

The final and definitive version of this paper can be found in Thought, 2017.

How Verbal Reports of Desire May Mislead

We each have many desires, and we sometimes report the existence of such desires. I might tell you that I don't want any dessert, that she does want to marry you, or that my parents want to buy a house. In this paper I highlight two noteworthy features of assertions about our desires, and then highlight two ways in which they can mislead us into drawing unwarranted conclusions about desire. Some of our assertions may indicate that we are sometimes motivated independently of desire, and other assertions may suggest that there are vast divergences between our normative judgements and our desires. But I suggest that some such assertions are, in this respect, potentially misleading, and have in fact misled authors such as Russ Shafer-Landau (2003), Jack Woods (2014), and Tim Scanlon (1998).

Before I begin, a few quick notes on terminology: I shall use the words "want" and "desire" interchangeably, picking my usage by ear. I shall assume that our desires can be divided into instrumental desires – when we want something as a means to an end – and non-instrumental desires – when we want something for its own sake.¹ I shall often make claims about psychological states *motivating* us, which means that they incline us to act. Finally, in section 3, I discuss "normative judgements": these include our various views about what is good, bad, right, wrong, and what we have reason to do or feel. Our normative judgements include, but are not exhausted by, our moral views.

1. Two Features of Desire Discourse

¹ Some have denied the existence of instrumental desires (e.g. Chan 2004), but I shall set such views aside as implausible. We certainly say things that suggest that we have instrumental desires, as when I say "many people want to win the lottery". That assertion seems true, and it would be very difficult to maintain that such desires are non-instrumental, given that they would clearly be extinguished by relevant changes in belief.

The first noteworthy feature of assertions about desire I shall label *the ambiguity effect*. The ambiguity effect is that assertions about what we don't want are ambiguous. To see this, first note that there can be conflicts of desire. So long as we focus on pro tanto desires and not overall preferences, it is clear that one can want to go for a beer right now (for the taste), but also want to not go for a beer right now (in order to work). To this extent, whether you desire that P is independent of whether you desire that $\neg P$: you can see something attractive about one of these, the other, both (conflict), or neither (disinterest). With this simple point in hand, we can see the ambiguity of assertions like "S doesn't want P" (cf. Broome 2013: 274). Such an assertion might mean that S fails to want P, or might instead mean that S does want $\neg P$. To illustrate the first possibility, imagine that I ask whether you want my office repainted. You might say that you don't, not because you are against this possibility, but rather because you fail to be at all in favour of it. Here, "S doesn't want P" means that S fails to desire that P, not that S desires that $\neg P$. In contrast, imagine that I say "I don't want to visit the dentist". I might well here be indicating not that I fail to want to go to the dentist, but instead that I positively want not to go. Here, "S doesn't want P" means that S desires that $\neg P$, not necessarily that S fails to desire that P. We can see this by noting that I might more emphatically say "I *really* don't want to visit the dentist", which surely conveys the presence of a strong desire not to visit the dentist, not (impossibly) the *strong absence* of a desire to visit the dentist. The important upshot of all this is that we should acknowledge that someone might truthfully assert "S doesn't want P" even though S does want P: if this assertion indicates that S desires that $\neg P$, it is consistent with S also desiring that P. I might really want not to go to the dentist but *also*, reasonably, have some instrumental desire to go.

The second noteworthy feature of some assertions about desire I shall label *the quantity effect*. The quantity effect is that for pragmatic reasons, we will sometimes felicitously fail to assert, or even deny, the existence of desires that we in fact think do exist (here my discussion runs parallel to Mark Schroeder's discussion of negative reason existentials (2007: 92-103). To see this, begin by

noting that our desires are numerous. You might simultaneously desire to go for a walk (to think), to go for a run (for exercise), and to stay at your desk (to write). Indeed, given that we often have weak desires even regarding very remote possibilities, our desires are presumably *extremely* numerous: you might well have *some* desire to become a hermit (no more emails!), some desire that your desk be on fire (for the warmth), and perhaps even – if you consider the matter – some desire to eat your car (for the iron content) (Schroeder 2007: 95-6).

