

## Introduction

This Element is written for the student who is new to these topics, though not to philosophy. It is an opinionated introduction to two influential theories of the nature of ethics. It aims to introduce these theories in ways that are sufficiently accessible to whet the appetite, yet sufficiently accurate so as not to mislead. To aid the reader, terms that first appear in **bold** feature in a glossary at the end of the Element.

## 1 Motivations and Methodology

### 1.1 Moral Debates

On 23 June 2016 the UK government asked 46,500,001 of its residents and citizens the following question:

*Should the United Kingdom remain a member of the European Union or leave the European Union?*

Prior to the referendum, the Britain Stronger in Europe campaign argued for remaining, claiming that membership of the Union brought jobs, lower prices, and protections for workers' rights. The Vote Leave campaign argued for leaving on the basis that it would allow the UK to control its borders, make its own laws, and save £350 million per week.<sup>1</sup> Emotions ran high; accusations of dishonesty and treachery flew; opponents were branded 'disgraceful', 'shameful', 'abhorrent', even 'enemies of the people'. When the votes were counted, 17,410,742 people had voted to leave, 16,141,241 to remain. The consequences of this decision are still unfolding.

On 26 July 2017 the president of the United States announced that

*the United States Government will not accept or allow Transgender individuals to serve in any capacity in the US Military.*

The president argued that the ban would allow the military to focus on 'decisive and overwhelming victory' and remove the burden of the 'tremendous medical costs and disruption that transgender in the military would entail'. The ban was strongly condemned by the American Civil Liberties Union, who called it 'outrageous' and 'desperate'. A *USA Today* editorial called it 'medically unethical', 'cruel', and 'senseless'. Others pointed out that the cost of transgender-related healthcare would amount to between 0.004 and 0.017 per cent of the

<sup>1</sup> See *The Brexit Collection: 2016 referendum* (LSE Digital Library) here: <https://digital.library.lse.ac.uk/collections/brexit/2016>.

Defense Department's total healthcare spend. The debate about the permissible roles of transgender people in the US military continues.<sup>2</sup>

Each of these cases exemplifies a moral or ethical debate – an issue framed in terms of what we *should* do, what is *permissible*, what is *right* or *best*. From the large to the small scale, these issues pervade our lives. Our responses to them shape how we act individually and collectively. Such issues have three interesting features. First, they are intimately connected to motivation, action, and emotion. We want to know what we *should do*, or what the *right thing* to do is, because we want to know what *to do* – and our motivations and emotions affect what we *do* do. Second, they involve genuine moral disagreement. In each case there is a moral question at stake to which the disputing parties are attempting to find an answer. Finally, such issues are not settled by fiat. Merely thinking that leaving is the right thing to do does not make it so. The correct answers to moral questions are independent of our minds.

## 1.2 Metaethics and the Attractions of Subjectivism and Expressivism

Suppose you wanted not to *settle* these disputes but to understand the general class into which they fall. In other words, suppose you wanted to understand what a moral issue is and what is going on when people engage in moral debates. Then you would be what philosophers call a **metaethicist**. Subjectivism and expressivism are metaethical theories. In particular they are theories about what is going on when people make moral judgements such as the judgement that the UK should leave the European Union. According to **subjectivism** such judgements report attitudes, in something like the way that psychiatrists provide reports on the mental states of their patients. So to say that leaving is the right thing to do is to report on one's own positive feelings or emotions, or perhaps the feelings of one's tribe, class, or society. Ethics is a branch of psychology. According to **expressivism**, by contrast, moral judgements *express* attitudes, in something like the way that cheering expresses support of a football team. So to say that leaving is the right thing to do is to express a positive feeling or emotion towards leaving. Ethics is a branch of rhetoric, whose goal is to shape others' attitudes. (These characterisations are rough and will be refined as we progress.)

<sup>2</sup> CNN reported the President's comments on July 27 2017; see here: <https://edition.cnn.com/2017/07/26/politics/trump-military-transgender/index.html>. Events are the enemies of deadlines, and sometime after this Element was drafted, but before it was published, an even more significant event – the Covid-19 pandemic – highlighted multiple further examples of moral disputes, concerning, for example, the wisdom of various lockdown measures, how to ration scarce medical resources, and how to balance public health with economic prosperity. What I say in the text about moral debates applies just as well to these (as I write) startlingly salient problems.

The rest of this Element is concerned with the differences between, and plausibility of, these two theories; but before battle is joined, it is worth noting some shared features that make both attractive enough to investigate further.

