



How to make progress in moral theory? The limits of Gradualist Dutilitarianism

Christian Seidel¹

Received: 12 December 2025 / Accepted: 29 March 2026

© The Author(s) 2026

Abstract

According to the Gradualist Dutilitarian Project, the moral theory that we have most reason to adopt in order to make progress in moral philosophy emanates from a conjunction of two claims: (1) Dutilitarianism, the claim that what makes an action right or wrong depends on both consequences and deontological concerns; and (2) Gradualism, the view that moral rightness and wrongness come in degrees. Peterson’s argument for the Gradualist Dutilitarian Project is based on the Conditional Rescue Claim, which states that *if* you believe in Dutilitarianism, *then* you ought to believe in Gradualism — because Gradualism is necessary to make Dutilitarianism internally coherent when faced with a challenging result developed within an Arrowian framework. I will argue that the argument for the Conditional Rescue Claim fails. By carefully distinguishing between two ways of individuating theories, as well as between theories and classes of theories, the apparent incompatibility between the Arrowian result and the guiding idea of Dutilitarianism (as being a compromise) disappears.

Keywords Moral theory · Gradualism · Dutilitarianism · Sidgwick · Parfit · Arrow

One way to understand the point of moral philosophy is to conceive of it as a search for the most plausible moral theory. Among the multitude of possible theories, Martin Peterson’s eponymous article considers

Dutilitarianism “[W]hat makes right acts right and wrong ones wrong depends on consequences *and* duties” (Peterson, 2025, 2; original emphasis).

Dutilitarianism is designed to be a “compromise” between utilitarian and deontological theories (Peterson, 2025, 2). In his paper, Peterson argues for what I will call

the Conditional Rescue Claim *If* you believe in Dutilitarianism, *then* you ought to believe in

Gradualism “[M]oral rightness and wrongness come in degrees” (Peterson, 2025, 2)

✉ Christian Seidel
christian.seidel@kit.edu

¹ CompPhil²MMAE, Department of Philosophy, Karlsruhe Institute of Technology (KIT), Douglasstr. 24, 76133 Karlsruhe, Germany

I call this a rescue claim because Peterson develops it in an attempt “to render the utilitarian theory *internally* coherent” (Peterson, 2025, 2; original emphasis). The challenge to internal coherence, as he conceives it, is this: Within an Arrowian framework, Peterson establishes that the moral content of any utilitarian theory is unilaterally determined by the moral content of *one* of its basic theories, and that the existence of a single “dictatorial” basic theory undermines the very idea of Utilitarianism as being a compromise between two theories. But “in response to this formal result” (Peterson, 2025, 3), “utilitarians are best advised to give up [a certain Arrowian condition, viz.] the Completeness part of Ordering” (Peterson, 2025, 11). The ordering condition requires the utilitarian theory to yield transitive and complete rankings (Peterson, 2025, 8).¹ For Peterson, this is supposed to guarantee binary moral determinacy: any alternative is either morally right or morally wrong and thus has a determinate deontic status. Giving up this condition opens the possibility that “an act’s deontic status falls in a gray area” (Peterson, 2025, 6), where it is “either indeterminate whether it is right or wrong, or a bit of both” (Peterson, 2025, 6). Since, according to Peterson, the latter — gradualist — option is the more tenable view, giving up (the Completeness part of) Ordering in response to the Arrowian challenge amounts to embracing Gradualism. In that sense, Gradualism comes to the rescue of Utilitarianism.

It is important to highlight the conditional nature of the Conditional Rescue Claim. By itself, it dialectically supports neither Utilitarianism nor Gradualism. Indeed, because “one philosopher’s *modus ponens* is another philosopher’s *modus tollens*” (Putnam, 1994, 280), the Conditional Rescue Claim (if successfully established) can also be read as a *reductio* of Utilitarianism: if — as others have argued (e.g. Carlson, 2025) — we have independent reason to doubt the truth or even coherence of Gradualism, we have reason to reject Utilitarianism altogether.

