

Preface

Within philosophy, the field of moral psychology has exploded in recent decades. In this short Element, there is no way to cover all of these developments. Instead, we will focus very selectively on the following big questions:

Are we ultimately motivated by our own self-interest in everything that we do? Is there any evidence for genuinely altruistic or selfless, as opposed to egoistic, motivation? [Section 2]

What is involved in being a good person, in having good character, and in meeting the standards of virtue? What are the virtues and vices, and are they possessed by most of us? [Section 3]

What motivates us to act? Is it always our desires, or could a belief, or even something else, motivate us instead? [Section 4]

When we judge something to be the right thing to do, are we always motivated to act accordingly? Or is it possible to be indifferent to what we genuinely think we ought to do? [Section 5]

When we act on our moral judgments, are we typically acting on the basis of moral principles and moral reasons? Or is our behavior largely driven by nonrational intuitions or gut reactions, followed by a made-up story that we tell ourselves afterwards about the moral principles and reasons which supposedly guided our thinking? [Section 6]

To be sure, these are not the only big questions in moral psychology, but they are among the central questions that have dominated the field in the past fifty years.

Two omissions are worth noting. First, we will not look at moral epistemology, including the justification of moral beliefs and the implications of moral disagreement, since these issues are covered by Tristram McPherson in another Element.¹ Second, we will not delve into the literature on moral responsibility, since that is also covered in another Element currently in progress.

Intended Audience and Background

Moral Psychology is intended for a broad audience. I have tried to make it suitable for undergraduate students with some background in philosophy, as well as intellectually curious readers more generally. At the same time, I hope it will be of interest to academics whose work intersects with topics in moral psychology.

Unlike a scholarly monograph, this volume mostly focuses on offering an overview of several central debates in moral psychology. The goal is to provide a nonspecialist with a good lay of the land that can serve as a starting point for diving into some of the more advanced material in this field.

¹ McPherson 2020.

While much of the Element is an overview of the various topics, each section will end with my own view on the debate. I have written on various topics in moral psychology over the past fifteen years. One of the things that excited me about this project was the chance to go back to my own work – which is scattered across journal articles, edited volume chapters, handbook overviews, encyclopedia entries, and book reviews – and try to incorporate it into a coherent picture (I hope!) about how our minds work when it comes to moral matters.

1 Introduction to Moral Psychology

What is the field of moral psychology? There are many different ways of characterizing it, and I do not want to claim that any one way has to be adopted. Personally, I prefer to think of moral psychology broadly as the field of study in moral philosophy which is centrally concerned with better understanding what is involved in moral thinking, and how that thinking does and does not give rise to morally relevant action.

There is a lot packed into this. Let me first expand a bit on the ‘moral’ aspect, and then on ‘moral thinking.’ Finally, I will say more about how the field of moral psychology relates to moral philosophy.

1.1 The Moral

Let’s start at a general level with the distinction between the normative and the descriptive.² Here are some *descriptive* statements:

“ $2+2=4$ ”

“A water molecule is composed of two hydrogen atoms and one oxygen atom.”

“George Washington was the first president of the United States of America.”

“Benjamin Franklin was the first president of the United States of America.”

These statements, as descriptive, only aim to capture the way the world actually is. Not all of them succeed in doing so, such as the fourth one about Benjamin Franklin. For descriptive statements, the key concern is whether the world is as each statement says it is.

Normative statements, on the other hand, are concerned with evaluating the way the world is:

“There should be more charitable organizations.”

“He is a cowardly person.”

“The criminal deserves the death penalty.”

“If only he had acted for better reasons.”

² Section 1.1 draws on Miller 2011, with permission of Bloomsbury Publishers.

Sometimes the descriptive and the normative come together. The statement that a person is cowardly may at once describe the person's actual character *and* also evaluate him as a person. But we know that the descriptive and normative domains often come apart. For descriptively the world as it *is*, falls short in all kinds of ways from the way that it *should be* normatively. There is war, but there should be peace. There is murder, but there should not be any murder. Examples abound.

