In motivation-based virtue theory, the moral properties of acts, ends, and the outcomes of acts are derivative from the goodness of motives; in DM theory, what makes human motives good is their similarity to divine motives' (282-283). There is then, over the course of the book, a movement from Zagzebski's attempts to establish the internal coherence and superiority of an initially-non-theistic motivation-based virtue theory over competing metaethical theories to her attempts to show its theistic variant's greater consonance with Christianity and superiority over other theological theories in dealing with some of the problems that Christianity is traditionally supposed to face. So, Zagzebski suggests that if her arguments are good, the first half of the work should appeal to all; the second to all who share with her the premise of the truth of Christianity. The whole is written in a clear and accessible style, one which should be understandable by anyone from the undergraduate level upwards.

[3] *Divine Motivation Theory* would not be the stimulating and substantive work of Moral Philosophy cum Philosophical Theology that it is were there not a number of points at which one might raise objections. In what space remains to me here, I shall offer a couple of objections to points Zagzebski makes near the

beginning of her argument.

^[4] One objection that one might have threatens to be too general to count as an objection to any particular move made within the game that Zagzebski is playing and, as such, I raise it now primarily as an aid to our pushing it to one side so that we can concentrate on the other points. This is what might be called the ‘cart before the horse’ objection that one can raise to virtue ethics *per se*: surely, virtuous persons get to be virtuous because they regularly-enough choose to perform virtuous actions (or, germane to Zagzebski’s variant, act from virtuous motives); virtuous actions (motives) don’t get to be virtuous because they’re regularly-enough chosen (acted from) by virtuous persons. If, like this reader, you incline towards this thought, push it to one side. Push it to one side to enable you to concentrate on the details of Zagzebski’s DM theory, but push it to one side also because DM theory must be the version of virtue theory that offers one most promise of overcoming this worry. On DM theory, ‘God is good in the same way that the standard meter stick is one meter long’ (285). Virtuous human persons do indeed get to be virtuous because they regularly-enough choose to perform virtuous actions (act from virtuous motives); virtuous actions (motives) don’t indeed get to be virtuous because they’re regularly-enough chosen (acted on) by virtuous human persons. But virtuous actions (motives) *do* get to be virtuous because they’re regularly enough, indeed unfailingly, chosen (acted on) by the virtuous God. Why isn’t this last claim obvious? Indeed, why might it seem obviously false to an atheist? Because that’s the real essence of virtuous actions (motives), something that we characteristically don’t have in mind and may well not even know about when we secure reference to virtuous actions (motives) through the use of examples of the actions done (motives acted on) by paradigmatically virtuous human people, what Zagzebski calls exemplars. So, push the ‘cart before the horse’ worry to one side. If Zagzebski is right, eventually cart and horse find their union in God.

^[5] Zagzebski has many interesting things to say about emotion, the starting point of her theory. She tells us that, ‘My position is that an emotion is a state that has both cognitive and affective aspects that are not separable states’ (59). ‘Corresponding to each emotion there is a thick concept under which the agent represents the intentional object of her emotion. Examples are *rude*, *dangerous*, *pitiful*, and *contemptible*. Pity is an emotion consisting in feeling pity for something seen as pitiful; love is an emotion consisting in loving something seen as lovable; fear is an emotion consisting in feeling afraid of something seen as dangerous; contempt is an emotion consisting in feeling contempt for something contemptible, and so on. The thick properties of the situations are properties of the intentional object of the feeling, not the cause of the feeling. This distinguishes concepts such as *rude* from concepts that refer only to a certain response in the agent, such as *irritating* or *nauseating*. Feeling irritated or nauseated has a cause, but no intentional object’ (61). This has as a consequence that ‘there are concepts [*rude*, *dangerous*, et al.] the possession of which is not a purely cognitive state’ (62) but also motivates – given another thesis of Zagzebski, that emotions are what motivate – when one believes the concepts instantiated. Yet, Zagzebski admits,

people can be 'unable to see someone as rude because their brain damage affects their ability to feel emotion. They can judge that a person is rude, because they learn to identify the descriptive features of situations that others see as rude, but they do not see anything as rude, and they have no motive to respond in the way people who are offended characteristically do' (69). So, Zagzebski seems to be contending that people who suffer certain sorts of brain damage can form correct judgements as to the extension of the concept of rude but, merely because they fail to be emotionally affected or immediately motivated having come to a belief that it is instantiated in a particular case, fail to count as possessing that concept. One might be more chary than Zagzebski of denying such people possession of the concept. I am.