First, subjectivism and expressivism cook with the same naturalistically respectable set of ingredients. According to both, all we need to cite in order to understand our moral practice is the natural world, our reactions to it, and the things we do with those reactions. Such elements are required anyway for any scientifically respectable account of human beings. Subjectivism and expressivism merely redeploy them to account for a very particular thing that human beings do to each other, namely moralise. This austere set of ingredients contrasts with the expansive set favoured by some (non-naturalist) moral realists, who hold that, as well as the properties countenanced by natural science, there exist irreducible moral properties such as goodness and rightness.<sup>3</sup> Ockham's razor famously asks us not to multiply entities beyond necessity; subjectivism and expressivism are more clean-shaven than most.

Second, this naturalistically kosher set of ingredients looks to be the correct set with which to explain the peculiar institution of morality. Both subjectivism and expressivism foreground our emotions, responses, attitudes, and motivations, and the thought that morality is concerned with such things is both intuitive and empirically grounded. Intuitive because we expect someone with moral stances to *care* in ways consistent with those stances. Empirically grounded because common experience and scientific studies show that moral judgements co-occur with, and are causally influenced by, emotions. One study, for instance (cited in Prinz 2006), found that people are more likely to make strong moral judgements after being primed to feel disgust (by being made to sit at a filthy desk).

A final appealing feature of subjectivism and expressivism is that the ingredients that they cook with are multiple, fecund, and permissive. The diversity of human emotional responses allows each theory to account for moralities with distinct emotional tones, such as sin-based moralities that foreground guilt, or paternalistic moralities that foreground pride and disappointment. Further, different emotions can be deployed to explain different types of moral verdict. For example, the difference between an action being merely bad and being shameful lies in the distinct emotion invoked by the latter. Finally, these ingredients are permissive insofar as they promise to explain other types of issues with which humans engage. For example, it is tempting to think that whether something is *beautiful*, *fashionable*, or *funny* can likewise be understood in terms of descriptions or expressions of attitudes. Subjectivism and

---

<sup>3</sup> Non-naturalist moral realists include Shafer-Landau (2003) and Enoch (2011).

expressivism are therefore worth studying not merely as theories of ethics but as adaptable models that can apply to other practices (especially areas where taking a realist view runs up against Ockham's razor).

So subjectivism and expressivism start as genuine contenders. In order to see whether this promise bears fruit we need to define our terms and consider our methodology. These are the tasks for sections 1.3 and 1.4.

### 1.3 Terminology and Initial Characterisations

Subjectivism and expressivism are primarily theories of moral judgements. By a **moral judgement** I mean the type of speech act performed when someone assertorically utters a moral sentence such as 'Murder is wrong'. So saying these words while performing a play or word-associating doesn't count; but reflecting on the issue of unlawful killing, coming to consider that it is not permissible, and then voicing that thought by saying 'Murder is wrong', does count.

What about the thought thus voiced or expressed? What is involved in *thinking* that murder is wrong? I call this a **moral commitment**. It is a mental state, and its nature is one of the points at issue between subjectivists and expressivists. Commitments are more enduring than judgements: I can think that murder is wrong long before I voice this opinion. Commitments are mental states that people *have* or *are in*, whereas judgements are acts of speech that people *make* or *perform*. Nevertheless the connections between them are close – a standard assumption is that judgements express commitments.<sup>4</sup>

Subjectivism and expressivism both talk about attitudes, so this term also needs elucidation. By an **attitude** I mean a mental state that provides us with a goal and some motivation to pursue it. This is in contrast with cognitions, which provide us with a depiction of the way the world is that is independent of any direction as to what to do about it. Desires are attitudes, insofar as my desire for chocolate, say, provides me with a goal (to have chocolate) and some motivation to act in ways that I believe will help me secure that goal. This motivation need not be conclusive – I may have a stronger desire to lose weight, for example – but it is real, nonetheless. In contrast my belief that the Netherlands is flat is a cognition insofar as it represents the lay of the land but gives me no guidance as to what to do about that fact. How I end up acting will largely depend on a complex interplay between my attitudes and cognitions. If I desire chocolate and believe there to be some in the fridge, then, other things

<sup>4</sup> Some use the term 'moral utterance' to refer to the speech act I call a moral judgement and 'moral judgement' to refer to the mental state I am calling a moral commitment. Like most metaethical terminology, usage is treacherous, but the important distinction to remember is that between the speech act and the mental state.

being equal, I will go to the fridge. Things are seldom equal, of course, since opposed attitudes may result in me channelling my efforts in other directions. In every case, however, action seems to require both a goal-setting attitude and a cognition that provides the agent with an idea of how to pursue that goal.