With this point in hand, let us turn to Grice's maxim of quantity, which says to make assertions only when they are informative (Grice 1989: 26). Because we have *some* desire for so many things, learning that someone has some desire that P is rarely informative. As a result, Grice's maxim tells us not to assert "S desires that P" if we mean only to convey the bare fact that S has some desire that P. Further, with this expectation in place, if someone *does* assert "S desires that P", we are in a position to infer more than the mere fact that S has some desire that P. We should infer that this assertion has been made because this desire is especially worth noting, most probably because it is very strong, or else because it is of some notable kind. For example, if you tell me that you want to go to the pub, I might reasonably infer that you don't merely mean that you have some desire to go (who doesn't?), but rather that you have a fairly strong desire to go. Or if I ask about your new job and you happily tell me that you now want to go to work, I might reasonably infer that you don't merely mean that you have some desire to do so, but rather that desire to do so not merely because of the salary.

Vice versa – and more surprisingly – with such expectations in place, we might sometimes deny that someone has a desire if the desire is only weak or else not of any notable kind. For example, you might deny that you want to go to the pub, really meaning by this that the desire in question is very heavily outweighed. And you might deny that you want to go to work, really meaning by this only that you gain no pleasure from your work, even though you would, if pressed,

acknowledge that you want to go in order to pay your rent. In this way, we will sometimes verbally deny that people have desires even though we know that they do in fact have them.

With these two points noted – the ambiguity effect, and the quantity effect – I now turn to two ways in which they can lead us into drawing mistaken inferences about desire.

2. The Belief/Desire Theory of Motivation

The belief/desire theory of motivation says that a person can be motivated to φ only by the desire that P combined with the belief that φ ing will secure P (Davidson 2001b; Smith 1994: 92-129; Stalnaker 1987:15).² This view is widely agreed to be attractive, but faces a simple common-sense objection (e.g. Fisher 2011:136-7; May 2013; Platts 1997: 256; Searle 2001: 169-170; Shafer-Landau 2003:123; Skorupski 2010:241; see also Schueler 1995): don't people sometimes do things that they don't want to do? For example, as we ordinarily talk, it seems that many of us don't want to get out of bed in the mornings, don't want to do chores around the house, and don't want to fulfil costly moral promises. Nonetheless, we are clearly often motivated to do these things.

In response, we might suggest that the belief/desire theory of motivation should be understood as using the term "desire" in some broad stipulative sense (see e.g. Davidson 2001a:3-4; cf. Schueler 1995: 29-30). But this threatens to trivialise the belief/desire theory of motivation by stipulating the meaning of "desire" so as to secure the truth of the theory (e.g. Platts 1997: 256; cf. Nagel 1970). Our two effects above suggest a more principled and illuminating response.

First, we should recall the ambiguity effect. We might *say* that we don't want to do these things, but plausibly this expresses a desire *not* to do them, and that is consistent with our *also* desiring to do them, at least for instrumental reasons. For example, Shafer-Landau imagines someone saying

² This theory is sometimes labelled "The Humean Theory of Motivation". I avoid this label in order to make clear that my focus is not on any other Humean claim(s), such as the claim that desires are not subject to rational assessment.

“I didn't want to (stand in front of that bullet, accept the blame, remain chaste), but duty called” (Shafer-Landau 2003:123; for similar examples, see Fisher 2011:136; Skorupski 2010:241). But such assertions would most plausibly be understood as conveying a desire *not* to do such things, and that is consistent with your *also* wanting to do them (out of a sense of duty, even if for no other reason). As a result, the belief/desire theory of motivation is (surprisingly) consistent with the possibility that you might be motivated to stand in front of a bullet, accept blame, or remain chaste, and yet truly assert that you don't want to do those things.

Second, we should recall the quantity effect: there are pragmatic reasons why we might be disinclined to assert the presence of some desires. For example, I hardly need to be told that you have some desire to get out of bed in the mornings when it is common knowledge that doing so is instrumental to so many things that I know you want (e.g. food). As a result, you might be disinclined to assert that you want to get out of bed in the mornings. But this disinclination can be explained by conversational pragmatics in a way that permits that you do nonetheless have this desire. As a result, the belief/desire theory of motivation (surprisingly) permits that you might be motivated to get up and yet be disinclined to assert that you want to do so.

The ambiguity effect and the quantity effect show how the belief/desire theory of motivation is consistent with cases like those above. Your assertions might suggest that you sometimes do unpleasant things that you don't want to do, but we can insist, for principled reasons, that such assertions are misleading.

3. The Relationship Between Normative Judgement and Desire

One crucial question in moral psychology surrounds the relationship between normative judgement and desire. Many claim that our normative judgements are related to our desires, but the intimacy of this relationship is hotly debated. At one extreme, some claim that the presence of a normative judgement entails the presence of a related desire, perhaps because one of these states can be reduced to the other (e.g. Blackburn 1998; Davidson 2001b:22-23; Gibbard 2003;

Gregory 2017). A natural way of testing such claims is to point to cases where agents apparently judge an act favourably but don't desire to do it (e.g. Scanlon 1998:39; Smith 1994:117-125; Stocker 1979).³ But here we should be careful: our effects above might lead us to misidentify some putative cases of this kind.