^[6] Although I have not suffered any obvious brain damage (obvious to me that is; perhaps it is of the nature of such damage that others would be better judges of it if I had suffered from it), I incline to think that I approach in my attitude to rudeness the people Zagzebski has in mind here and yet also think that I possess the relevant concept; my grasp on it is no less sure for approaching these people in my lack of affective response; if I completely lost my affective response, I wouldn't lose the concept; and I might very well be a better moral agent were I to approach them more closely or even completely. So it is that I would maintain that I recognise that certain instances of humour, for example, are 'off colour'; 'gauche'; 'rude'; 'offensive'; et al., but I fail myself to be greatly emotionally moved by this recognition and I fail to be directly motivated by it to any great extent. I am relatively emotionally blind, as it were, to offensiveness. I am not completely emotionally blind, I admit, but I regard this lack of completeness in my emotional blindness not as the slender thread which yet links me to the concept but as quite possibly a failure on my part, one which I might do well to remedy so as to leave me in possession of the concept but unencumbered by emotion in doing what I ought in situations where it is instantiated. I find at least somewhat plausible the claim that humans tend to have prudish sensibilities which prevent them enjoying what really are funny jokes by making them feel offended, the view espoused, for example, by certain fans of Monty Python in response to criticism of *Life of Brian*, the final scene of which depicts Brian, the Jesus figure, singing a cheerful ditty endorsing the fundamental meaninglessness of life whilst being crucified. The view that one ought not to be emotionally affected by that which is offensive is one which it strikes me arguments might draw one towards and thus it is a view that should be allowed at least to be put on the table by someone who claims to have achieved the goal it sets for one. According to Zagzebski, by trying, a la these fans of Monty Python, to articulate the thought that offensiveness doesn't affect and directly motivate oneself and it oughtn't to affect and directly motivate others, one would show oneself not to be in possession of the concept of offensiveness. I would understand it were someone to argue that the view that offensiveness oughtn't to affect or directly motivate was just plain mistaken, perhaps because emotions are our guide to moral properties and so by emotionally blinding oneself to offensiveness, one erodes one's ability to make judgements of offensiveness. I would understand it were someone to argue

that the view was mistaken because in fact at the source of all motivation is an emotional response and thus by emotionally blinding oneself to offensiveness, one erodes one's motivation to avoid offensiveness. Both of these things may be true and if they are true, they count against the view that I suggested I find at least somewhat plausible, but Zagzebski seems committed to the point that nobody can actually achieve the end that the view sets for them, because it's impossible for someone to have the concept of offensiveness yet succeed in taking this approach to it. That I find hard to believe.

^[7] Let us move on. According to Zagzebski, these emotional responses are 'what gets us going' (71); 'I am convinced that an affective state is necessary to motivate' (123). I imagine the following situation. Someone is deliberately rude to me in a public bar. As a result I feel like punching him (I am not, I have admitted, entirely unmoved emotionally by offensiveness, especially after a few drinks and when I discern that this offensiveness has been deliberately directed at myself); I feel warm towards the fondly-imagined prospect of his punched-out form prostrate on the floor; I feel rather cold towards the prospect of my 'backing down' and leaving. I consider two possible exemplars of nobility: Jesus and the sorts of character normally played by Arnold Schwarzenegger. I don't feel at all emotionally drawn to the prospect of behaving closer to the exemplar of Jesus (whose words of advice about turning the other cheek I call to mind); quite the opposite. Rather, I find myself drawn to the exemplar set for me by the types of person Arnold Schwarzenegger characteristically plays. But I do call Jesus' words to mind and I recognise that Jesus offers me the better role model. I leave the bar without having attempted to 'terminate' anyone. I have behaved in a way then to which I was not emotionally drawn. What 'got me going' then was not an emotion, *pace* Zagzebski; I was motivated by non-affective moral reflection to act contrary to the promptings of my emotions. I have of course made this situation up, but similar situations are surely familiar to all of us from everyday life.