Let's focus on the normative domain a bit more. Moral statements are a central part of the normative domain. But they are not the only part, as we can see below:

Prudential (long-term self-interest):

"Smoking is bad for you."

"Exercise will be good for you in the long run."

Aesthetic:

"That painting is beautiful."

"The room is tastefully designed."

Legal:

"This law is unjust."

"He should be sent to jail for breaking the law."

Religious:

"Communion should be served every Sunday."

"It is better for adults to be baptized rather than children."

Furthermore, these different normative domains can come into conflict with each other. For instance, it might be *legally* permitted in some societies to own slaves, but still be *morally* wrong to do so. Hence, in addition to each of these specific normative domains, there is also the idea of what we should do or what would be good to do, *all things considered*.

Unfortunately, philosophers have had little success in trying to clarify what makes a normative statement a specifically *moral* one. I do not have anything better to offer myself, and so in what follows I will simply adopt the attitude that we know a moral statement when we see it.

1.2 Moral Thinking

A number of different elements in our minds play an important role in our moral thinking. Here I will highlight three of them: beliefs, desires, and judgments.

We all have a host of *moral beliefs*. I might believe, for instance, that my friend is trustworthy, world peace is good, and this politician is corrupt. These count as *moral* beliefs because moral concepts are part of the contents of the

beliefs (*trustworthy, good, and corrupt*). Of course, we have plenty of nonmoral beliefs too, such as my belief that $2+2=4$, which can be very important as well to our moral thinking. My thoughts about whether someone told a lie or took a bribe or killed an attacker in self-defense involve nonmoral beliefs, but they can clearly factor heavily into my moral deliberation.

At the same time, I have plenty of moral *desires* too: desires with contents where moral concepts play a role. So I want to do the right thing, I want to become a better person, and I want my leaders to stop being so corrupt. Here too we have plenty of nonmoral desires, and again they can be very important to our moral thinking. An example is my desire to not say or do anything that would be embarrassing. Such a desire can influence my thinking so that I rationalize away why I didn't speak up in front of a group on an important issue.

A popular way to distinguish between beliefs and desires has been to appeal to differences in their *direction of fit*. Very roughly, the thought is that beliefs have a mind-to-world direction of fit – they aim to represent or capture the way the world is. If the belief is false, that is because it has failed to capture the world accurately. My belief that I teach at Wake Forest University aims to represent the way the world is, and fortunately it succeeds in so doing (since I really do teach at Wake Forest University!).

Desires, on the other hand, have a world-to-mind direction of fit. They aim to change the world (as seen by the desire) to have the world reflect the desire. Failing to capture the world is no fault of the desire. Rather, the fault is with the world in not reflecting what the desire aims for it to be like. If I want Wake Forest to win a national championship in college football (which I do), it is no fault of my desire if that does not ever happen (and, alas, it has not yet). I want the world to change so that Wake Forest has a national championship to its name, even if sadly it does not.³

Beliefs and desires, in turn, factor into the formation of moral *judgments*. Suppose I am deciding between three charities when trying to figure out where to make a donation to famine relief. I want famine to be eliminated. I also want to donate to the most efficient charity I can. On the basis of doing some research, I learn that one of the charities spends a lot of money on administrative costs. Another has a history of corruption. Eventually I form the moral judgment that I should make a donation to the third charity.

In this example, my moral judgment was formed through a process of moral deliberation, which is often understood as conscious and reflective. Sometimes,

³ If we adopt a very broad understanding of 'desire,' then it could include wishes, wants, hopes, and intentions. Emotions such as fear and anger have been treated as belief-like or as desire-like by different philosophers, and we will not explore those debates here. For more on direction-of-fit accounts as well as their problems, see Sobel and Copp 2001.

though, moral judgments can be formed more immediately, such as when it is just obvious to me that, for instance, slavery is wrong. To the extent to which I reason to that conclusion, it would appear that the reasoning is happening under my conscious radar screen, subconsciously.

