Rethinking neutrality: a conceptual analysis

Matt Watson
Lecturer, TC Beirne School of Law, University of Queensland

This article attempts to determine whether there exists a coherent, plausible, and ultimately compelling explication of what it is to act neutrally. I argue that there is – an account I label neutrality of volition, and according to which an actor acts non-neutrally where she either acts for the purpose of differentially helping or hindering a particular party in a given contest, or acts or in the belief that there is a substantial likelihood that her action will have this effect. Along the way, I suggest that political philosophers concerned with whether justice requires that the state’s laws and policies be publicly justifiable, as well as legal commentators who note that oftentimes laws of general application have disparate impacts, would do well to cease framing their arguments in the language of neutrality. I conclude by arguing that debate over the proper interpretation of neutrality is not merely a matter of semantics. Having identified an account of neutrality that aligns with our ordinary understanding of the concept, and which is also internally consistent, we have access to a conceptual tool that we can use to make better sense of a wide array of actions in the political sphere and beyond, while also avoiding an unhelpful conflation of neutrality with distinct concepts such as fairness, public justifiability, anti-perfectionism, equal impact, and non-discrimination.

Keywords: neutrality, conceptual analysis, political philosophy, legal philosophy

1 INTRODUCTION

What does it mean to act neutrally? We are familiar with canonical examples of neutrality such as that of Switzerland vis-à-vis foreign military conflicts. It is also common to hear natural persons described as neutral in respect of some contest going on around them. We might say, for instance, that a good judge must be neutral between the parties in the case before her. Or we might remark that a parent has remained neutral in a quarrel involving his two children. Legal and political philosophers, as part of the so-called ‘neutrality versus perfectionism’ debate, also argue for or against the idea that justice requires the state to be neutral between competing conceptions of the good life.1 In addition, legal theorists and political activists commonly insist that

some facially neutral laws are anything but, in light of how they disproportionately disadvantage marginalized persons.

What, if anything, unites instances of neutrality? It seems clear that many of our judgments about the neutrality or non-neutrality of an action or actor in a given scenario are made without reference to an explicit theory of what it is to act neutrally. The question that this article seeks to answer is whether such a comprehensive account of neutrality is even available. Is neutrality an ‘essentially contested concept’, such that it is impossible to lay out a single conception of neutrality that captures all of our judgments and intuitions about what neutrality is and how it can be realized in practice? Or can we articulate a precise, formal conception of neutrality that is able to account for all of our judgments about the neutrality or non-neutrality of various possible actions, while remaining internally consistent?

I answer the latter question in the affirmative. I attempt to construct a sound theory of neutrality by beginning with the intuition that acting neutrally means refraining from differentially helping or hindering one party over others. This idea is then refined over a multi-step process in order to advance a conception of neutrality – which I label ‘neutrality of volition’ (or, ‘volitional neutrality’) – that tracks our judgments about the neutrality or non-neutrality of various hypothetical actions. I then contrast neutrality of volition with competing interpretations of neutrality common in the political and legal philosophy literature – namely, neutrality of effects, neutrality of intentions, neutrality of justification, and Alan Patten’s ‘neutrality of treatment’. I conclude the article by arguing that debate over the proper interpretation of neutrality, although by definition a semantic argument, is not merely a matter of semantics. In seeking to identify an account of neutrality that aligns with our ordinary understanding of the concept, and which is also internally consistent, we seek access to a conceptual tool that we can use to make better sense of a wide array of actions in the legal and political spheres and beyond, while also avoiding an unhelpful conflation of neutrality with various other, distinct concepts such as fairness, impartiality, equal impact, anti-perfectionism, and public justifiability. By getting clear on the correct way to conceive of neutrality, that is, we avoid diluting the explanatory power that the concept, correctly understood, can offer.

2 VOLITIONAL NEUTRALITY

We will flesh out our conception of neutrality in a series of steps. By engaging in this iterative process we can revise our working account of neutrality in the light of possible objections and counter-examples, until we are left with a conception of neutrality that is able to meet all such challenges.

Perfectionism (Oxford University Press 1993); and George Sher, Beyond Neutrality: Perfectionism and Politics (Cambridge University Press 1997).

2. I will have a word to say about the distinction between acting neutrally and being neutral below, after we have fully articulated the volitional neutrality account of what it is to act neutrally.

3. According to W B Gallie, ‘Essentially Contested Concepts’ (1956) 56 Proceedings of the Aristotelian Society 167, 169, essentially contested concepts are ‘concepts the proper use of which inevitably involves endless disputes about their proper uses on the part of their users’, which disputes (according to John N Gray, ‘On the Contestability of Social and Political Concepts’ (1977) 5 Political Theory 331, 344) ‘cannot be settled by appeal to empirical evidence, linguistic usage, or the canons of logic alone’.
2.1 What Does it Mean to Act Neutrally?

As the starting point in our attempt to build up a comprehensive account of neutrality, consider the following proposition, which appears to capture a basic intuition we have about acting neutrally.

\( N_1 \): Acting neutrally means not acting in a way that helps or hinders one party more than another.

In order to assess whether \( N_1 \) amounts to an adequate conception of neutrality, we will need to provide an explanation of what it is to ‘help’ a ‘party’. Imagine, for example, that you are walking down a street and the person in front of you unknowingly drops his wallet. You kindly pick it up, call out to him, and return the wallet. Clearly, your returning the wallet has helped the other person. But it seems equally clear, I think, that it would be inapt to say that since you helped that person, and did not provide equivalent help to the pedestrian behind you, say, you therefore acted non-neutrally. Consider, along the same lines, a corporate CEO who decides to promote Employee A to a lucrative managerial position, as a reward for achieving the highest sales figures over the last quarter. The boss’s decision here can easily be regarded as one that (broadly speaking) helps Employee A. But can we safely say that the boss therefore acted non-neutrally in promoting Employee A? The answer, it seems, is that it depends on how we frame the question. If we have in mind the question of whether the employer acted non-neutrally in promoting Employee A when she might conceivably have given the manager job to me – a perfect stranger with no connection to the organization, let’s say – then it would seem quite odd to label that action as non-neutral merely because it helped A more than it helped me. If, instead, we focus on the question of whether she acted non-neutrally by giving the promotion to Employee A and not her other employees B and C, then it becomes at least plausible to say that the action was non-neutral, on the grounds that it helped A more than it helped A’s co-workers B and C. What makes the CEO’s action amenable to being described as ‘non-neutral’, then, is not that it benefited A but not me (although that is true), but rather that it benefited A but not B or C. In selecting Employee A but not Employee B or C for the promotion, the CEO has helped one but not all of her employees who were in a sense competing for the job. In offering the position to Employee A instead of to me, by contrast, she is simply opting not to provide an unearned windfall to a stranger.

What all this suggests, then, is that in order for the notion of neutrality to have purchase, it must be the case that the parties among whom a given actor may be said to act neutrally or non-neutrally stand in a particular sort of relationship to one another. They must each be participants (knowingly or unknowingly) in what could accurately be described as a contest (broadly defined).

The foregoing observations reveal that \( N_1 \) is in need of revision. Consider, therefore \( N_2 \).

\( N_2 \): Acting neutrally means not acting in a way that differentially helps or hinders a party to a contest.

The definition of neutrality contained in \( N_2 \) reflects the apparently widespread sense that the essence of neutrality is refraining from helping (or hindering) one party to a
dispute more than one helps (or hinders) all the other parties. At this point, it will be helpful to nail down what is meant by helping or hindering a party to a contest. I will understand an actor to ‘help’ a party to a contest when the actor, generally via the provision of some sort of aid or assistance, increases that party’s chances of success in the contest. Conversely, ‘hindering’ a party will be understood as decreasing that party’s chances of success, ordinarily by way of saddling that party with some kind of obstacle or disadvantage. Now, in the context of some zero-sum competitions, the very act of helping party D may necessarily and directly hinder party E, such as where an actor intervenes to pull on D’s side of the rope in a two-person tug of war between D and E. In most contests, though, if I give an advantage to F but not to G or H – if, for example, F, G, and H are facing off in an endurance race and I provide F but not G or H with a cup of water part way through the race – that will not directly hinder either G or H, but it may well still decrease both G’s and H’s chances of success in the race, since they have both been denied a benefit that was conferred on another competitor.

In turn, we will need to know how to understand a party’s ‘chances of success’ in the relevant contest. In a zero sum, two-person contest, each party’s chances of success are their odds of defeating the other and thereby winning the contest. In other sorts of contests, however, one party’s success will not necessarily be another party’s failure. As a general definition applicable to any contest, then, a party’s chances of success amounts to the likelihood that that party will meet with success, however ‘success’ is defined in the context of that specific contest – whether that be bettering one’s competitors, achieving some objective benchmark, impressing one’s audience, or the like.

N2, however, is susceptible to the following counter-example. Imagine for the moment that a boxing trainer attempted to remain neutral in a sparring match between two fighters by studiously avoiding intervening in the contest in any way, except to give each of them an exactly equal amount of some sports drink between rounds. Totally unexpectedly, one of the fighters had an allergic reaction to a certain chemical in the drink, which caused him to become weak and sluggish, thus greatly reducing his chances of winning the match. Here, since the differential effect that the trainer’s action caused was not something he in any way foresaw or aimed at when he acted, we do not think he acted non-neutrally. An account of neutrality that trains its attention merely on the effects of the actor’s action, however, would force us to say that he did.

Let us, accordingly, consider N3.

N3: Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest.