So if Peterson’s line of argument is intended to work *in support of* Gradualism rather than against it, he must assign some initial dialectical weight to Utilitarianism. Though he does not state this explicitly, I take him to believe in

the Initial Commitment When we try to make progress in moral theory, Utilitarianism is among the most plausible starting points.

Given the Initial Commitment, the Conditional Rescue Claim implies

the Gradualist Utilitarian Project The most plausible moral theory — the theory that we have most reason to adopt in order to make progress in moral philosophy — emanates from a conjunction of Utilitarianism and Gradualism.

I will argue that Peterson’s case for the Gradualist Utilitarian Project ultimately fails because the argument for its centrepiece, the Conditional Rescue Claim, is far from conclusive and even inconsistent.

¹ A ranking is complete if and only if, for any two alternatives, one is at least ranked as high as the other.

1 Contrasting the Gradualist Dutilitarian Project with competing approaches: two distinctions

Let me begin by locating Peterson's approach in the landscape of other projects that aim for progress in moral theory, in particular Sidgwick's project in *The Methods of Ethics* (Sidgwick, 1981) and Parfit's projects in *Reasons and Persons* (Parfit, 1987) and *On What Matters* (Parfit, 2011).

When Peterson motivates Dutilitarianism, he contrasts it with toy versions of consequentialist and Kantian theories that have highly implausible deontic implications (Peterson, 2025, 2). One might therefore think that the primary aim of the Gradualist Dutilitarian Project is *deontic adequacy*: to develop a moral theory whose deontic content² best fits our well-considered deontic judgments and thus (ideally) covers all and only those deontic judgments we have most reason to accept as true. This project is similar to Parfit's search for a "Unified Theory" in Part I of *Reasons and Persons* (Parfit, 1987, 112). Dutilitarianism would then be a candidate Unified Theory. If that were all that mattered for progress in moral theory, it would be enough to consider only the deontic content and to

individuate theories extensionally Two moral theories T_1 and T_2 are different if and only if their deontic contents differ, i.e., if and only if there exists some action φ in some decision situation S such that φ 's deontic status in S according to T_1 differs from its deontic status in S according to T_2 .

However, as Peterson emphasizes in contrasting Dutilitarianism with theories of moral uncertainty, this is *not* the only thing that matters. What is distinctive about Dutilitarianism, he repeatedly highlights, is that "dutilitarianism is a claim about what *makes* acts right or wrong" (Peterson, 2025, 2–3; emphasis added; cp. initial definition of Dutilitarianism quoted above). This fits well with a received view in moral philosophy according to which moral theories identify the right- and wrong-making features of actions such that we are to

individuate theories substantively Two moral theories T_1 and T_2 are different if and only if they specify different sets of right- and wrong-making features (i.e., properties in virtue of which an action has its deontic status).

The systematic rationale for substantive individuation is that we not only expect a moral theory to deliver correct deontic verdicts, but also to justify them correctly: a moral theory ought to tell us not only *what* is right or wrong, but also (perhaps more importantly) *why* — i.e., in virtue of what — it is right or wrong. For this reason, moral theories are often specified by statements of the form " φ is right if and only if *and because* F" rather than " φ is right if and only if F."

Besides the systematic rationale for individuating theories substantively, there is also an "internal" rationale that is specific to Peterson's case for the Gradualist

² The deontic content of a theory T is the set of deontic verdicts (judgments on the deontic status of actions) that T entails in all possible decision situations.

Dutilitarian Project. Unless we take Dutilitarianism to be a claim about right- and wrong-making features — more precisely, the claim of

Substantive Dutilitarianism There are two irreducible kinds of features that make actions morally right or morally wrong: (1) features relating to the consequences of actions, and (2) features figuring in deontological concerns.