Finally, with respect to moral judgments, what should we say about them? Are they beliefs or desires? It turns out that this is a source of much controversy in the philosophical literature. The majority view, called *moral cognitivism*, holds that moral judgments are beliefs. But a minority view, called *moral noncognitivism*, holds that our moral judgments are, no surprise, desires. To make matters more complicated, there is some recent work exploring whether a hybrid account might be best, where moral judgments have both belief-like and desire-like dimensions. While not a focus of this Element, we will return to issues about cognitivism versus noncognitivism in Section 5.⁴

1.3 Situating the Field of Moral Psychology within Moral Philosophy

It is common to distinguish between three branches of contemporary moral philosophy: ethical theory, meta-ethics, and applied ethics.⁵ How does moral psychology relate to them? First, a brief word of explanation about each.

Ethical Theory. Let's start with the claim that slavery is wrong. There are many possibilities for what makes it the case that slavery is wrong, including the pain and suffering it causes, how it violates the dignity of persons, and how it is cruel and inhumane. The ethical theorist attempts to sort through these options in order to arrive at the most promising account of the feature(s) which makes slavery wrong. More generally, the ethical theorist attempts to develop an understanding of what the relationship is between moral properties (such as wrongness) and nonmoral properties (such as causing pain). For example, on one ethical theory wrongness might be understood as a matter of causing pain. On another ethical theory, it might be understood as a matter of going against the commands of God. Leading ethical theories include utilitarianism, Kantian ethics, virtue ethics, and divine command theory.⁶

Meta-Ethics. Meta-ethics can be understood as the nonmoral study of the metaphysics, epistemology, and semantics of the moral. Unlike ethical theory, meta-ethical approaches are developed by examining the practice of morality from a disengaged perspective and typically refrain from making moral

⁴ For more on moral cognitivism and non-cognitivism, see van Roojen 2018.

⁵ Section 1.3 draws on Miller 2011, with permission of Bloomsbury Publishers.

⁶ For a helpful introduction to ethical theory, see Timmons 2012.

claims. In other words, meta-ethics raises and attempts to answer questions *about* morality. To use an analogy, a scientist arrives at first-order scientific conclusions, whereas a philosopher of science examines the practice of science as such, and does not make any scientific discoveries. So too is the meta-ethicist not concerned, in the first instance, with arriving at new ethical claims, but rather with the answers to various questions about morality, such as the following:

- Do moral facts exist?
- If so, are they objective?
- If they are not objective, who or what created them?
- How do we learn the content of morality, if there is such content to learn in the first place?
- What is the meaning of moral terms?
- Are moral statements capable of being true or false?
- If so, are any of them true?

A central goal of the leading meta-ethical positions is to answer questions such as these.⁷

Applied Ethics. As its name suggests, applied ethics examines the moral status of specific human actions and practices, including those which have become prominent in societal debates such as abortion, the death penalty, cloning, stem cell research, animal consumption, access to scarce medical resources, and so forth.

Where does moral psychology fit into these three leading areas of moral philosophy? The answer is that it is not confined to any one of them, but rather is important to all three of them.

Hence, as we will see in Sections 4 and 5, debates about motivation have been highly influential in meta-ethics, for instance. In normative ethics, Kantian, utilitarian, virtue ethical, and other approaches have had much to say about what our moral psychology should look like. And in applied ethics, one example where moral psychology is important is in trying to figure out how best to motivate people to change their behavior when it comes to, say, donating to famine relief agencies. These are just a couple out of a myriad of examples which highlight the importance of moral psychology to the different branches of moral philosophy.

Enough stage-setting, however. Let us dive into our first central issue in moral psychology.

⁷ For a helpful introduction to meta-ethics, see Shafer-Landau 2003a.

2 Do We Ultimately Only Care About Ourselves? Egoism and the Alternatives

Let's begin with the following case, which we will return to in every section of this Element:

The Hospital Visit. Sally breaks her leg in a skiing accident. She has been in the hospital for several days without any visitors, so she texts her friend, Franklin, tells him what happened, and mentions the name of the hospital in the hope that he will come to visit her.