It is worth pausing for a moment here and noting that the notion of differentially helping or hindering remains in need of some further elaboration. On one interpretation, for instance, an actor might be understood to intend to differentially help or hinder a party if the actor intends, via their action, to cause that party’s (but not all

4. See Alan Montefiore, Neutrality and Impartiality: The University and Political Commitment (Cambridge University Press 1975) 5–6; Raz (n 1) 110–33.
5. We could imagine, for instance, a contest among five people for four jobs. Here, it cannot be said that one person’s securing a job implies the failure of the one unsuccessful applicant. Further, while in some contests what counts as success might be nothing short of outright victory over an opponent, in others success might instead be a matter of, say, performing at an objectively high calibre, or entertaining one’s audience. As such, it is possible, in certain sorts of contests, for all the parties to meet with success.
parties’) chances of success to be, respectively, greater or less than they objectively were at the moment before the actor performed the action. Let this idea be $N4$.

$N4$: Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest, where an actor is understood to intend to help a party whenever the actor intends that the party will face, as a result of the actor’s action, a chance of success that is greater than the chance of success that party faced at the moment before the actor performed the action.$^6$

$N4$ should be rejected. To see why, consider a scenario in which an actor is observing two people playing a game of chess. Suppose that player I has player J right where she wants him, and will checkmate him in three moves. Player J cannot avoid this fate, but is presently unaware that he is about to lose. Suppose further that our observer is likewise unaware of J’s inevitable defeat. As far as she can make out, the match is a toss-up. She wishes to see I win, however, and so, with the intention of guaranteeing that result, she surreptitiously removes from the board one of player J’s pawns, wrongly believing that this will ensure that player I will capture J’s newly exposed king with I’s next move. She therefore estimates, in her ignorance, that by stealing the pawn from J she has increased I’s odds of winning from roughly even to 100 percent. In this scenario, since it cannot be said that the actor intended her action to bring I’s chances of success to a level that was higher than the chances of success that I actually faced at the moment before the actor acted – the chance of success I faced at the moment before the actor acted, and the chance of success the actor intended her act to see I face thereafter, being instead an identical 100 percent – we are thus forced to deny, pursuant to the conception of neutrality laid out in $N4$, that the actor intended to differentially help player I. $N4$ would therefore force us to say that the actor acted neutrally.

$N4$ is thus clearly in need of revision. Consider the following refinement.

$N5$: Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest, where an actor is understood to intend to help a party whenever the actor intends that the party will face, as a result of the actor’s action, a chance of success that is greater than the chance of success the actor believed the party faced at the moment the actor performed the action.$^7$

$N5$ improves upon $N4$ by focusing on the difference, if any, between the chances of success the relevant actor intended a party to face following the actor’s action, and the chances of success the actor subjectively believed that party to be facing at the time the action was performed. $N5$, for example, allows us to come to the obviously true conclusion that the actor in the chess example above non-neutrally intended to differentially help player I, since she intended her action to cause I to enjoy a chance

6. An intention to hinder a party would of course be understood in complementary fashion as an intention on the part of the actor that the party will face, as a result of the actor’s action, a chance of success that is less than the chance of success that party faced at the moment before the actor performed the action.

7. The fact of multiple post-action time points is not a problem given our metric of ‘chance of success’. That is, if an actor intends to cause a party to lose a few battles post-action before later gaining the upper hand and winning a string of later battles and ultimately the war, we should understand the actor as intending to raise the party’s chances of success to whatever he estimates are the party’s overall odds of winning the war.
of success post-action (100 percent) that is greater than the chance of success she
understood I to be facing at the time the action was performed (50 percent).

For ease of reference going forward, we can shorten N5 to read:

N5: Acting neutrally means not acting with the intention of differentially helping or
hindering a party to a contest.

This definition of neutrality is identical to N3, although in moving from N3 to N5 we
have given a vital clarification of how N3’s language of ‘differentially helping or
hindering’ is to be read.

The definition of neutrality provided by N5 is, however, ambiguous when it comes
to the notion of ‘acting with the intention of differentially helping or hindering a
party’. For example, on one interpretation, an intention to help or hinder a party might
be understood as requiring that the actor performs an action ‘for the purpose’ of
thereby increasing a party’s chances of success, and in that sense aims at that result.
Let this be N6.

N6: Acting neutrally means not acting for the purpose of differentially helping or hindering a
party to a contest.

N6 too, however, is vulnerable to counter-examples. Consider, for instance, the
following scenario. A civil war rages between Sinhalese and Tamils in Sri Lanka. The
government of Pakistan has no preference when it comes to who will prevail. Instead,
the Pakistani government views the Sri Lankan war as merely an excellent opportunity
for it to demonstrate to its rival India the awful power of a new bomb that the Pakistani
army has recently developed. With the aim of making crystal clear to India just how
devastating such a bomb would be to Indian ground forces, were they to one day
invade Pakistan, the Pakistani government orders its air force to drop one of these new
bombs on an encampment of troops in Sri Lanka. The government orders the bomb
dropped on a site near the northern tip of Sri Lanka, purely because this is the closest
location to the Pakistani air force base from which the bomber will depart that is
currently occupied by deployed troops. The site, let us assume, happens to be a key
Sinhalese army base and command centre, a fact of which the Pakistanis are well
aware. Let us, further, stipulate that Pakistan takes no other action in respect of the Sri
Lankan conflict. While it knows that its bombing of the Sinhalese forces will cause
massive destruction and an extraordinarily large number of casualties, it takes no
action to inflict a similar blow against the Tamils.

What should we say about the neutrality or non-neutrality of the Pakistani
government in this example? I believe we should say that Pakistan has acted
non-neutrally. True, it did not act for the purpose of differentially helping one side in
the Sri Lankan civil war, but neither did it refrain from intervening in the conflict
despite expecting that its intervention would have precisely that effect; neither, that is,
did it exhibit the sort of forbearance that we do, it seems, require of an actor before we
are willing to say that they have remained neutral.

One way in which to account for the foregoing would be to revise N6 as follows.

N7: Acting neutrally means not acting with the intention of differentially helping or hindering a
party to a contest, where an intention to differentially help or hinder a party is understood to

---

exist whenever the actor acts in the belief that there is a substantial likelihood that their action will differentially help or hinder a party.

Imagine, however, that a tennis fan is attending a match between Rafael Nadal and Roger Federer. As the match wears on and things look increasingly dire for his favoured Federer, the fan yells out, across the stadium and loud enough for both players to hear, ‘Come on, Roger!’ He does not actually believe that shouting this encouragement will increase Federer’s chances of winning the match, but does so – reasoning that there is no harm in trying – just on the off chance that it might. Here, I think, we want to say that this fan has therefore acted non-neutrally. However, N7 tells us that the fan’s action was a neutral one, since he did not act in the expectation that his action would cause a differential effect. The upshot, in short, is that any satisfactory conception of neutrality will have to account for our belief that whenever an actor acts with the specific aim of differentially benefiting a particular party to a contest (even where he or she does not expect that result to eventuate), the actor acts non-neutrally. Consider then the following revision.

N8: Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest, where an intention to differentially help or hinder a party is understood to exist whenever the actor either acts for the purpose of differentially helping or hindering a party, or acts in the belief that there is a substantial likelihood that their action will differentially help or hinder a party.

According to N8, where an actor does not believe that her act will differentially benefit a party to a contest, her act will still be non-neutral if it is taken for the reason that it might do so – i.e., if it is taken because the actor believes that performing the act will make that result more likely than it otherwise would be. N8 therefore asserts that acting neutrally requires both refraining from performing actions one expects will cause a differential benefit among the relevant parties, and refraining from performing actions aimed at having that effect. In other words, both sorts of forbearance are necessary for an action to be neutral, and thus neither one is on its own sufficient to make an action neutral.9 We can say, then, that N8 defines a neutral act as one where the relevant actor acts with an intention – broadly construed – to have an equivalent effect on all parties.10

N8, however, threatens to label as non-neutral actions that our ordinary understanding of neutrality insists are in fact neutral. Imagine, for instance, that some club is holding an election to appoint its new president. Two members are vying for the position. As they are each equally popular among the other members of the club, the

9. There are, then, three possible scenarios, all of which, where they obtain, are sufficient to make out ‘intent’ as that term is expansively understood on the volitional neutrality conception. Firstly, an actor might believe that her action will cause a differential benefit and act so as to bring about that result. Secondly, the actor might perform an action in the belief that doing so will cause some differential benefit, albeit without acting for the purpose of bringing about that benefit but rather due to some other motivating reason. Thirdly, the actor might choose to perform an act with the aim of causing a differential benefit, despite not believing that the performance of that act is likely to cause that effect. (An actor acts neutrally only in the fourth and final possible scenario in which her action is neither performed for the purpose of causing a differential effect, nor taken in the belief that the action is likely to cause such a result.)

10. This includes, but is not limited to, cases in which the actor intends to have no effect at all on the parties.
outgoing president estimates that at present both candidates face about a 50 percent chance of election. The outgoing president is also aware that neither candidate has plans to put forward to the membership any substantive proposals for how they would manage the club’s activities if elected. The current president has, however, met privately with both candidates, and has been told by each about the sorts of objectives they would pursue as president. The outgoing president happens to know that Candidate K’s vision would resonate strongly with a majority of the members of the group, whereas Candidate L’s would not. Against this background, and purely out of a desire to see that the voting members make an informed decision about who should be their next president, the outgoing president decides to host an ‘all candidates’ meeting’, at which both candidates can present their platforms.

N8, it seems, would have to describe the outgoing president’s convening of the all candidates’ meeting as a non-neutral act, since she acted in the belief that as a result of the meeting a sizeable proportion of members initially planning on voting for Candidate L would instead decide to vote for Candidate K, in light of what they would regard as her superior platform. However, this strikes most of us, I think, as incorrect. It would be inappropriate to label the outgoing president’s actions as non-neutral, since while she did expect that holding the meeting would lead to K’s chances of winning the presidency increasing from 50 percent to, say, 80 percent, she made no attempt to sway the minds of voters in one direction or the other, and instead merely gave to each candidate equal time and opportunity to inform the voters of their respective visions for the club’s future. We should, then, offer a refinement to N8.