—, several claims about Dutilitarianism make less sense. First, Peterson introduces Dutilitarianism as a *contender* for (or an *alternative* to) utilitarian and deontological theories. But those theories typically are claims about (singular) kinds of right- and wrong-making features — particularly in response to worries raised by results on deontic co-extensionality (known as “consequentializing”). Thus, Dutilitarianism would not even be on a par (nor truly compete) with utilitarian or deontological theories unless it is understood as Substantive Dutilitarianism.

Second, conditional on the belief that any moral theory can be consequentialized by a one-dimensional (consequentialist) theory, the distinction between one- and multidimensional theories (which Peterson relies on when he later argues for his preferred version of multidimensional Dutilitarianism) loses its point unless we individuate theories substantively. Take an arbitrary dutilitarian theory T_{duti} and suppose there is a deontically co-extensional one-dimensional theory T_{one} . Individuating theories extensionally commits us to saying that T_{duti} is *identical* to T_{one} . But if T_{one} is one-dimensional, this implies that *any* (and in particular any multidimensional) T_{duti} is one-dimensional as well; and *mutatis mutandis*, in the special case where T_{duti} is multi-dimensional, T_{duti} being identical to — by being co-extensional with — T_{one} implies that the one-dimensional theory T_{one} is multi-dimensional as well. I take this to be a *reductio* of the distinction between one- and multidimensional theories — unless we individuate theories substantively and understand Dutilitarianism as Substantive Dutilitarianism.

Third, and most importantly, unless we do so, it is difficult to see how Dutilitarianism can be a “meaningful” (Peterson, 2025, 2) or “genuine *compromise*” (Peterson, 2025, 10; cp. p. 13) between utilitarian and deontological theories. If a dutilitarian theory T_{duti} turns out to be deontically co-extensional with another (say, utilitarian) theory T_{uti} , then individuating theories extensionally forces us to say that T_{duti} simply *is* T_{uti} . But if a dutilitarian theory is *identical* to a utilitarian theory, what does it mean to think of the dutilitarian theory as a “compromise” between this utilitarian theory and a theory of another type? Indeed, Peterson’s own argument relies on this point (Peterson, 2025, 2, 10, 13). He cannot make that point unless he individuates theories substantively and understands Dutilitarianism as Substantive Dutilitarianism.

So the Gradualist Dutilitarian Project is committed to Substantive Dutilitarianism. Indeed, it is this commitment that renders Peterson’s project a *distinct contender* among the competing approaches to making progress in moral theory: Like Sidgwick’s attempt “to frame a complete synthesis of practical maxims” in *The Methods of Ethics* (Sidgwick, 1981, 12) and Parfit’s quests for the Unified Theory in *Reasons and Persons* (Parfit, 1987, 112) and for the Triple Theory in *On What Matters* (Parfit, 2011, 411–419), the Gradualist Dutilitarian Project, with its search for a compromise between theories, is guided by the motivation to reasonably reduce the plurality of

competing moral theories. In this sense, all these projects share a commitment to increased *unification* in moral theorizing. They also share the view that moral theories identify fundamental right- and wrong-making features: Recall that Substantive Dutilitarianism is a claim about what *makes* actions right or wrong. This aligns with the Sidgwickian approach to moral theorizing, according to which part of the point of a moral theory — or rather of its constitutive principle — is to identify “ultimate reasons” (Sidgwick, 1981, 77–78, 83–85, 411; cp. pp. xxviii, 3, 8–9). It also aligns with Parfit’s approach to moral theorizing in *On What Matters*, according to which theories identify *higher-level* right- and wrong-making properties under which all more particular right- and wrong-making properties (e.g., “causing pointless suffering”) “can be subsumed, or gathered” (Parfit, 2011, 414; cp. p. 369), and which all lower-level right- and wrong-making properties “have in common” (Parfit, 2011, 415). Hence, it is only in virtue of Substantive Dutilitarianism that the Gradualist Dutilitarian Project is on a par with these other unificatory approaches in moral theorizing.