When Franklin reads Sally's text, he decides that he should go visit Sally in the hospital that afternoon. A few hours later, he walks into her hospital room. Sally is very glad to see him, and they have a very enjoyable time together.⁸

Franklin decided to visit his friend Sally in the hospital. That much we have already said. *Why* did he ultimately decide to go? This we do not know from the description of the case.

If the position in moral psychology called *psychological egoism* is correct, then we do know a bit more. The answer has to be that Franklin wants to benefit himself in some way. This is because, according to psychological egoism, everyone is ultimately motivated to benefit themselves in everything they do.

In the coming subsections, we will first unpack what the psychological egoist is claiming, and also see what some of the alternative positions are. Then we will look at three important philosophical objections which can be raised against the view. The section ends by shifting to the empirical literature, first with a brief summary of some results from dictator games, and then by reviewing a fascinating strand of research on empathy.

2.1 Clarifying the Positions

In Book II of Plato's *Republic*, one of the main characters, named Glaucon, recounts a famous myth about the Ring of Gyges:

a shepherd . . . saw there was a corpse inside that looked larger than human size. It had nothing on except a gold ring on its hand; he slipped it off and went out . . . while he was sitting with the others, he chanced to turn the collet of the ring to himself, toward the inside of his hand; when he did this, he became invisible to those sitting by him, and they discussed him as though he were away . . . he immediately contrived to be one of the messengers to the king. When he arrived, he committed adultery with the king's wife and, along with her, set upon the king and killed him. And so he took over the rule.⁹

⁸ This case is inspired by an example from Stocker 1976. ⁹ Plato 359d–360b.

The main lesson of the story is not supposed to be about the shepherd. It is supposed to be about us. The claim is that if we had such a ring, we too would do whatever we thought would benefit ourselves. That might not take the same form as the shepherd's behavior, but nevertheless the focus would still be on ourselves.

Such a claim is an expression of psychological egoism. The view can be stated as follows:

(PE) The ultimate goal of each person's actions is the pursuit of his or her self-interest, subjectively understood.

What is self-interest? Loosely, we can understand it as what will benefit us in some way. Unfortunately, there are a myriad of theories trying to give a more precise definition. For instance, a *hedonistic* account of self-interest might describe benefits in terms of pleasure and the avoidance of pain. So, hedonistic psychological egoism could be written as:

(PE_{hed}) The ultimate goal of each person's actions is the pursuit of his or her pleasure and avoidance of pain, subjectively understood.

But a hedonistic account could be replaced with a variety of alternative accounts of self-interest, such as self-interest being what satisfies my desires or what promotes my flourishing. Fortunately for the purposes of this Element, we can remain at the generic level of just talking, as (PE) does, about self-interest.

Why the 'subjectively understood' in (PE)? Because it seems clear that we can be mistaken about what is in our self-interest. Someone might be motivated to experience pleasure from eating a fruit, for instance, only to bite into it and discover that it had gone bad and tastes rotten. For the psychological egoist, this action still fits the view since the goal is self-interest, from the person's own perspective.

Note that psychological egoism is both an empirical and a universal claim. It is empirical in that it attempts to describe how human beings actually are, rather than how they should be. It is universal in that it is making a claim about *all* human beings, without exception. Hence, strictly speaking, to refute psychological egoism one would need just one person performing one action for which (PE) is not true. At the same time, to really develop an interesting position in opposition to (PE), it would be better to claim something more substantive. For instance, one might try to argue that in certain conditions, *most* human beings are capable of acting in a way that does not ultimately involve the pursuit of their own self-interest. One possibility, to be explored later, is that sometimes we might be motivated selflessly by love for our closest family members.

Finally, psychological egoism is *not* the same thing as selfishness. A really selfish person never helps others in need. But it is quite possible for someone to be a psychological egoist and also deeply committed to charity work or volunteering for a good cause. Such a person would just be helping others as a means to pursuing his own self-interest as he sees it.