^[8] Zagzebski has a reply to this line of thought: 'We have emotions automatically, but we learn them in part by imitation, and they change under the influence of the emotion of admired others' (52). 'I surmise that the move from "I want to be like R and not like S" to "R is better than S" is not only genetically primitive, but also basic to moral thinking' (53). So in fact, according to Zagzebski, what must have happened in my case had this story been true is that, at some earlier stage in my life, I was emotionally drawn to the pattern of life set for me by Jesus (someone whom I admired) – 'I want to be like Jesus and not like The Terminator', I thought – an emotional drawing which left its imprint on me in the judgement I made in the bar, a judgement at that stage shorn of its affective aspect, 'Better to be like Jesus than The Terminator'. Well perhaps. I cannot remember making such comparisons as those surmised by Zagzebski as being the first 'moves' in my moral development, but that's hardly conclusive; I suppose that if push comes to shove, she can always claim that this was going on at a subconscious level or in a childhood now lost from recall. Even so, it does seem that an affective state is *not* necessary to motivate at the time of the choice to perform every (any?) particular action. Let's leave this to one side for a moment

then and turn to look at Zagzebski bravely taking on what look like two difficult cases for her theory.

^[9] Zagzebski points out that ‘there are two virtues that appear not to have an emotion-disposition as a component because they are almost entirely directed toward the control of emotion—courage, which controls fear, and temperance, which controls the desire for pleasure. These virtues require special attention, since they appear to be counterexamples to my claim that every virtue has an emotion as a primary component’ (112). The solution comes in saying that ‘Courage and temperance . . . have higher-order emotions as their primary components. They are emotions whose intentional objects are situations that include the agent’s own emotions’ (124). ‘My conjecture is that the entire situation is the intentional object of some emotion’ (126). Straight off the bat: is it really true to the phenomenology of one’s moral experience that one has as an intentional object in such cases a representation of a situation and oneself and one’s emotions in it? If not, then there’s nothing for the putative higher-order emotion to latch onto and thus the theory falls down.

^[10] Let’s turn from courage and temperance to another virtue, prudence, and an example of its being deployed merely so as to make the prudent decision about how to structure one’s life so as to get most pleasure from it, all other things being supposed for the sake of the example to be equal.

^[11] Let us suppose then that as an inveterate smoker I am suddenly made aware of the probable long term effects smoking will have on my health. I have a certain emotional response to a possible far-future self as I picture him in my mind’s eye suffering from some smoking-induced disease; I have another emotional response to a possible close-future self as I picture him in my mind’s eye enjoying his next cigarette. I make the decision, let’s say, to stop smoking that very day. The force and vivacity with which I’m emotionally drawn to my smoking my next cigarette is far greater than that by which I’m repelled from my suffering some smoking-induced disease in the far future, yet I nevertheless act as I do. Hume ‘solves’ – only in the sense of distracts one from – the problem that I seem to act contrary to the prompting of my strongest emotion by talking of calm passions. Similarly, according to Zagzebski, an emotion is at the root of my action even in this case, not the weak repulsion of an imagined future disease-ridden self, but the emotional attraction I feel towards being the sort of person who distributes pleasures over his or her life so as to maximise the area under the graph of pleasure against time, rather than the gradient on it at a particular moment of choice. But, introspection suggest to me, I might feel no such emotion towards such an intentional object even if I do hold it in my mind (as most often I do not). I just see that it’s more reasonable to be the sort of person who takes short-term losses in pleasure for greater long-term gains. I may or may not feel any emotion about being the sort of person who is prudentially reasonable in that way. In fact, in this case, the negative emotion caused by my nicotine craving may well blot out all other emotional responses leaving naked prudential reason and willpower alone to explain why it is that I am, nevertheless, able to stop smoking. So it seems to me that reason, as well as emotion, could, does, and sometimes