Before I explain this, I should first clarify our focus. Our question is *not* about the extent to which someone can judge an act favourably and yet fail to be *motivated* to do it. The relationship between normative judgement and motivation (“judgement internalism”) has been heavily discussed (for a review, see Björklund et al. 2012). But we are interested only in the relationship between normative judgement and *desire*. If the presence of a desire fails to guarantee the presence of motivation, then even the most extreme view about the relationship between normative judgement and desire fails to guarantee any necessary connection between normative judgement and motivation. And there are at least some grounds for thinking that the presence of desire does fail to guarantee the presence of motivation: certainly, on the standard Humean theory desires generate motivation only if they are combined with suitable means-ends beliefs. We might also think that depression can leave our preferences intact but rob us of the motivational energy needed to pursue those goals (Swartzler 2015). Perhaps there are still other conditions which need to be met for our desires to generate motivation (Gregory 2017; Sinhababu 2009, 2017; Toppinen 2017). Since the relationship between desire and motivation is indirect and potentially complex, in what follows I stay wholly focused on our central issue, which is the relationship between normative judgement and desire, not motivation.

Back to our main theme. I claim that the two effects described in section 1 might mislead us about the relationship between normative judgement and desire. I illustrate the point with two examples:

³ I say “judge *favourably*” so as to abstract away from issues about whether our desires are supposedly related to normative judgements about *values*, about *normative reasons*, about what we *ought* to do, or something else. For related discussion, see (Gregory 2013).

First, expressivism. On at least one canonical way of thinking about expressivism, it is committed to there being a *very* intimate relationship between normative judgement and desire. For according to one canonical formulation of expressivism, it commits us to the claim that normative judgements just *are* desires, and as such that assertions such as “murder is wrong” express our desires. In a recent paper, Jack Woods argues against all forms of expressivism (2014). To understand Woods’ argument, first take *the rain assertion*: “It’s raining but I don’t believe it’s raining”. The rain assertion seems incoherent since the second conjunct reports the absence of the very state of mind its first conjunct expresses (for more on such assertions, see Green and Williams 2007). Woods then examines *the murder assertion*: “Murder is wrong, but I don’t disapprove of murder”. Woods claims that if expressivism were true, the murder assertion should strike us as incoherent in the very same way as the rain assertion, since according to expressivism its second conjunct reports the absence of the very state of mind its first conjunct expresses. But since the murder assertion does not strike us as incoherent in this way, expressivism must be false.

To assess Woods’ argument we should first note that expressivism – or at least the canonical version of it that I described – analyses normative judgements as *desires* (wants), whereas the murder assertion made reference to *disapproval*. To this extent, it is open to my kind of expressivist to claim that a fairer test for their view would be an assertion that makes explicit reference to desires, as such. Woods acknowledges this when he writes: “I have also only canvassed the most straightforward attitudes to insert on the right-hand side of the conjunction. However, it is easy to see from my examples how to construct similar cases with ... other, less plausible, accounts of the conative or affective attitude expressed by moral assertions” (Woods 2014:6). I think that Woods must be imagining more remote possibilities than the one I suggest when he says that other options will be “less plausible”: many canonical formulations of expressivism make reference to our desires (e.g. Blackburn 1984:187-9), and it is not at all obvious that such views are less plausible than those that make reference to disapproval.

So instead of the murder assertion above, we should assess the canonical version of expressivism by appeal to *the desire assertion*: “Murder is wrong, but I don’t want not to murder”.⁴ Is that assertion incoherent in the manner that the rain assertion is? Plausibly, no. But the quantity effect can explain why this is true. The desire assertion might very commonly be heard as expressing the absence of just one particular kind of desire not to murder (i.e. a self-interested desire), and that is consistent with one’s also having other desires not to murder. Perhaps the ambiguity effect also plays a role here: perhaps we might interpret the desire assertion as reporting the presence of a desire not to not murder (i.e. to murder), and expressivism is certainly consistent with the possibility that agents may have desires that conflict with their moral commitments.