N9: Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest, where an intention to differentially help or hinder a party is understood to exist whenever the actor either acts for the purpose of differentially helping or hindering a party, or acts in the belief that there is a substantial likelihood that their action will, sufficiently directly, cause a party to be differentially helped or hindered.

Clearly, N9 contains some ambiguity about the circumstances in which an actor’s action will, in a ‘sufficiently direct’ manner, ‘cause’ a differential effect. Providing a precise definition here is not an easy task. It is not surprising, for instance, that we should meet with difficulty in drawing a line between actions that cause their effects in a sufficiently direct manner, and those that, for our purposes, do so too indirectly, given that causation is a notoriously difficult philosophical concept. What is more important to note, however, is the fact that this sort of ambiguity or uncertainty about the requisite directness of causation seems to be a feature of our ordinary understanding of neutrality. We are able to invoke the concept of neutrality in everyday life without having any particularly well-considered view about just how directly a given

---

11. It is worth noting that Alan Patten’s ‘neutrality of treatment’ conception of neutrality, which we will discuss at length below, also distinguishes between direct and indirect effects. For example, Patten explains in Equal Recognition: The Moral Foundations of Minority Rights (Princeton University Press 2014) 116 that neutrality of treatment focuses only on ‘the direct effect that a policy has on making a conception of the good more or less realizable … [and] does not look at the further, indirect effects that arise because of the way in which people react to that relative treatment’.

12. This sort of vagueness, I would submit, is also apparent in Alan Patten’s contention that ascertaining the neutrality of a policy involves asking whether that policy ‘can be expected, in conjunction with other necessary inputs, to make a particular conception of the good more successful’: ibid 115 (emphasis added).
action must cause a differential benefit before it is appropriate to conclude that the actor who foresaw that benefit acted non-neutrally. It does seem, however, to be part of our ordinary conception of neutrality that this distinction – between effects that are sufficiently directly caused, and those that are caused instead only quite indirectly – matters. (This distinction between direct and indirect effects, for instance, explains our judgment that although the club president acted neutrally in calling the meeting, Pakistan acted non-neutrally in bombing the Sinhalese army base.) The inherent vagueness here is thus a fact about our everyday understanding of neutrality that should be taken as a given – something that a careful exposition of the concept (such as this article aims to provide) should identify, rather than attempt to remove.

However, if pressed for an account – one that coheres with our judgments about the neutrality or non-neutrality of the sorts of real-world examples we’ve considered above, and is also capable of relatively precise application – of how one can identify differential effects that are ‘sufficiently directly caused’ for the purposes of assessing the neutrality or non-neutrality of actions, I would offer the following. An action – \( p \) – does not cause a differential impact – \( q \) – on the relevant parties in a sufficiently direct way if it would be misleading to say, without qualification, that ‘\( p \) caused \( q \)’. For example, where \( p \) is not a proximate cause of some differential effect (even if it is a ‘but-for’ cause), and instead sets up background conditions against which the intervening actions and choices of other actors give rise to the differential effect, then there would be something misleading about claiming simply that ‘\( p \) caused the effect’. It would be more accurate, for instance, to say that the differential benefit was caused by the subsequent actions of the other actors.\(^{13}\) By contrast, if saying – without qualification – that a given action caused a particular differential benefit is not apt to mislead in this sense, then we should regard that action as having ‘sufficiently directly’ caused that effect for the purposes of \( N^9 \).\(^{14}\)

2.2 The Subjective Nature of the Inquiry

The volitional neutrality account that we are developing, as mentioned, inquires into the subjective aims and beliefs of the actor. An actor therefore acts non-neutrally, pursuant to \( N^9 \), when she subjectively believes her act is likely to cause a differential benefit, even if a reasonable person would not have expected that act to have that effect in the circumstances. And conversely, an actor cannot be said to act non-neutrally simply because his action is one that a reasonable person would have regarded as likely to cause a differential benefit – although the actor does act non-neutrally if the action he performs is one that he too judges likely to have such a consequence.

13. In the case of the all-candidates’ meeting, for instance, merely calling the meeting, even if we assume it was a but-for cause of K’s win, did not sufficiently directly cause K’s win; rather, we should say that K’s performance at that meeting, or the voters’ reaction to that performance, caused K’s win.

14. This rule of thumb, I would suggest, does not deny the possibility that an actor might act non-neutrally in light of his performing an action that he expected would rather indirectly cause a differential impact. It is not the case, for instance, that saying ‘\( p \) caused a differential effect’ will always be misleading unless \( p \) caused the effect in the completely direct sense in which one billiard ball striking another directly causes the latter to move. Even if we insist on a dichotomy between direct and indirect causes, in other words, it is not the case that actions which are best regarded as indirectly causing a differential benefit are incapable of ‘sufficiently directly’ causing such effects for the purposes of determining their neutrality or non-neutrality.
Another important point to note about volitional neutrality is that, by inquiring into an actor’s subjective intentions, it forces us to distinguish between cases involving an actor who desires some state of affairs to come about, but does not act for the purpose of bringing it about (nor for the purpose of making it more likely to arise), and cases in which an actor’s desire that some state of affairs comes about does lead him to act for the purpose of bringing about that state of affairs (or for the purpose of making it more likely to arise). If I performed an action that I did not believe was likely to differentially help or hinder either Peter or Paul, and neither did I act for the purpose of making either result more likely, then I acted neutrally, even if at the time I performed the action I would have been pleased to learn that it would in fact differentially benefit Paul, say.\textsuperscript{15} Since I did not act because of my desire that Paul meet with success, my merely having that desire does not render my action non-neutral.

Admittedly, distinguishing between motivating and non-motivating desires in this manner is inherently difficult. I submit that the appropriate inquiry is whether a desire to cause a differential effect was \textit{any part} of the actor’s reasons for acting. Now, it might initially seem too demanding to say that an actor acts neutrally only if no part of her reason for action involves a desire to see a particular party to a contest differentially advantaged. However, this is, I think, quite in line with how we ordinarily understand what it takes to be truly neutral. Neutrality, in other words, is a rather exacting standard; in order to maintain her neutrality an actor must refrain from performing an action if she believes that it is likely to cause a differential impact on the relevant parties, or if any part of her reason for performing the action would be the possibility that it will have such an effect.\textsuperscript{16} The demanding nature of neutrality is what Robert Nozick appreciated, for example, when he wrote that the liberal state ‘scrupulously must be neutral’ between its citizens.\textsuperscript{17}

This idea that maintaining one’s neutrality is a relatively demanding business leads on naturally to an objection that might be raised against the volitional neutrality account. It might be alleged that volitional neutrality errs in focusing on whether the relevant actor subjectively believed, at the time she acted, that there was a substantial likelihood that her act would cause a differential effect. Mustn’t we ask – as well, or instead – whether it was objectively reasonable for the actor to conclude that her proposed course of action was unlikely to differentially impact the relevant parties? Surely we shouldn’t let an actor off the hook, so to speak, by endorsing their action as neutral in a case where it is apparent that a reasonable person, having taken all reasonable steps to inform themselves on the matter, would have expected the action to have a differential effect? Why adopt a conception of neutrality that makes negligence a full defence to a charge of non-neutrality?

The reason that this line of thinking is mistaken is due to the infelicity of invoking the concept of negligence in the context of neutrality. If non-neutrality was, by definition, morally blameworthy, then it is true that we would not want to adopt a conception of neutrality that allowed actors to escape moral sanction by invoking their own negligence. However, while negligence is an inherently moralized notion,

\textsuperscript{15} We will take up below the separate question of whether it is correct to say, in these circumstances, not only that my action was neutral, but also that I was a neutral actor.

\textsuperscript{16} Neutrality is not so demanding, however, (and as we will see in just a moment) as to require an actor to take all the measures that a reasonable person, committed to ensuring that her action did not cause a differential effect on the parties, would take to that end.

\textsuperscript{17} Nozick (n 8) 33 (emphasis added).
neutrality is not. To act negligently, for instance, is to fail to take all reasonable measures to prevent the occurrence of some undesired event. A negligent actor is, necessarily, a proper subject of moral disapprobation. But a neutral actor is not, necessarily, a proper subject of approbation. On the contrary, I take it as axiomatic that neutrality can, in certain contexts, be a morally impermissible posture to adopt. Such was the sentiment famously expressed, for instance, by John F. Kennedy, ostensibly borrowing from Dante, when he promised that ‘[t]he hottest places in hell are reserved for those who in time of moral crisis preserve their neutrality’. In short, the question of whether a given action is neutral or non-neutral is orthogonal to the question of whether the action is right or wrong, fair or unfair.

2.3 Actions Required by the Contest Itself

To recap, we left off, before delving into the ramifications of our expansive interpretation of ‘intention’, having tentatively offered the following definition of neutrality.

\[ N9: \text{Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest, where an intention to differentially help or hinder a party is understood to exist whenever the actor either acts for the purpose of differentially helping or hindering a party, or acts in the belief that there is a substantial likelihood that their action will, sufficiently directly, cause a party to be differentially helped or hindered.} \]

For ease of reference going forward, we can simply read ‘acting with the intention’ in this liberal manner, thus allowing us to shorten \( N9 \) to read:

\[ N9': \text{Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest.} \]

As it turns out, before we can arrive at a fully satisfactory account of neutrality, we will have to make two refinements to \( N9' \). The first refinement relates to actions required of an actor by the very nature of the relevant contest, and the second relates to non-volitional acts. Taking these in turn, consider a scenario involving the refereeing of a soccer match. Well into the second half, the game remains scoreless, with the two sides appearing evenly matched. The referee estimates that each team has an equal chance of winning. Imagine that at this point in the game a player on Team M clearly handles the ball within his own penalty area. Dutifully enforcing the rules of the game, the referee blows his whistle and awards Team N a penalty kick. The referee is well aware that by doing so he has just caused Team N’s chances of success to rise steeply (and Team M’s chances to correspondingly drop). For this reason, \( N9' \) would force us to conclude that the referee has acted non-neutrally. It seems clear, however, that we


19. That ‘neutral’ and ‘fair’ are far from synonymous is easily demonstrated. We can think of many examples – such as a case in which one’s older and stronger child gratuitously picks a fight with her younger sibling – where fairness demands a non-neutral stance, and where maintaining one’s neutrality would be unfair.