The opposite kind of motivation from egoistic motivation is *altruistic*. One version of psychological altruism claims that:

(PA) One ultimate goal for at least many people can be the pursuit of what would benefit another person, subjectively understood, regardless of whether the actor would benefit or not.

Less abstractly, if I am altruistically motivated, then I am ultimately concerned with what is good for another person. Whether I would benefit in the process is not my concern. For instance, when Franklin goes to the hospital to visit Sally, if he is concerned with just helping her during this difficult time, then that counts as being altruistically motivated.

Now, it is compatible with (PA) that one can be altruistically motivated and try to help someone else out while *also* benefiting from the action at the same time. So long as the benefit is merely a side effect or by-product, rather than the goal of the action, it would not detract from the action's being altruistic. To help with the distinction between a goal and a by-product, I like to use the analogy of driving my car. My goal is to arrive at my destination, but a side effect (in this case a negative one) is that my car emits exhaust into the air. So too might it be the case that Franklin's goal is to just help Sally, and a by-product of the helping is that he also feels good about what he did.

Psychological altruism can be formulated in a variety of ways. As stated in (PA), it is a claim about *many* people. A very weak version of the view would be about just one person. A very strong version would be about all people. The very weak version might be true, but it is not very interesting. The very strong version is unlikely to be true. It is not clear that newborns are able to think in those terms. Also, psychopaths may be incapable of altruistic motivation, even if the rest of us are.

Of course if psychological egoism is true, then *none* of us is capable of altruistic motivation, or motivation whose ultimate goal is to benefit others. A stark way to bring out the contrast between the two views is as follows:

(PE*) For all human beings, altruistic motivation does not exist.

(PA*) For most human beings, altruistic motivation does exist.

Again, this is not the only way to develop the contrast between psychological egoism and altruism, but adopting this approach gives us a sharp juxtaposition between them.

It is worth stressing that the altruist is happy to grant that for much of the time, we are motivated by the pursuit of our own self-interest. Furthermore, she can and should grant that we often have more than one motive for our actions, so that cases of altruistic motivation can also be cases of egoistic motivation. Franklin might visit Sally both because he cares about her and wants to do what he can to help her for her own sake, *and* because he also does not want to feel bad as a result of not visiting her in the hospital.¹⁰ The independent presence of the egoistic motive does not negate the existence of the altruistic motive.

Egoistic and altruistic motivation might seem like the only two kinds there could be. Not so fast. First, note that altruistic motivation is characterized in terms of *benefiting* another person. At least in principle, there could be motivation concerned with another person, rather than oneself, but not with benefiting that person – rather, it is focused on harming him, for instance. We will return to this in more detail in the next subsection.

More significantly, at least in theory there could be *dutiful* motivation, which is neither egoistic nor altruistic. If Franklin visits Sally because he thinks it is right and he is ultimately motivated to do the right thing, then that would count as dutiful.¹¹ Thanks to the work of Immanuel Kant in particular, dutiful motivation has been widely discussed in moral psychology. As we will explore in Section 3, Kant thought that it was the only kind of highly praiseworthy motivation there is.

Psychological altruists can be neutral about whether dutiful motivation exists. As far as their view is concerned, they do not have to take a stand. Psychological egoists, though, must reject dutiful motivation, along with altruistic motivation and any other kinds there might be, because they only accept motives which aim at benefiting oneself.

One final note of clarification about psychological egoism before we move on to assessing it. It is important to not confuse it with *ethical egoism*, which is the following view:

(EE) The ultimate goal of each person's actions should be the pursuit of his or her self-interest.

¹⁰ "Feeling bad" can take a variety of different forms. For instance, a person could be motivated to do something to avoid feeling guilty, to avoid feeling ashamed, or to avoid feeling embarrassed. These would all count as egoistic motivations. Thanks to Dale Miller for noting the need to clarify this.

¹¹ Note that the "ultimately" is important here. If he is motivated to do the right thing, but that in turn stems from a deeper desire to make a good impression on other people, then his motivation would ultimately count as egoistic.