Desires are one kind of attitude, but there are others. Emotions, preferences, intentions, aversions, phobias, approvals, and disapprovals all fall into this category insofar as they provide motivational direction. Conversely, perceptions and intuitions seem to be cognitions insofar as they depict the world. The attitude/cognition distinction has a venerable history, being traceable from Plato, through David Hume, to modern philosophers such as Elizabeth Anscombe and David Lewis.<sup>5</sup>

The final terms in need of definition are **report** and **express**. According to subjectivism moral judgements report attitudes; according to expressivism they express attitudes. So what is the difference? Roughly, when we report attitudes, we describe them. To describe is to represent the world as being a certain way or to make a claim about the way things are. So when we report *attitudes*, we make a claim about the ways things are *psychologically*, and when we report our *own* attitudes, we make a claim about the way things are with *our own* psychology. Descriptions can be accurate or inaccurate, true or false. By contrast when we express attitudes, we do something other than report them. We voice them, externalise them, lay them out before others. These metaphors get us somewhere, but they remain mysterious (and potentially misleading) – we shall have to return to this issue later (see Section 5.2.2). For now, it is worth noting some structural features of expression. First, when we express attitudes, we do not report or describe. Thus when we express attitudes there is no possibility of that expression being ‘accurate’ or ‘inaccurate’ – or at least no sense of it being accurate or inaccurate *as a description*. Likewise, there is no possibility of it being true or false, agreed or disagreed with – or at least, no sense in it being these things *qua description*. (Whether there are other senses in which expressions of attitudes can be these things is discussed in Section 5.2.3.) Second, it is commonly assumed that whatever it means to say that moral judgements express attitudes, regular non-moral judgements – such as my judgement that grass is green – express beliefs in just the same way.<sup>6</sup> The thought here is that to describe is to express a belief, and regular non-moral judgements function to describe.

These assumptions suggest improved definitions of expressivism and subjectivism. According to expressivism moral judgements express our attitudes;

<sup>5</sup> See Plato’s *Republic*, Hume’s *A Treatise of Human Nature* (book 2), Anscombe (1957), and Lewis (1988).

<sup>6</sup> See, for example, Gibbard (2003: 75) and Schroeder (2008: 18).

and remembering that moral commitments just are the mental states expressed by moral judgements, we can characterise expressivism as:

E. Moral judgements express attitudes. Hence moral commitments are attitudes.

According to subjectivism moral judgements report or describe attitudes. To describe is to express a belief, so moral judgements express beliefs about attitudes. Thus:

S. Moral judgements express beliefs about attitudes (i.e. they describe attitudes). Hence moral commitments are beliefs about attitudes.

As we shall see, these characterisations are still deficient but they highlight the fact that one point of contention between expressivism and subjectivism concerns the nature of moral commitments: are they attitudes or beliefs about attitudes?

## 1.4 Methodology

### 1.4.1 *Moral Practice, Semantics, and Metasemantics*

Subjectivism and expressivism are theories of moral judgement, but much else besides. Moral judgements are one small part of the phenomenon of **moral practice**, which includes the myriad ways in which moral concepts feature in human thought, language, and interaction. Other parts of this practice include moral phenomenology (what it is like to intuit a moral fact), moral deliberation (the thought processes that lead to forming or revising one's moral views), moral debate (the public exchange of moral views), moral argument (inferring moral claims from others), and moral guidance (having one's feelings and actions shaped by moral thought and debate). Any subjectivist or expressivist worth their salt will have something to say about all these things. Nevertheless I shall follow tradition and focus primarily on moral judgements. This is justifiable both insofar as it makes the issues tractable and because in most cases the account of the other aspects of moral practice is strongly suggested by the account of moral judgements.

Even with this narrower focus, however, we need to consider what it means to be a *theory* of moral judgements. Subjectivism and expressivism are standardly taken to be theories that *give* or *explain* the *meaning* of these judgements. So when expressivists claim that moral judgements express attitudes, they are not claiming that this is a contingent or accidental feature of their usage. They are claiming that the fact that moral judgements express attitudes is a necessary part of the explanation of their meaning.