These points may be easier to see if we – unlike Woods – focus not on assertions about wrongness, but instead on assertions about goodness. Take the *philanthropy* assertion: “philanthropy is good, but I don’t want to give my money away”. This sentence surely sounds perfectly coherent, unlike the rain assertion. But the quantity effect and the ambiguity effect give the expressivist ample explanatory resources to explain why this is true. They might claim that this assertion reports only that the agent lacks some notable – self-interested – desire to give their money away, and that is consistent with their also desiring to give their money away, a desire they express with the first conjunct of the philanthropy assertion. And the ambiguity effect permits them to insist that such an assertion sounds coherent because it is naturally heard as reporting the presence of a desire not to give one’s money away, not the absence of a desire to give it away, and again, this is consistent with expressivism which permits that agents can have desires that conflict with their moral commitments.

⁴ I use “want” rather than “desire” since “want” is the more common English word, and is therefore more helpful for eliciting intuitions about the coherence of certain assertions.

In short, Woods is wrong when he suggests that there is no expressivist-friendly explanation as to why the various normative assertions sound more coherent than the rain assertion: principles of charity encourage us to hear the relevant claims about desire as restricted in scope, or else encourage us to disambiguate such assertions in a manner that renders them coherent. It is not obvious that any similar interpretive charity is possible with respect to the rain assertion. So at least one way of formulating expressivism has the resources to escape Woods' objection, by appealing to the quantity effect and the ambiguity effect.

I now turn to my second illustration of the effects above as they apply to claims about the relationship between normative judgement and desire. It is illustrated by the following quote from Scanlon (1998:39; see also Baker 2014: 25 for a similar example):

“I might see something good about drinking a glass of foul-tasting medicine, but would not therefore be *said* to have a desire to do so” (my emphasis)

Return to our two effects above. It is surely true that you might think the medicine is good and yet say “I don't want to take the medicine”. But this assertion is ambiguous, and is plausibly intended to convey the presence of a desire not to take the medicine rather than the absence of any desire to take it. At least, this is the natural way of reading the case if the foul-taste of the medicine is at all relevant: such a taste might *outweigh*, but would not *eliminate*, a desire to drink the medicine, and so bears on whether you desire *not* to take it, not on whether you desire to take it. If this assertion indicates only the presence of a desire *not* to take the medicine, it is perfectly consistent with the claim that your value judgement here generates a desire to drink the medicine.⁵

⁵ Neil Sinhababu (2009:490-491) says something somewhat similar about a similar case of Scanlon's, though he relies on the unconventional claim that we should distinguish two different *kinds* of desire – positive desires, and aversions – each with their own properties (2009:470, and 490-491). My claims in the text require no such metaphysical distinction but instead only a much more modest claim about the grammar of English.

Depending on the exact envisaged circumstances, the quantity effect may also arise here. If the beneficial potential of the medicine is common knowledge, that might well make you disinclined to assert that you want to take the medicine just because this desire is obvious. Perhaps we aren't supposed to imagine the case in this way, but if we exclude this possibility then Scanlon's claim is itself far less plausible. If, for example, the foul-taste of a liquid is well-known but its medical potential is not, you might well be very strongly inclined to point out your surprising desire to drink it. It is only if we think of the liquid as *medicine* that Scanlon's claim may seem plausible, and a natural explanation of this is that labelling it as such renders the desire to drink it too obvious to be worth stating.

Our effects might make you inclined to assert that you "don't want" to take the medicine and disinclined to assert that you "do want" to take the medicine, but for all that, you might well want to take the medicine just as you judge you should. We should be careful when drawing conclusions about the relationship between normative judgement and desire from cases like the one above.

How widely do these points apply? The ambiguity effect will apply in many cases: whenever we are given a supposed case where an agent judges P favourably but "fails to desire P", we should always be careful as to whether the intuitive claim is that the agent desires that $\neg P$ or that the agent fails to desire that P: only in the latter case will it bear on the relationship between judging P favourably and desiring that P.

The quantity effect will also apply in many cases, but not all. It applies most straightforwardly in cases where your desire to do something is obvious but your desire not to do it is less so. In such circumstances, you might be inclined to highlight the latter desire rather than the former, and to that extent your assertions might suggest that you fail to want something even though you in fact do want it. For example, you might judge that you ought to go to work, but be disinclined to

assert that you want to do so. Here your disinclination to assert that you want to go might be best explained by the obviousness of that desire rather than its absence.

In summary, the ambiguity effect and the quantity effect show that we should be careful when investigating the relationship between normative judgement and desire. Some examples we might use to show a discontinuity between these states might be misleading, and to that extent the relationship between normative judgement and desires may be closer than it seems.