But there is also an important difference that likewise originates in Substantive Dutilitarianism. On the Sidgwickian account, each moral theory (or rather, its constitutive principle) identifies ultimate reasons of a *single*, unitary kind.³ On the Parfitian account, each moral theory — including Parfit’s Triple Theory — describes a *single* (complex) higher-level wrong-making property (Parfit, 2011, 25, 369, 414).⁴ According to Substantive Dutilitarianism, however, there are *two irreducible kinds* of right- and wrong-making features. So we might say that the very point of the Gradualist Dutilitarian Project — its *differentia specifica* — is to provide an alternative to Sidgwickian-Parfitian attempts at *monistic unification* in moral theorizing. And it is only in virtue of understanding Dutilitarianism as Substantive Dutilitarianism that this alternative can get off the ground, because Substantive Dutilitarianism asserts that there are *multiple* irreducible kinds of right- and wrong-making features. So Substantive Dutilitarianism (which relies on individuating moral theories substantively) distinguishes the Gradualist Dutilitarian Project from competing approaches to making progress in moral theory.

To the distinction between two ways of individuating moral theories, let me add a second distinction. I have been careful to speak of utilitarian *theories* and deontological *theories* in the plural, because I want to highlight that the terms “consequentialism,” “utilitarianism,” and “deontology” (or, using Peterson’s preferred term, “duty ethics”)

³ Those reasons are provided, e.g., by the fact that an action promotes an ultimate end (individual happiness, universal happiness, or the agent’s own excellence; Sidgwick, 1981, 7–9, 77–78) or by the fact that an action “conformed to certain precepts or principles of Duty, intuitively known to be unconditionally binding” (Sidgwick, 1981, 3).

⁴ A possible objection is that in Parfit’s Triple Theory there are *three* kinds of higher-order wrong-making properties — (1) the fact that an act is disallowed by a principle which is “one of the principles being universal laws would make things go best”; (2) the fact that an act is disallowed by a principle which is “one of the only principles whose being universal laws everyone could rationally will”; (3) the fact that an act is disallowed by a principle which is “a principle that no one could reasonably reject” (Parfit, 2011, 412–413) — and that taking the conjunction of these to be a *single* complex property (as Parfit does) is obscuring and misleading. But note that Parfit’s case for the Triple Theory rests on the claim that these properties are co-instantiated in the very same actions. In this sense, each of the properties is extensionally reducible to the other. It would then not be true that in Parfit’s Triple Theory there is more than a *single irreducible* kind of right- or wrong-making features. So even on this understanding, Parfit’s project is distinct from the Gradualist Dutilitarian Project.

refer to a *class* (or family) of theories (Seidel, 2019). There are many distinct ways to spell out the core commitments of consequentialism, utilitarianism, and deontology, and each of these specifies a possible consequentialist, utilitarian, or deontological theory: act-, rule-, or motive-utilitarian theories; consequentialist theories with hedonistic or other axiologies; Kantian and Rossian deontological theories, etc. In moral philosophy, we often try to make progress by identifying problems with entire *classes* of theories. We then make statements about utilitarianism (“Utilitarianism does not respect the separateness of persons.”) or deontology (“The paradox of deontology is a decisive objection against deontology.”), because we claim that these problems affect *all* members of the class in question.

Peterson comes close to acknowledging this point; he concedes that “both theories [i.e., utilitarianism and duty ethics] can be precisified in numerous ways,” but he thinks “that will not matter here” (Peterson, 2025, 4). There are two mistakes here. The first is that the sentence should read “both *classes* of theories can be precisified in numerous ways.” And the second is that this *does* matter for his argument. To prepare us for seeing this (later in section 2), let us carefully and consistently distinguish between claims about theories and claims about classes of theories. If Dutilitarianism “is a compromise between utilitarianism and duty ethics” (Peterson, 2025, 2), if both utilitarianism and deontology (or duty ethics) are classes of moral theories and if a compromise between X and Y is categorically of the same kind as X and Y, then Dutilitarianism must itself be a *class* of moral theories — the class of dutilitarian theories.