20. This definition of neutrality is identical to \( N3 \), although in moving from \( N3 \) to \( N9' \) we have of course given several vital clarifications of how \( N3’ \)s language of ‘acting with the intention of differentially helping or hindering’ should be interpreted.
do not regard referees who merely enforce the rules of the game as acting non-neutrally wherever they make a call that they apprehend will differentially affect the competitors’ chances of success. On the contrary, we are inclined to think that where the accepted rules of a particular game dictate that a referee must take an action that differentially alters the parties’ status quo chances of success – for example, by awarding a clear penalty to Team N in our soccer example – the referee would be acting non-neutrally if, for the purpose of ensuring that a particular party’s chances of success do not differentially increase, she refrained from taking that action. It is necessary, then, to revise $N9'$ to take account of these sorts of situations. To this end, we can offer $N10$.

$N10$: Acting neutrally means not acting with the intention of differentially helping or hindering a party to a contest, unless the actor believes that the rules of the contest itself require her to perform an action that has such an effect, in which case the actor acts neutrally in performing that action provided she does not do so for the purpose of causing a differential effect, and acts non-neutrally if she refrains from performing the action in order to avoid differentially helping or hindering a particular party.

2.4 Non-Volitional Acts

As mentioned, in order to put the finishing touches on our account of neutrality, we will need to acknowledge the fact that some actions are non-volitional. Imagine, for instance, that a man is asked to sit as a model for two artists. The artists are each given 30 minutes alone with the model to sketch his portrait. A panel of judges will afterwards select the winning portrait based on its likeness to the model. The model, having sat stock still for the first 30-minute session with artist O, receives news, just before sitting for artist P, that a family member has died. The man, overcome by his grief, sobs throughout the session with artist P. He does not wish to cry, but cannot stop himself from doing so. Imagine further that, as the man is aware, his inability to sit still for artist P hinders P’s ability to create a true to life portrait, and thus reduces P’s chances of success in the contest.

This scenario, according to $N10$, is an example of non-neutral action, since the model believed that his crying was likely to cause a differential impact on the chances of success of the two artists. But the judgement that the man acted non-neutrally seems wrong. The reason for this, I submit, is that it is part of our understanding of neutrality that an actor cannot have acted non-neutrally unless he exercised control over the relevant action. As such, we should refine $N10$ accordingly.

$N11$: Acting neutrally means not performing a volitional action with the intention of differentially helping or hindering a party to a contest, unless the actor believes that the rules of the contest itself require her to perform an action that has such an effect, in which case the actor acts neutrally in performing that action provided she does not do so for the purpose of causing a differential effect, and acts non-neutrally if she refrains from performing the action in order to avoid differentially helping or hindering a particular party.

21. Likewise, for example, we do not regard a judge as acting non-neutrally just because, in ruling that certain evidence of the plaintiff’s is inadmissible, she apprehends that her ruling will cause the plaintiff’s chances of winning the lawsuit to decrease.
In *N11*, we have finally found a definition of what it is to act neutrally that coheres with our core intuitions about neutrality – that neutrality requires refraining from purposefully or knowingly tipping the scales of a particular contest and instead letting the chips fall where they may, so to speak – and which is not susceptible to counter-examples.

2.5 The Neutrality Determination Exercise

It may be helpful to concisely lay out the procedure that should be employed, according to the neutrality of volition account of neutrality, in order to ascertain the neutrality or non-neutrality of a given action. (For simplicity’s sake, let us imagine that the actor does not believe that the rules of the contest itself require her to perform an action that would cause a differential effect.)

The Neutrality Determination Exercise: Determining the neutrality or non-neutrality of a given action requires determining whether the actor performed a volitional act with the intention of causing a party to the relevant contest to face a chance of success that was better or worse that the chance of success the actor believed that that party faced at the moment before the actor performed the action, while having no intention of causing all other parties to face a chance of success that was, respectively, equally better than, or equally worse than, the chance of success the actor believed that those parties faced at the moment before the actor performed the action.

According to volitional neutrality, as defined by *N11*, if the exercise above yields a negative finding, the actor’s action was neutral. If it yields an affirmative finding, then the action was non-neutral.

2.6 Rounding Out Our Understanding of Volitional Neutrality

One implication of the volitional neutrality account that is worth noting is that an actor may act neutrally even where, for instance, he chooses to adopt this stance due to a desire to appear neutral to others, combined with a belief that his favoured party is likely to succeed in the contest regardless of whether or not he intervenes on that party’s behalf. This fact highlights the disconnect between neutrality and impartiality. The actor in this scenario, for instance, is clearly not impartial as between the parties to the contest, and yet he has still acted neutrally. He possesses the sort of mental state that often accompanies and motivates non-neutral acts. However, he has not actually performed a non-neutral act, since he has not performed an action for the reason that it may, or in the belief that it will, differentially benefit any party to the contest. In short, a neutral action that is performed for strategic reasons does not therefore cease to be a neutral action.

It is also the case, on the volitional neutrality account, that the neutrality or non-neutrality of a given act may at times be very difficult for outside observers to ascertain. This is so in light of the inherent difficulties involved in identifying the particular effects that a given actor aimed to cause via her action, as well as the difficulties in determining an actor’s subjective expectations as to the likely consequences of her actions.\(^\text{22}\) In addition, we run into further potential complications when

\(^{22}\) Added to this is the difficulty of determining precisely how likely an actor must regard it as being that her action will cause a differential benefit before we may say that she ‘intended’ to cause that benefit. Volitional neutrality is somewhat vague on this score, claiming only that the
we seek to ascertain the neutrality or non-neutrality of an action taken by a collective entity. There is considerable controversy, for instance, about whether it even makes sense to speak of the intention of Parliament, say. While there is not space to tackle the issue here, I believe we can sensibly speak of the intention of a collective, and so it does make sense to ask whether a given legislature, say, acted neutrally or non-neutrally in passing some law. But it is an implication of neutrality of volition that if collective entities do not have aims or expectations, then ‘neutral’ and ‘non-neutral’ are not adjectives that properly apply to them.

A final clarificatory point pertains to the relationship between neutral actions and neutral actors. Volitional neutrality, as articulated and defended above, is an account of what it is to act neutrally. ‘Neutral’ and ‘non-neutral’ are therefore adjectives that can properly be applied to the actions of volitional agents. But what should we say about how we determine the neutrality or non-neutrality of actors? The simplest and most natural answer here would be that the neutrality of actors is determined by the neutrality of their actions. Thus an actor is neutral, relative to a given contest, just in case the actions he performs in the context of that contest are themselves neutral ones.

While this position has the benefit of both simplicity and clarity, we must, unfortunately, reject it in favour of a less straightforward account. It is commonplace, for instance, to hear a person described as ‘neutral’ where they are, say, watching a football game without any particular desire to see either team win. The corresponding notion that persons are non-neutral if they hope for the success of a particular party to a given contest is also very commonly expressed. This idea that we are non-neutral in respect of a contest where we are partial to a particular party to that contest – even if we take no action intended to differentially benefit that party – is in fact so ingrained in the way we speak about neutrality in everyday life that I believe we must account for it if we are to fully explicate the concept of neutrality.

In other words, while there is only one sense in which we should speak of an actor’s having acted neutrally or non-neutrally – whether an actor acts neutrally, as we’ve seen, is determined solely by her volitional state when performing the action – there appear to be two senses in which we can speak of the neutrality or non-neutrality of actors. Thus, where it can be said that, in respect of a given contest, an actor has not performed any volitional acts in the belief that his actions will, or for the reason that they might, cause a party to that contest to be differentially helped or hindered, we should say both that that actor has acted neutrally, and that the actor therefore is neutral. At the same time, however, such an actor may, in a different sense, be considered non-neutral if, despite his having acted neutrally with respect to the contest, he happens to be partial to a particular party to that contest. Likewise, where an actor performing a volitional act does so either in the belief that his action will, or for the reason that it might, cause a party to a contest to be differentially helped or hindered, we should say both that that actor has acted non-neutrally, and that the actor therefore is non-neutral. But we may nevertheless regard the actor as neutral in a question is whether the actor regarded that consequence as ‘substantially’ likely. I think this vagueness must be tolerated, as greater precision could only be bought at the cost of arbitrariness.

23. While there is not space to tackle this issue, I believe we can sensibly speak of the intention of a collective. For a well-argued defence of this notion, see Richard Ekins, The Nature of Legislative Intent (Oxford University Press 2012).

24. This is assuming that the actor does not believe that the rules of the contest itself require her to perform an action that would cause such an effect, in which case, as we noted, the actor still
different sense if, despite his having acted non-neutrally in respect of the contest, he is impartial between the parties to that contest.

3 VOLITIONAL NEUTRALITY POSITIONED IN THE LEGAL AND POLITICAL THEORY LITERATURE

Volitional neutrality, as expressed in \( NII \) above, is clearly not the only possible interpretation of what it is to act neutrally. Other prominent conceptions of neutrality that have been put forward in the political and legal theory literature include neutrality of effects, neutrality of intentions, neutrality of treatment, and neutrality of justification. It will be instructive to concisely outline the ways in which neutrality of volition is at odds with, and preferable to, each of these competing conceptions of neutrality.

