

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 H₂O, but is another substance, XYZ. XYZ has the same properties of taste and appearance as H₂O, 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 H₂O 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 trustworthy, 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)