should, motivate. Of course, again one might claim that at some stage in my life, now forgotten, I said to myself, 'I want to be like J. S. Mill (or rather J. S. Mill's conception of the rational agent) when it comes to structuring pleasures over time, not like Sid Vicious', the dim emotionless echo of which guides me here when I say to myself, if I do say to myself, 'Better to be like J. S. Mill's conception of the rational agent than like Sid Vicious'. Again, one cannot show this to be false, but of course one has no reason prior to accepting Zagzebski's claims to believe it to be true.

[12] Here is another interesting thing that Zagzebski says about emotion. 'A belief is true just in case the belief corresponds to reality. Similarly, I propose that an emotion is good just in case it "fits" its intentional object. . . . This means that emotions have something analogous to truth value . . . fitness is not in the mind of the perceiver. The emotion fits just in case the standard judgement expressing it is true' (76–77). 'To see things right is to be in an affective state that is such that the standard judgement expressing it is true. So we can truly say that one person's affective state is right and another's is wrong. What we cannot say is that the kinds of dispositions human beings have are the right dispositions to have, as if the dispositions of other imaginable intelligent creatures would be wrong to have. We have these dispositions because that is part of what it is to be human. The dispositions cannot be wrong unless it is wrong to be human, and I assume it is senseless to say that; but for the same reason, it is senseless to say it is right' (82). This might raise a worry revealed by the following thought experiment.

[13] Martians land and start zapping everything that moves with ray guns. These Martians, we discover, have as a universal feature of their species what we might call an inverted emotional spectrum. Zapping people with ray guns is, we are free to posit, cruel and wantonly destructive, but Martians, in contrast to most humans, feel affectively drawn towards the cruel and wantonly destructive. Wouldn't we incline to say in such a case that we humans were right in our affective responses to cruelty and wanton destruction; our emotions did fit these things in a way that Martians did not? If there was some simple-to-administer 'tweak' by which we could flip someone's emotional spectrum, then shouldn't we do so for Martians and not for humans? Isn't the fact that we should do so for Martians because it's wrong to be an (untweaked) Martian in a way that it isn't wrong to be an (untweaked) human? Isn't the reason why it's wrong to be an (untweaked) Martian and it isn't wrong to be an (untweaked) human because – regardless of what affective dispositions any creatures happen to have towards it – cruelty and wanton destruction are in fact wrong?

[14] If this train of thought is right, anyone accepting the arguments of the first half of *Divine Motivation Theory* would, it seems to me, be likely to be moved to accept something along the lines of the 'ideal observer theory' that Zagzebski discusses towards the end of her book and put something pretty much like the theistic God in the place of the ideal observer. There's a species-transcendent standard by which (untweaked) Martians' affective responses are wrong and (untweaked) humans' affective responses are right; the best way to explain this is by saying that (untweaked) Martians don't resemble God in this regard in the way

that (untweaked) humans do and it's God's reactions that determine what's really right and what's really wrong when it comes to affective responses. If this is correct (and of course one can raise objections to it), it would show simply that there was scope for a moral argument for the existence of God framed in Zagzebski's terms, something that would itself merely indicate that if her arguments in the first half of the book are good, then they have even more interesting implications than she herself would wish to claim for them. As such, this point may be mischaracterised as an objection to her ideas and arguments.

^[15] I have not touched on many of the themes Zagzebski weaves together in the rich tapestry of her text, but I hope that, in picking out what little I have picked out, I have given an indication of the significance of what she says there. As always when reading her, one finds that even when one disagrees with what is being said, one cannot help but be glad that it's been said and said so clearly.