In this sense subjectivism and expressivism are *semantic* theories, concerned with *meaning*. Yet we need to be careful here, because the terminology of ‘semantics’ risks ambiguity. For by ‘semantic theory’ we can mean either **first-order semantic theory** or **metasemantic theory**. I think that subjectivism and expressivism are best understood as the latter, but to explain why we first need to understand the distinction.

Suppose you are a monolingual English speaker interested in the meaning of the judgement ‘Schnee ist weiß’. I tell you that ‘Schnee ist weiß’ is German and means that snow is white. It seems there has been some progress. I have given you a first-order semantic theory for this judgement: a theory that tells you what it means, that *gives* you its meaning.

Suppose on the other hand that you are a bilingual English and German speaker who knows full well what ‘Schnee ist weiß’ means. Yet still, you want to know *why* this type of judgement, as typically made by actual speakers of German, means what it does. Why is it that just these words allow speakers of German to make just this claim about just these things? After all there is nothing necessary about the fact that ‘Schnee’ refers to snow – the same string of letters could have easily referred to snowshoes, or pyramids. One tempting answer is that the judgement has the stable (that is, non-context-specific) meaning it does in virtue of characteristically expressing or propounding a particular belief, viz. the belief that snow is white. This is not a complete theory of meaning, to be sure – it explains the *meaning* of a type of judgement in terms of a suspiciously similar notion of the *content* of a belief characteristically expressed – but it is a substantive theory, nonetheless.<sup>7</sup> It is a *metasemantic* theory insofar as it provides an account of the contingent facts that explain why a particular type of judgement, as typically made by a group of people, has the stable meaning that it does.

There is a distinction, then, between theories that *give* you the meaning of a type of judgement and theories that *explain* why that judgement has the meaning it does. I take subjectivism and expressivism to be theories of the latter sort.<sup>8</sup> Subjectivists claim that we can explain the meaning of moral judgements in terms of their characteristic ability to express beliefs about attitudes. Expressivists claim that we can explain the meaning of moral judgements in terms of their characteristic ability to express attitudes. If we define the **semantic function** of a set of judgements as the characteristic function that

<sup>7</sup> This is sometimes called ‘ideationism’ or a ‘psychologising’ or ‘dog-legged’ approach to meaning. See Blackburn (1984: 39–45), Laurence (1996), and Ridge (2014: 107).

<sup>8</sup> For the recently popular idea that key issues in metaethics are metasemantic rather than semantic, see, for example, Chrisman (2012) and Ridge (2014).

explains their stable meaning, then we can further refine our two theories as follows:

S\*. The semantic function of moral judgements is to express beliefs about attitudes (i.e. to describe those attitudes). Hence moral commitments are beliefs about attitudes.

E\*. The semantic function of moral judgements is to express attitudes. Hence moral commitments are attitudes.

Unfortunately, the existing literature on subjectivism and expressivism often misses the distinction between first-order semantics and metasemantics. It is common, for example, to take expressivism to be a semantic theory that *gives* meanings or paraphrases of moral judgements.<sup>9</sup> This is a mistake. Expressivism is no more a theory that gives you meanings than Marxism is a theory that gives you capital. Both theories explain their targets rather than producing them.

There are many reasons why the distinction between first-order semantics and metasemantics is overlooked, but one salient from the current context is that S\* is a metasemantic theory that quickly generates claims in first-order semantics as well. According to S\* the meaning of moral judgements is explained in terms of those judgements expressing beliefs about attitudes. Further, it seems a plausible independent assumption that, where the semantic function of a judgement is to express a belief, the meaning of that judgement is the same as the content of the belief expressed.<sup>10</sup> It follows that, according to S\*, the meaning of moral judgements is given by claims about attitudes. Thus S\* quickly generates first-order semantic claims (or ‘analyses’) of the sort:

S1. ‘Murder is wrong’ means ‘I disapprove of murder’.

Nevertheless the metasemantic claim S\* remains distinct from the first-order semantic claim S1. The tendency, when discussing subjectivism, to focus the latter, and then to consider expressivism a *rival* to subjectivism, is one source of the mistaken thought that expressivism, too, provides psychological paraphrases of moral judgements.

#### 1.4.2 *First Desideratum: Accommodation*

So subjectivism and expressivism are metasemantic theories, that is, theories of those facts in virtue of which moral judgements have the meaning that they do.

<sup>9</sup> For example, Cuneo (2006: 35), Olson (2010), and Wright (1992: 11).

<sup>10</sup> See Schroeder (2008): 32–3.