3.1 Neutrality of Effects

According to the ‘neutrality of effects’ conception of neutrality, an action, in order to count as neutral, must not ‘have the effect – this effect being intentional or not – of favoring” a particular party.\(^{25}\) This view of neutrality, however, clearly conflicts with our ordinary understanding of what acting neutrally entails. As we saw early on with the boxing trainer example, for instance, it is quite implausible to hold that an actor acts non-neutrally wherever their action leads on to differential effects, since those effects may well have been neither aimed at nor foreseen (or even foreseeable). The problem, in a nutshell, is that the neutrality of effects account leads us astray by focusing exclusively on an action’s consequences, while taking no note of the volition of the actor at the time she performs the action. Volitional neutrality, by focusing not on the effects of an action, but rather on the actor’s aims and expectations – i.e., her volition at the time she performs the action – is therefore both clearly at odds with neutrality of effects, and superior to it.

3.2 Neutrality of Intentions

Another conception of neutrality – generally referred to in the literature as ‘neutrality of intentions’, or, sometimes, ‘neutrality of aim’ – focuses on the actor’s objective in performing a given act. For example, in the introduction to an edited collection on political neutrality, Roberto Merrill defines ‘neutrality of aim’ as holding that ‘the state should do nothing with the aim of promoting a conception of the good life over others’.\(^{26}\) Now, proponents of political neutrality did not create a new term to describe their view of what liberal justice requires of states vis-à-vis their citizens’ conceptions of the good. Instead, they co-opted a term that is readily understood in non-political contexts, as we have already seen. As such, and so as to have before us a conception of neutrality capable of application not merely in the political realm, we can regard the neutrality of intentions account as specifying that an actor, in order to act neutrally,

acts neutrally in performing that action provided she does not do so for the purpose of causing a differential effect.

\(^{25}\) Roberto Merrill, ‘Introduction’ in Roberto Merrill and Daniel Weinstock (eds), Political Neutrality: A Re-evaluation (Palgrave Macmillan 2014) 2.

\(^{26}\) ibid.
must refrain from taking any action aimed at differentially helping or hindering a party to a given contest.27

That volitional neutrality is a superior account of neutrality than is neutrality of intentions, however, was made clear above when we contemplated the scenario involving Pakistan dropping a bomb on Sinhalese armed forces. Since the Pakistani government did not actually aim to differentially help the Tamils, neutrality of intentions would force us to conclude that Pakistan acted neutrally. Instead, however, we concluded that on a sound conception of neutrality an action will be non-neutral not merely where it is performed with the aim of differentially affecting the parties to a contest, but also where the actor expects that such will be the likely result.28 Otherwise put, the mental state of the actor that suffices for a finding of non-neutrality is broader than that which ‘neutrality of intentions’ contemplates. And neutrality of volition, of course, insists on precisely this point.

3.3 Neutrality of Treatment

As part of an original and meticulously argued account of the ‘moral foundations of minority rights’, Alan Patten has made the case that, properly understood, a liberal conception of justice places the state under a pro tanto obligation to be neutral between all ‘worthwhile’ conceptions of the good29 (and between the ‘identity-related component[s]’30 thereof – including, as Patten sees it, ‘preferences’31 for activities that are encouraged or mandated by those conceptions32).33 As part of this larger project, Patten attempts to ‘reinterpret’34 the concept of neutrality, ultimately offering up an account that he calls ‘neutrality of treatment’.35

Patten explains that ‘[i]n contrast with the … traditional approaches’ of neutrality of intentions36 and neutrality of effects, neutrality of treatment requires that the state’s policies be ‘equally accommodating’ of cultures and conceptions of the good.37 It goes on to explain what it means for a policy to be ‘accommodating’ of a culture or conception of the good: ‘If the state adopts some policy that can be expected, in

27. Proponents of political neutrality import this notion into the political realm when they argue that it is a requirement of liberal justice that states adopt the stance described above in respect of ‘conceptions of the good’ (or, often, permissible conceptions of the good) – see, for example, Rawls (n 1) 460; Steven Wall, ‘Neutrality and Responsibility’ (2001) 98 Journal of Philosophy 389, 390 – these conceptions of the good standing in the position of competing ‘parties’ to a ‘contest’ in the sense that they compete for adherents among the citizenry.
28. Notice that this is not to endorse neutrality of effects, which, as we saw above, errs in asserting that an action’s merely having such a differential effect suffices to make it a non-neutral act.
29. Patten (n 11) 128–36.
30. ibid 158.
31. ibid 156–64.
32. For further discussion, see Chiara Cordelli, ‘Neutrality of What?’ (2017) 20 Critical Review of International Social and Political Philosophy 36.
34. Patten (n 33) 249–72; Patten (n 11) 112–48.
35. Patten groups together, in a manner that I would suggest invites confusion, ‘neutrality of aim’ and ‘neutrality of justification’ as two subspecies of ‘neutrality of intentions’: Patten (n 11) 112.
36. ibid 115–19.
conjunction with other necessary inputs, to make a particular conception of the good more successful, then, in my terminology, its policy is "accommodating" toward that conception."37 Patten clarifies that whereas "the basic idea of neutrality of effects is to equalize across outputs of the policy process, the idea of neutrality of treatment is to equalize across the state's inputs", while acknowledging that "for [his] view to get off the ground, there must be some metric of inputs that does not simply reduce back to outputs".38 It appears that ultimately state action is "equally accommodating" and thus neutral, on Patten's account, if, all else being equal, that action could be expected to equally impact on the success of the relevant conceptions of the good.

Consider, for instance, the following example that Patten provides:

Imagine that a field belongs to a local public authority, which, up to now, has prohibited its use for team sports. The local authority now decides to relax that policy and to allow any group of people to sign up to use the field for the team sport of their choosing. The local community contains people who would like to play softball on the field and people who would like to play cricket. … Imagine that the distribution of preferences for softball and cricket is heavily tilted toward softball. Large numbers of people want to play softball, and only a handful want cricket. With these preferences in the background, the policy can be predicted to have a very unequal impact on softball and cricket lovers. It is now much easier (we might suppose) to play softball: finally there is a suitable place to play. But for the few would-be cricket players, it is not much easier to play their preferred game, at least not a proper game. It is true that they now have access to one of the inputs (a field) that they need to realize their preference. But so few people want to play cricket that there is no realistic prospect of ever getting a game together.39

According to Patten, the local government’s policy here is equally accommodating of these different preferences and is therefore consistent with neutrality of treatment. This is true because the policy extends exactly the same rights to bearers of each sporting preference, and not because of any conjectures about the impact of the policy on the popularity or realizability of the two sports.40

All else being equal, in other words, opening up a field to all comers would not be expected to differentially impact on the success of softball or cricket (whether measured by the popularity of the sport or its realizability in the lives of citizens). And so such a policy is neutral, on the neutrality of treatment account, even if the local government expects that the scheme is likely to have that differential effect in the actual real-world circumstances in which it is implemented.

37. ibid 115. This sentence might seem to suggest that Patten’s neutrality of treatment differs from neutrality of effects only in focussing on an action’s expected, as opposed to actual, effects. But Patten makes clear that that is not how he understand neutrality of treatment. He writes, for instance, that "[i]n a wide range of different situations, it is possible to identify a sense in which the state might adopt policies toward rival conceptions of the good that are equally accommodating even though they can be expected to have different impacts on the success of those conceptions": ibid 115–16.
38. ibid 115.
39. ibid 117.
40. ibid. For a helpful discussion of how the success of conceptions of the good can be understood in terms of the popularity or the realizability of these conceptions, and of how the two measures differ, see Patten (n 33) 256–59 and Patten (n 11) 114–17.
In fleshing out his conception of neutrality, Patten provides some helpful insights and analysis. He does well, I think, to look beyond the realm of the political and to inquire into the neutrality of private actions.\footnote{41} By paying regard to how we conceive of neutrality in such everyday scenarios, for instance, he increases the likelihood that his account of state neutrality will cohere with how neutrality is understood in less specialized contexts. Relatedly, Patten is correct in arguing that neutrality should not be contrasted ‘exclusively with perfectionism’, and should instead be ‘oppos[ed] ... to a broader range of uses of political power’.\footnote{42} However, Patten’s neutrality of treatment conception ultimately must be rejected in favour of neutrality of volition.

As we saw, Patten argues that in order to determine the neutrality of an action, we must focus our inquiry solely on the nature of the treatment extended to the parties itself, divorced from all of the following considerations: the effects such treatment could be expected to cause once we factor in the particular circumstances of the affected parties; the effects such treatment actually causes; the intentions of the actor in extending the treatment; and the reasons that could be given to provide a justification for extending that treatment.\footnote{43} It should be clear by now, however, that a satisfactory account of neutrality cannot afford to regard as irrelevant the intentions or real-world expectations of actors.

If, for instance, we alter Patten’s example slightly and stipulate that the local government not only foresaw the unequal benefit that making the field available to all team sports would have when it came to the popularity of softball versus cricket, but also acted \textit{for the express purpose of bringing that particular differential benefit about}, then surely we would want to say that the local authority acted non-neutrally. Surely, that is, even if I treat two parties in a manner calculated to appear even-handed, by extending treatment that seems, from a high enough level of abstraction, incapable of differentially affecting their fortunes, but which is actually employed with the aim of doing precisely that, then I have acted not neutrally but non-neutrally.\footnote{44} Since

\begin{itemize}
\item \footnote{41}{See Patten (n 11) 115ff.}
\item \footnote{42}{ibid 106.}
\item \footnote{43}{I am sceptical that Patten is here giving us an account of neutrality that is capable of practical application on its own terms. For instance, Patten, as we saw, believes that the extent to which a policy ‘accommodates’ a conception of the good is determined by the extent to which that policy ‘can be expected, in conjunction with other necessary inputs, to make a particular conception of the good more successful’. But how could we respond if we were pressed to state with specificity what these ‘other necessary inputs’ consist of? I suspect that we could not hope to give an accounting here that was comprehensive and principled, as opposed to partial and ad hoc.}
\item \footnote{44}{Patten’s failure to offer up a conception of neutrality that coheres with our belief that an actor acts non-neutrally whenever she acts with the specific aim of differentially benefiting a particular party to a contest can be understood in light of his larger project of invoking neutrality as an attractive political ideal that can help to explain why minority cultural rights are a requirement of justice. For instance, Patten rejects neutrality of intentions as an account of what justice requires of states, since, as he puts it, neutrality of intentions ‘sweeps too broadly’, in the sense of ‘counting as neutral’ – and thus, as he sees it, endorsing as just – ‘policies that seem, intuitively, to be nonneutral’ – which, on his view, makes them presumptively unjust – such as state establishment of religion where such establishment is carried out purely to enhance the state’s perceived legitimacy rather than to promote the chosen religion (Patten (n 11) 113). Preoccupied with the issue of whether or not neutrality of intentions amounts to a normative standard that states are bound, as a matter of liberal justice, to live up to – and finding in the negative – Patten fails to recognize that neutrality of intentions is nevertheless correct in making the conceptual claim that action taken with the aim of differentially benefitting a particular party...}
\end{itemize}
volitional neutrality – which regards as non-neutral an action taken in the expectation that it will cause, or taken for the purpose of causing, a differential effect – can readily account for this intuition while Patten’s neutrality of treatment cannot, the former is superior to the latter as a conception of what it means to act neutrally.