Conclusion

In this paper I highlighted two noteworthy features of assertions about desire: that assertions about what we “don’t want” are ambiguous, and that pragmatic influences might encourage us to fail to assert, or even deny, the presence of desires that we do believe exist. These effects might lead us to wrongly deny the belief/desire theory of motivation, and to wrongly understate the relationship between normative judgement and desire.

Bibliography

- Baker, Derek. 2014. ‘The Abductive Case for Humeanism Over Quasi-Perceptual Theories of Desire’. *Journal of Ethics and Social Philosophy* 8 (2): 1–29.
- Björklund, Fredrik, Gunnar Björnsson, John Eriksson, Ragnar Francén Olinder, and Caj Strandberg. 2012. ‘Recent Work on Motivational Internalism’. *Analysis* 72 (1): 124–37. doi:10.1093/analys/anr118.
- Blackburn, Simon. 1984. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford: Clarendon Press.
- . 1998. *Ruling Passions: A Theory of Practical Reasoning*. Oxford : Oxford ; New York: OUP Oxford.
- Broome, John. 2013. *Rationality through Reasoning*. Chichester, West Sussex ; Malden, MA: Wiley Blackwell.
- Chan, David K. 2004. ‘Are There Extrinsic Desires?’ *Noûs* 38 (2): 326–50. doi:10.1111/j.1468-0068.2004.00472.x.
- Davidson, Donald. 2001a. ‘Actions, Reasons, and Causes’. In *Essays on Actions and Events*, 3–19. Oxford: Oxford University Press.
- . 2001b. ‘How Is Weakness of the Will Possible?’ In *Essays on Actions and Events*, 21–42. Oxford: Oxford University Press.
- Fisher, Andrew. 2011. *Metaethics: An Introduction*. Durham: Acumen.
- Gibbard, Allan. 2003. *Thinking How to Live*. Cambridge, Mass.: Harvard University Press.

- Green, Mitchell S, and John N Williams. 2007. *Moore's Paradox: New Essays on Belief, Rationality, and the First Person*. Oxford; New York: Clarendon Press ; Oxford University Press.
- Gregory, Alex. 2013. 'The Guise of Reasons'. *American Philosophical Quarterly* 50 (1): 63–72.
- . 2017. 'Might Desires Be Beliefs About Normative Reasons?' In *The Nature of Desire*, edited by Julien A Deonna and Lauria Federico, 201–17.
- Grice, Paul. 1989. 'Logic and Conversation'. In *Studies in the Way of Words*, 22–40. Cambridge, Massachusetts: Harvard University Press.
- May, Joshua. 2013. 'Because I Believe It's the Right Thing to Do'. *Ethical Theory and Moral Practice* 16 (4): 791–808.
- Nagel, Thomas. 1970. *The Possibility of Altruism*. Princeton, N J: Princeton University Press.
- Platts, Mark de Bretton. 1997. *Ways of Meaning: An Introduction to a Philosophy of Language*. Cambridge, Mass.; London: The MIT Press.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, Mass.; London: Belknap.
- Schroeder, Mark Andrew. 2007. *Slaves of the Passions*. Oxford; New York: Oxford University Press.
- Schueler, G. F. 1995. *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge, Mass.; London: MIT Press.
- Searle, John R. 2001. *Rationality in Action*. Cambridge, Mass.: The MIT Press.
- Shafer-Landau, Russ. 2003. *Moral Realism: A Defence*. Oxford: Clarendon.
- Sinhababu, Neil. 2009. 'The Humean Theory of Motivation Reformulated and Defended'. *Philosophical Review* 118 (4): 465–500.
- . 2017. *Humean Nature: How Desire Explains Action, Thought, and Feeling*.
- Skorupski, John. 2010. *The Domain of Reasons*. Oxford: Oxford University Press.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Oxford University Press.
- Stalnaker, Robert. 1987. *Inquiry*. Cambridge, Mass.: MIT Press.
- Stocker, Michael. 1979. 'Desiring the Bad: An Essay in Moral Psychology'. *Journal of Philosophy* 76 (12): 738–753.
- Swartzter, Steven. 2015. 'Humean Externalism and the Argument From Depression'. *Journal of Ethics and Social Philosophy* 9 (2): 1–16.
- Toppinen, Teemu. 2017. 'Enkrasia for Non-Cognitivists'. *Ethical Theory and Moral Practice*, 1–13.
- Woods, Jack. 2014. 'Expressivism and Moore's Paradox'. *Philosophers' Imprint* 14 (5): 1–12.