In a more expansive sense they are also theories of the moral practice that grows up with and surrounds such judgements. Yet how are such theories to be judged?

Like many metaethicists, I accept two desiderata for metaethical theories.<sup>11</sup> The first is accommodation. A good metaethical theory should be able to **accommodate** the forms of ordinary moral practice and the assumptions of those who engage with it. To accommodate a form or assumption is to provide an explanation of why a practice with that form or assumption has arisen and (ideally) to justify that practice carrying on with that form or assumption (to accommodate is to learn to live with, to cohabit).<sup>12</sup> The forms of moral practice include the fact that moral judgements often come to us unbidden, the fact that moral judgements are made using sentences of indicative form, the fact that moral judgements feature in public debates, and the fact that people offer reasons in support of moral claims. Furthermore, people who engage in moral practice make assumptions when doing so. For example, they assume that some actions are wrong, others right, that there is such a thing as moral disagreement, that debate can help discover answers to moral questions, that merely making a moral judgement does not make it correct, and so on. These assumptions are revealed in people's behaviour as much as their explicit beliefs.

I shall refer to these forms and assumptions as the 'features' of moral practice. The reason that a good metaethical theory needs to accommodate them is that, if it did not, it would fail to be a theory of *moral* practice at all. For example, a metaethical theory that could not explain why moral sentences have indicative form and predicts instead that they have imperative form (like 'Shut the door!') seems to be a theory of something other than our actual practice.

There are hundreds of features of moral practice that good metaethical theories should accommodate. To make things tractable, I will consider just three. These are the practicality of moral judgements, moral disagreement, and moral mind-independence. The next three subsections say a little about each. The rest of this Element will then consider whether various versions of subjectivism and expressivism can accommodate them. Of course, the result will not be exhaustive. For each feature of moral practice that is not discussed there will be a dimension of the evaluation of these theories that is missing. So my approach is necessarily partial; but I hope that by focusing tightly on three features this

<sup>11</sup> A canonical statement is Timmons (1999: 12).

<sup>12</sup> Note that there are two senses of 'accommodation' in play here: a weaker sense that demands only *explanation* and a stronger sense that also demands *justification* or *vindication* of the relevant form or assumption. The argument I give in the next paragraph seems only to support the former, and there is genuine controversy about the latter: see Section 1.4.4, Loeb (2007), and Sinclair (2012).

Element will provide the reader with an indication of the sort of accommodation strategies subjectivists and expressivists can deploy, as well as the characteristic problems with those strategies – thus equipping the reader with the tools to complete a more thorough assessment of both views.

### 1.4.2.1 Practicality

Consider first the fact that moral judgements are practical. There are at least two parts to this. First, moral judgements answer practical questions. Return to the issue of the UK's relationship with the European Union, as set out in the 2016 referendum question. What are we to do: leave or remain? This is a practical question, seemingly answered by the claim that the *right* thing to do, the thing we *should* do, is leave (or remain). Moral judgements seem to have just the right sort of meaning – **normative meaning** – to help us answer such practical questions. By contrast, many non-moral judgements lack such content. Suppose I was told that leaving will lead to higher prices. That wouldn't settle the issue of what to do, since I may not care about higher prices.<sup>13</sup>

The second dimension of practicality concerns the connection to motivation. There seems to be a necessary connection between making a (sincere) moral judgement and being appropriately motivated by it. The necessity of this connection is revealed by the fact that, when faced with an agent who (apparently sincerely) judged that leaving was the right thing to do and yet who showed no indication of supporting the leave campaign, no indication of regretting the fact they felt no such motivation, and even voted remain, then we would begin to doubt that they understood what it means to call an action 'right'. It is important to note that the necessary connection here is between moral judgement and appropriate *motivation*, not moral judgement and appropriate *action*. A motivation is an attitude – an internal push to action – but how I end up acting depends on how all my motivations interact, and some motivations may be overridden by others. It would be an absurdly strong thesis to say that there is a necessary connection between moral judgement and appropriate *action*, for this would entail that people always act in accord with their moral judgements, as if people always lived up to their moral ideals. Note also that even the connection between moral judgement and *motivation* is not without exception. We are comfortable with the thought that the connection can break down in exceptional cases, for example in corrupted sadists (who are now explicitly motivated away from good and towards evil), jaded politicians (who have been on the political scene for too long, and seen too much, to remain moved by their ideological beliefs), or motivationally drained parents

<sup>13</sup> This paragraph summarises ideas from Blackburn (1998: 70, 90–1).