Consider, in order to drive home the point, another example that Patten himself employs. A certain philanthropist, Patten asks us to imagine, is ‘faced with a decision about how to allocate money between two worthy projects’.45 According to Patten, one option here:

would be for the philanthropist to give each project the amount that is calculated to bring the projects as close as possible to equal levels of success. A second would simply be to give each project equal amounts of money. One decision rule equalizes across outputs; the other across inputs.46

Since, as we noted, neutrality of treatment requires the equalization of inputs, the philanthropist acts neutrally on the neutrality of treatment account only if he opts to allocate the money in the latter, as opposed to the former, manner.47 Let’s suppose the philanthropist did so, giving an identical monetary donation to each cause. The problem, however, is that to embrace neutrality of treatment would be to declare the philanthropist’s actions here to be neutral without knowing anything about the aims or expectations behind them.

It is quite possible, for instance, to imagine cases in which a philanthropist’s donating equal sums of money to two different charitable projects would, once we factor in her intentions in doing so, strike us as non-neutral. Imagine, for example, that one project involves buying up and restoring to their former glory some wetlands, while the other project aims to purchase the same property in order to fill it in and erect low-cost social housing. The two projects are in direct competition with one another, since the lands can only be purchased by one group, and the first to raise sufficient funds to cover the asking price will secure the property. Imagine that the philanthropist has it on very good authority that the wetlands project is presently just $150,000 short of being able to meet the asking price for the lands in question, while the social housing project is a million dollars shy of that mark. Imagine further that the philanthropist proceeds to make a precisely equal $150,000 donation to both causes, with the specific aim of ensuring that her favoured wetlands restoration project therefore secures the land at the expense of the rival cause. Here I think we would and should regard her actions as quite non-neutral, notwithstanding that Patten’s neutrality of treatment account seems committed to viewing them as equally accommodating and therefore neutral.48

is for that reason non-neutral. (We will have more to say below on how ‘neutrality’ is often employed as a moralised norm, in ways that are at odds with our conceptual understanding of the term.)

45. Patten (n 11) 115.
46. ibid.
47. Patten goes on to explain that ‘[t]he basic idea behind neutrality of treatment is to generalize from cases like that of the philanthropist’: ibid.
48. One way to defend the neutrality of treatment account in the face of this example might be to claim that the philanthropist’s equal donations do not in fact equally accommodate the two projects, once their respective pre-donation financial positions are taken into account. As mentioned, Patten is not entirely clear on the ‘other necessary inputs’ that appropriately factor into the question of how ‘accommodating’ a given ‘treatment’ is. So it might be claimed that the
To round out our discussion of how volitional neutrality is a better conception of neutrality than Patten’s neutrality of treatment account, it may be helpful to quickly note how volitional neutrality would regard the local government’s action in Patten’s softball hypothetical. Let’s assume that the local council did not act for the purpose of differentially benefitting softball (or softball enthusiasts) over cricket (or cricket enthusiasts), but did expect (given the respective sizes of the softball-loving and cricket-loving communities) that the practical effect of their decision to make the field available for team sports would be to increase the popularity and realizability of softball over cricket. Volitional neutrality tells us that the question of whether the council therefore acted non-neutrally is answered by determining whether the council expected their actions to sufficiently directly cause such a differential effect. Applying the rough and ready methodology outlined earlier, would it be misleading to say, without further elaboration, that the council’s decision caused the differential impact? If so – if the only non-misleading claim here would be that both the council’s action and the relative size of the two communities were major features of the causal story – then volitional neutrality tells us that the council’s action was neutral.49

To be clear, the fact that the council’s actions might strike us as justified would not be a good reason for thinking that their actions were also neutral. We might, that is, feel a temptation to reason from the premise that the council did nothing unfair in opening up a field to all comers to the conclusion that it cannot therefore have acted non-neutrally. But that simply does not follow. We might, for instance, empathize with an actor who is struggling doggedly to remain neutral while apprehending that even the most even-handed of interventions, in the context of the particular contest she is caught up in, will likely accrue to the differential benefit of one party over the others. We might wish to say that an actor who chose to take that minimally disruptive action acted neutrally, on the thinking that in the circumstances her action was ‘neutral enough’, and that it would be harsh to accuse her of having acted non-neutrally when she extended to the parties the most even-handed treatment she was capable of providing. But the better approach, of course, is just to say that in such cases there may be no possibility of occupying a neutral ground at all. We should admit, that is, that the actor’s action was non-neutral while insisting that (since, as we saw above, neutrality is not an inherently moralized concept) the bare fact that a given action was non-neutral says nothing at all, in and of itself, about whether the action was right or projects’ respective financial positions ought to be ruled in, with the result that the only way for the philanthropist to equally accommodate the two projects would be to donate more to the less cashed-up group. We might, that is, wish to alter neutrality of treatment so as to require an actor to take account of how far away from some benchmark of success each party stands, and then carefully calibrate their intervention so as to only extend treatment that will not disturb the parties’ relative positions vis-à-vis that benchmark. Such a recasting, however, would seem to open up all manner of practical difficulties in determining what may be rightly taken into consideration (and what rightly excluded) when working out how accommodating of a given party a proposed action would be. Moreover, it would arguably see neutrality of treatment collapse into a disguised version of neutrality of effects. In any event it would clearly run contrary to the whole spirit of Patten’s account of neutrality, premised as it is on ‘equalizing inputs’ rather than ‘outputs’. 49. Patten’s example has an obviously similar structure, for instance, to the ‘all candidates’ meeting’ scenario we examined earlier, and which we found to be an example of a neutral action.
wrong, and thus nothing at all about whether the actor should be criticized for performing it.\footnote{50}

Ultimately, then, if we accede to Patten’s attempt to re-orient our thinking when it comes to determining the neutrality of actions – if, as he counsels, we focus squarely and exclusively on the relevant ‘treatment’ in itself – we would be focussing on a red herring. Asking whether some treatment extended by the state in and of itself ‘accommodates’ a particular conception of the good more than others may well be the right question to ask if our concern is to determine whether that state action is \emph{just}.\footnote{51} But that is not the pertinent question if our concern instead is to determine whether the action is \emph{neutral}.

### 3.4 Neutrality of Justification

Steven Wall and George Klosko formulate the principle of justificatory neutrality as follows: ‘The state should not aim to do anything to promote any particular conception of the good, or give greater assistance to those who pursue it, unless a plausible neutral justification can be given for the state’s action’.\footnote{52} On this view, just because the effect of a given state action is more favourable to some conception of the good than others – and even if the state knew that that would be the action’s effect – this does not mean that the state has thereby non-neutrally favoured that conception. Instead, ‘the requirement that the state not favor any conception of the good simply rules out various arguments or considerations for justifying state action’.\footnote{53} Specifically, ‘the state should not justify what it does by invoking controversial claims about the worth or value of different conceptions of the good’.\footnote{54} As long as the state can provide a justification of its action that does not appeal to such proscribed grounds, that action will be neutral notwithstanding any differential impact it may have on the success of citizens’ various conceptions of the good.

Neutrality of justification proves on close inspection, however, to be an unsatisfactory account of neutrality. This is because, even within the political context, we will clearly want to label as non-neutral some actions that justificatory neutrality would have to endorse as neutral. To make the point by way of a somewhat whimsical example, suppose we discover an opera-loving alien race that credibly threatens to

\footnote{50. It is worth noting that one of Patten’s explicit aims in \emph{Equal Recognition} (n 11) is to work up a ‘reinterpretation and defence’ of the concept of neutrality, such that the rehabilitated account of neutrality represents a \textit{pro tanto} requirement of justice (see ch 4). Patten is in search, that is, of a moralized notion of neutrality such that the concept represents a norm of political ethics. This search for a ‘fit-for-purpose’ conception of neutrality, however, is at odds with the fact, which we noted earlier, that our ordinary understanding of neutrality contains within it an acknowledgement that to act neutrally can at times be to act unjustly.}

\footnote{51. It would appear that this view of what justice requires vis-à-vis conceptions of the good stands or falls with the ‘resourcist’ account of justice in distribution put forward, most famously, by Ronald Dworkin. See Ronald Dworkin, \textit{What is Equality? Part 2: Equality of Resources}’ (1981) 10 Philosophy and Public Affairs 283. Patten, for instance, has explicitly tied his claim that neutrality of treatment is a \textit{pro tanto} requirement of justice to such a resourcist account, writing that ‘neutrality of treatment is an aspect of a broader, resourcist conception of justice, and stands or falls with such a conception’: Patten (n 33) 42.}

\footnote{52. George Klosko and Steven Wall (eds), \textit{Perfectionism and Neutrality: Essays in Liberal Theory} (Rowman & Littlefield 2003) 8.}

\footnote{53. ibid 7.}

\footnote{54. ibid 8.}
annihilate Earth unless the Earth’s inhabitants turn en masse from Hollywood blockbusters to opera. If an individual state, out of a desire merely to save the planet from destruction, were to respond by taking action to ensure that their citizens regularly attend the opera and stop going to the cinema, we should say that although they have acted very sensibly, they have not acted neutrally – either as between opera and the cinema or as between citizens who at the time of the threat preferred opera and citizens who at the time of the threat preferred the cinema. Neutrality of justification, however, would force upon us the highly counter-intuitive conclusion that the state’s action here was neutral, since a justification of the action that does not appeal to controversial claims about the relative value of different conceptions of the good was available. In contrast, volitional neutrality would have no trouble identifying the state’s action as non-neutral, given that it was clearly taken with the intention of favouring opera.

Consider another example. The United States, we say, was not neutral (certainly not post-1941) in the Second World War. We know they were not neutral because they chose to act with the specific intention of assisting the Allies in their fight with the Axis Powers. Now, it might be argued that America’s involvement in that conflict could not have been justified other than by appeal to controversial conceptions of the good. But the important point to notice here is that we don’t need to determine whether that is so in order to decide whether the United States was neutral or non-neutral. We can see straight away that it wasn’t acting neutrally. As the neutrality of volition account allows us to acknowledge, the fact that the US was clearly acting with the aim of helping the other Allied forces defeat the Nazis is sufficient reason for us to regard it as non-neutral in that conflict. No further inquiries need be made. Thus when it comes to assessing the neutrality or non-neutrality of actions, reliance on the neutrality of justification account of neutrality will lead us to the wrong answer in specific cases, and, in all cases, will lead us to ask the wrong sorts of questions.

4 TAKING STOCK

We have mentioned briefly the ongoing debate in legal and political philosophy over state neutrality and perfectionism in order to see that the conceptions of neutrality that are employed there are ultimately defective, as they do not square at all well with our judgments about the neutrality or non-neutrality of various actions. To be clear, nothing in our explication of the concept of neutrality was meant to deny (nor to assert) that the sort of state action that proponents of a given defective account of neutrality are advocating in regards to conceptions of the good might be exactly what justice requires of states. For instance, we might firmly believe that states, as a matter of justice, must only take action that promotes particular conceptions of the good if they can justify those actions in non-perfectionist terms, as neutrality of justification requires. We might think that states are under a duty of justice to refrain from acting with the aim of promoting particular conceptions of the good full-stop, as neutrality of intentions holds. Or we might be convinced that states, as a matter of justice, are under at least a pro tanto obligation to equally accommodate all (worthwhile) conceptions of the good, as Alan Patten contends. But what we should not think, I have argued, is that...
any of these views about how the state ought to act vis-à-vis conceptions of the good amount to compelling conceptions of neutrality.

It might, however, be pointed out that scholars who argue for or against so-called ‘political neutrality’ (or, sometimes, ‘liberal neutrality’) seem to adopt whatever conception of neutrality they do on the basis, first and foremost, of their belief that that standard of conduct is normatively desirable. Since the particular standard of political morality they wish to espouse bears passing resemblance at least to the concept of neutrality as we tend to think of it in everyday life, they are happy to use the term ‘neutral’ as a shorthand for it. We could well imagine, then, that many of these writers might respond to the fact that their conception of neutrality turns out to be irreconcilable with some aspects of neutrality as we ordinarily understand that concept by claiming that this is more or less irrelevant for their purposes, and no challenge at all to their substantive claims about the orientation that governments should or must take towards citizens’ conceptions of the good.

As such, even if we agree that the accounts of neutrality these theorists advance – which are often explicitly moralized, and typically confined to contests between conceptions of the good that are both permissible and controversial – stretch the concept of neutrality beyond the ways in which we ordinarily understand it, shouldn’t we nevertheless let the scholars use the word as they see fit? Surely there is nothing scandalous in their employing the term ‘neutral’ as a shorthand for it. We could well imagine, then, that many of these writers might respond to the fact that their conception of neutrality turns out to be irreconcilable with some aspects of neutrality as we ordinarily understand that concept by claiming that this is more or less irrelevant for their purposes, and no challenge at all to their substantive claims about the orientation that governments should or must take towards citizens’ conceptions of the good.

As such, even if we agree that the accounts of neutrality these theorists advance – which are often explicitly moralized, and typically confined to contests between conceptions of the good that are both permissible and controversial – stretch the concept of neutrality beyond the ways in which we ordinarily understand it, shouldn’t we nevertheless let the scholars use the word as they see fit? Surely there is nothing scandalous in their employing the term ‘neutral’ as a shorthand for it. We could well imagine, then, that many of these writers might respond to the fact that their conception of neutrality turns out to be irreconcilable with some aspects of neutrality as we ordinarily understand that concept by claiming that this is more or less irrelevant for their purposes, and no challenge at all to their substantive claims about the orientation that governments should or must take towards citizens’ conceptions of the good.

The correct response here, surely, is to concede that no great scandal is occurring when, for example, legal and political philosophers employ the term ‘neutral’ to refer to state action that is endorsed by a Rawlsian notion of public justification,56 despite neutrality and public justifiability being orthogonal to each other.57 It is likewise no scandal that many critics of the legal and political system employ the term ‘non-neutral’ to describe any state action that impacts different populations differentially. (I certainly do not wish to be taken as saying that these invocations of neutrality are, like the ending of a sentence with a preposition, crimes against language up with which we simply must not put.) I do wish to insist, however, that using the label ‘neutrality’ to describe the various standards of conduct that justice allegedly requires in respect of the state’s orientation towards conceptions of the good (as opposed to speaking of, for instance, ‘public justification’, ‘anti-perfectionism’, or ‘equal accommodation’) is unhelpful; it obscures rather than illuminates,58 just as we obscure matters when we similarly conflate neutrality with fairness or impartiality.

The standard of conduct picked out by the term ‘neutrality’, for instance, is a familiar one to us because we often notice various sorts of actors striving to hew to this standard, or at least striving to make it appear that they are doing so. We see individuals and groups as well as governments stake a claim to neutrality in order to

---

56. Rawls’s defence of something like what has come to be known as the public justification principle is evident in his A Theory of Justice (Harvard University Press 1971). For a more direct exposition of ‘public justification’ see Rawls (n 1) 435–90.
57. To be sure, it is possible for a course of action to be neutral and also satisfy neutrality of justification’s requirement of public justifiability. But the point is that it is also quite possible for the actions of a state to be publicly justifiable but non-neutral (as in the opera example just discussed), or (although this will be a rare case) neutral but not publicly justifiable. The two standards of conduct are not mutually exclusive, but neither are they coextensive.
58. See János Kis, ‘State Neutrality’ in Michel Rosenfeld and András Sajó (eds), The Oxford Handbook of Comparative Constitutional Law (Oxford University Press 2012) 318.
rebuff accusations of favouritism, or in order to not upset one party to a conflict by appearing to take another’s side. We speak, therefore, of neutrality as a standard of conduct that morality may require of individuals, and that prudence, at least, will often recommend to us. When we also speak of neutrality in the highly specialized sense of public justifiability, or in the (competing and distinct) sense of equal impact, we court confusion in two directions. Firstly, co-opting neutrality to refer to what is in fact public justifiability, for instance, courts confusion about what ‘neutrality of justifica-
tion’ actually entails, since those engaging with these debates may wrongly assume that neutrality of justification has more in common with our ordinary understanding of what it is to act neutrally than is in fact the case. They may be rather uncertain, for instance, about just which aspects of neutrality as it is ordinarily understood are retained within the idea of neutrality as it is invoked by legal and political phil-
osophers. (This would appear to be precisely the reason that Rawls eschewed the term in his own writings.)

Secondly, and inversely, when talk of ‘political neutrality’ seeps from scholarly circles into the wider public discourse, it risks diminishing the explanatory power of neutrality as a concept, by causing all of us to lose sight of the precise contours and requirements of the distinct standard of conduct that neutrality properly understood constitutes – that standard being one we should very much want to be able to speak accurately and precisely about, given that it is, again, a standard of conduct that various state and non-state actors commonly believe themselves to have moral or prudential reasons for adopting.59

Suppose, however, that a legal or political philosopher advancing the cause of so-called liberal neutrality were to insist that even if volitional neutrality is the best available account of what it is to act neutrally, it remains crucial that we be able to call out state action that can only be justified by appeal to controversial conceptions of the good as non-neutral. This is crucial, we might imagine them to say, because such actions, in promoting certain controversial views of the good life over others, violate core liberal norms of equality, respect, autonomy, and so on. How should we respond to this rejoinder? The correct response, of course, is that even if we grant the claim that it is crucial to call out such perfectionist actions as violating the sorts of norms just listed, we should refrain from automatically using the term ‘non-neutral’ to do so. And that’s because such actions, as we’ve seen, are not necessarily non-neutral. (The actor or actors who performed such an action may have done so without foreseeing, and without aiming to bring about, the differential benefit that the action conferred on the relevant conception of the good or on those who pursue it.) Further, the claim that the state must refrain from acting on the basis of perfectionist premises can easily be made using the language of ‘anti-perfectionism’, or ‘public justification’, or ‘non-
sectarianism’, or the like. There is no need to adopt a moralized – and for that reason alone an implausible – conception of neutrality as public justifiability, and then use

59. We could, of course, simply choose to use the term ‘neutral’ to describe an actor that acts as, say, neutrality of justification requires, and then coin some new term for an actor who acts as volitional neutrality demands. However, given that neutrality is not a term of art, but rather a term that ordinary people have been using in everyday conversations for hundreds of years, employing it in this new way would foster a considerable amount of confusion – thus negating the promise of clarity that makes concepts like neutrality valuable as explanatory tools in the first place. By employing the concept of neutrality properly – that is, by conceiving of it in accordance with the volitional neutrality account we have laid out above – we forestall such confusion, and therefore conserve the explanatory power that the concept, correctly understood, possesses.
‘non-neutral’ as a kind of epithet that attaches to any actions that violate this principle of public justification. Refraining from doing so would instead acknowledge that the standard of conduct picked out by the term ‘neutrality’ (properly understood) is, even where we have in mind a state actor, conceptually distinct from the conduct that the notion of public justification requires.

There is a further, more politically salient reason to deploy the term ‘neutral’ advisedly. Consider, for example, a government that decides not to allow the teaching of creationism in publicly funded science classes. This decision is no doubt a sound one, and in my view can be justified without appeal to any controversial conception of the good. But none of this changes the fact that the government’s decision not to teach creationism in public schools is quite non-neutral in respect of creationism (provided, that is, that we regard creationism and evolution by natural selection as rival theories competing for adherents, and that the state apprehends that by excluding creationism from the classroom it hinders creationism from attracting proponents). Creationists, for instance, commonly lament that decisions such as these by education departments around the world hinder their ability to promote their views as an alternative to the theory of evolution. And no doubt they’re quite right on that score. If, instead of admitting as much, we insist that the state’s disdain for the pseudo-science that creationism peddles is perfectly ‘neutral’, since it can be given a public justification, we actually give succour to those convinced that the so-called neutrality of the contemporary liberal democratic state is being weaponized against them and many of their most cherished beliefs and practices. By giving the impression that we do not even notice that we are favouring sound science over pseudo-science, we give ammunition to those who argue that they are the victims of what is in fact a bigoted and intolerant ideology of militant secularism dressed up as the height of tolerance and neutrality.

Something rather similar can be said in respect of an imagined interlocutor who insisted on the efficacy of calling out as ‘non-neutral’ laws and policies that disparately impact certain populations. Indeed, the main reason, I suspect, that one might adopt the (quite out of favour) neutrality of effects conception of neutrality is not out of a conviction that it provides a good accounting of what acting neutrally involves, but rather because of an assumption that a commitment to the equality of persons requires such a view of neutrality. That is, we may be inclined to think that the effects – or at least the foreseeable effects – of an (especially government) actor’s actions often matter as much or more, from a moral point of view, as the mental state of the actor at the time the action was performed. We may, for example, be concerned to point out that facially neutral state action can differentially burden members of minority groups in unjust ways. That claim is no doubt true, and important. But – and this is the key point for our purposes – we need not and should not stretch the notion of neutrality beyond its proper bounds in order to make it. That is, endorsing the view that neutrality is a function of the subjective volition of the relevant actor – and not of the actual effects (nor the reasonably foreseeable effects) of her action – does not preclude us from acknowledging the differential impact that her neutral action may or may not have had. Nor does anything we have said above about volitional neutrality deny that the pursuit of social justice might require calling out instances of so-called

60. This because neither evolution by natural selection, nor the belief that the proper education of children entails teaching them the truth of evolution by natural selection, should be regarded as a conception of the good.
61. See Merrill (n 25) 3.
‘indirect discrimination’ in which neutral actions turn out to have unforeseen and yet substantially detrimental effects on certain parties.62

Indeed, it is a virtue of the neutrality of volition account we have been propounding that it allows us to recognize the possibility that a law or policy can be both neutral and discriminatory. This insight, for example, helps us to see that the objective of rooting out discriminatory laws and policies cannot be secured merely by requiring that the state act neutrally in creating and passing regulations. Even states that act in a scrupulously neutral manner will need to attend to the unintended consequences of their actions in order to fully combat discrimination.63

Finally, as with co-opting neutrality to refer to public justification, there is a prudential argument against co-opting neutrality to refer to equality of impact. The risk is that if we automatically decry as ‘non-neutral’ actions that cause disparate impacts, then others (and in particular ordinary citizens) may – drawing on their everyday understandings of what acting neutrally involves – wrongly assume that the indictment of a law, say, as non-neutral implies that that law was intended to have a discriminatory effect. This, in turn, will naturally lead some to believe that wherever a law is impeached as ‘non-neutral’ by scholars or activists, all that needs to be done to provide a full answer and defence of the law is to prove that those who passed it did not actually intend for it to have a discriminatory effect. Using the term ‘non-neutral’ to indict indirectly discriminatory laws, that is, actually makes it easier for supporters of those laws to rationalize them without having to face up to the way in which they are discriminatory.

In summary, then, political philosophers who wish to defend (or attack) the Rawlsian notion of public justification, as well as scholars and activists concerned to spotlight indirect discrimination, would be well advised to take care, in doing so, not to conflate these ideas with the conceptually distinct notion of neutrality. For one thing, there is no need to do so; we can easily say all we wish to say about these issues without invoking the idea of neutrality.64 What is more, as Rawls himself warned, ‘the term neutrality in these contexts is unfortunate; some of its connotations are highly misleading’.65

It might be thought that all of this still rather misses the point that proponents of neutrality in the political and legal theory literature are making.66 But by choosing not

---

63. Moreover, paying attention to the disparate impacts that neutral actions can give rise to is relevant to the neutrality of future actions. Consider, for example, an actor who has acted neutrally in performing a certain action for the first time. If it is clear that that action turned out, in defiance of expectations, to differentially burden a particular party to the relevant contest and would be likely to do so again, the subsequent performance of that act would therefore be non-neutral (assuming that its likely differential impact is now subjectively foreseen by the actor).
64. If, for example, one wishes to criticise perfectionist or discriminatory laws or policies as unjust, one can and should call them out for what they are – i.e., ‘perfectionist’, or ‘discriminatory’ – as opposed to automatically labelling them ‘non-neutral’ – which, on the volitional neutrality account of neutrality, they may or may not be.
65. Rawls (n 1) 191 (emphasis in original).
66. Their point being, I take it, that what they call ‘neutrality’ is a requirement of justice on a liberal conception, in light of certain core liberal norms such as equality, autonomy, and respect
to assess their claim that a liberal conception of justice requires the state to extend ‘equal accommodation’ to all cultures and conceptions of the good, as Patten asserts, or to refrain from taking any actions whose justification would entail appeal to perfectionist premises, as neutrality of justification contends, and instead inquiring into whether all actions that live up to these standards (whatever their virtues might be) are properly regarded as instances of neutrality, we do not miss the liberal neutralist’s point. Rather, we acknowledge the point they are making but deliberately adopt, for present purposes, a stance of agnosticism towards it – the better to draw attention to the fact that these writers make their point using language that is confounded and apt to confuse.

So where does all of this leave us? By way of analogy, it is indeed somewhat regrettable that the label ‘peanut’ has been affixed to the popular legume. Misnomers like this often lead to entrenched misunderstandings. Where we have better terms at our disposal – that is, terms that are more accurate and less apt to mislead – we should use them. But of course our failing to do so will not, alone, put an end to the possibility of making intellectual progress in the relevant field. In addition, there is no point denying that those writing about the so-called neutrality versus perfectionism debate will continue to readily use the term neutral to identify state action amenable to a non-perfectionist justification, just as would-be law reformers will continue to reserve the term neutral to identify state action that equally impacts all affected parties. The horse is now probably far enough down the track as to be out of sight of the barn.

5 CONCLUSION

We have, after careful consideration of numerous examples and counter-examples, been able to construct a comprehensive theory of neutrality – one we have labelled neutrality of volition, and according to which acting neutrally requires that an actor refrains from conduct engaged in for the purpose of helping or hindering a particular party in a given contest, as well as conduct that the actor believes is likely to cause such an effect – that is both internally consistent and able to make sense of our judgments about the neutrality or non-neutrality of various actions. This was the main goal of the article. Along the way, we also found that many theorists writing about neutrality as a political ideal are operating without the benefit of a coherent and plausible conception of what it is to act neutrally, and that in the interests of conceptual clarity it would be preferable if they spoke not of ‘neutrality’, but rather of ‘public justifiability’, or ‘equal accommodation’, or ‘equal impact’ (depending on the particular emphasis of particular scholars).

Further, we have cleared up, in the course of developing the volitional account of neutrality, a couple of potential misconceptions surrounding the concept of neutrality. It is not the case, for instance, that nothing turns on how we use the term neutrality in the context of political or legal debates; it is not merely a matter of semantics. Likewise, it is not the case that we should feel free to use neutrality however it best suits us on the grounds that neutrality is in any case an inherently vague concept, or

for persons, and due as well to certain empirical facts such as moral and cultural pluralism and the existence of what Rawls calls the ‘burdens of judgment’: Rawls (n 1) 487–88.

© 2021 The Author

Journal compilation © 2021 Edward Elgar Publishing Ltd
what Walter Bryce Gallie termed an ‘essentially contested concept’.\textsuperscript{67} We might have suspected, for instance, that it is simply impossible to explicate the concept of neutrality in such a way as to have before us a concept that coheres with our ordinary understandings of what it is to act neutrally and is at the same time sufficiently precise as to be capable of doing the sort of rigorous analytical work that proponents of ‘liberal neutrality’ in the political and legal philosophy literature, for example, employ their highly modified accounts of neutrality to do. But that is not so. We do have access to a precise, internally consistent conception of neutrality that keeps faith with our ordinary views about the neutrality of various actions – the conception we have called volitional neutrality.

\textsuperscript{67} Whereas Gray (n 3) 344, for instance, claimed that disputes over the proper usage of ‘essentially contested concepts’ ‘cannot be settled by appeal to empirical evidence, linguistic usage, or the canons of logic alone’, we have argued that ‘neutrality’ turns out to be a concept whose proper use can be determined by careful reference to linguistic usage and logical analysis, and so is not ‘essentially contested’ in the way that ‘art’ or ‘fairness’ are often alleged to be.