So in line with understanding Dutilitarianism as Substantive Dutilitarianism and taking the idea of a genuine compromise seriously, let us say that

T is a dutilitarian theory if and only if there exist a theory T_d from the class of deontological theories \mathcal{T}_D and a theory T_{uti} from the class of utilitarian theories \mathcal{T}_{UTI} such that T specifies exactly two irreducible features F_1 and F_2 that make actions morally right or morally wrong, where F_1 is the feature that makes actions morally right or morally wrong according to T_d and F_2 is the feature that makes actions morally right or morally wrong according to T_{uti} .

and that

Dutilitarianism (as a class) is the set of all dutilitarian theories (\mathcal{T}_{DUTI}) in this sense.

Note that both the Conditional Rescue Claim and the claim about the Gradualist Dutilitarian Project are most plausibly understood as claims about the *class* of dutilitarian theories. Consequently, in these claims (and thus in the goal of Peterson’s argumentation), “Dutilitarianism” must refer to Dutilitarianism (as a class).

2 Modelling moral theory choice in the Arrowian framework: two problems

We are now in a position to identify two problems with Peterson’s Arrowian argument for the Gradualist Dutilitarian Project.

This argument proceeds in four crucial steps: First, it transfers the framework of Arrow (2012) from aggregating voters' rankings on options (as representations of voters' preferences) to aggregating moral theories' rankings of acts (as representations of the theories' deontic content).⁵ Second, the argument applies the logic of Arrow's impossibility result to show that, for at least three options, there is no procedure for combining ("aggregating") a finite number of at least two basic theories into a hybrid (e.g., utilitarian) theory that simultaneously satisfies Ordering, Independence of Irrelevant Alternatives, Dominance, and Nondictatorship. One implication of this Utilitarian Impossibility Theorem is that any theory satisfying the first three of these conditions will be dictatorial: if T_{duti} is the hybrid utilitarian combination of a set of theories \mathcal{T} (where \mathcal{T} contains at least one deontological theory $T_d \in \mathcal{T}_D$ and at least one utilitarian theory $T_{\text{uti}} \in \mathcal{T}_{UTI}$, but no other kind of theory), then there will be a single basic $T \in \mathcal{T}$, such that T is decisive — i.e., for any pairwise choice, T 's (strict) verdict on that choice is sufficient for the hybrid utilitarian theory T_{duti} 's (strict) verdict on that choice. Thus, "any utilitarian theory that meets a small number of seemingly plausible conditions will turn one of the individual theories into a 'dictator theory' that unilaterally determines the ranking of the utilitarian theory" (Peterson, 2025, 3.)

But this, the argument continues in a third step, renders Utilitarianism internally incoherent because "forcing a hybrid theory to mimic one of its basic theories would square poorly with the intuition that a plausible hybrid theory should be a genuine *compromise* between basic theories" (Peterson, 2025, 10; original emphasis). Fourth and finally, the argument states that the best response to the tension between (a) the Arrowian conditions on utilitarian combinations and (b) the spirit of Utilitarianism as a compromise is to give up the Ordering condition⁶ — and thereby embrace Gradualism.

This reconstruction highlights what has been explained at the beginning: that Peterson's entire case for the Gradualist Utilitarian Project is devised as a response to a challenge to the internal coherence of Utilitarianism — the challenge being that, as spelt out in the third step, a violation of Nondictatorship is incompatible with the

⁵ This description invites the question whether this framework is appropriate if we are concerned with the *deontic adequacy* of moral theories. In social choice, the point of aggregating voters' preferences is to reach a collective decision that is reasonable and democratic by satisfying certain *procedural* constraints; but there are no additional constraints on the *content* of the aggregated collective preference ordering that comes about by satisfying these procedural constraints. By contrast, when we seek progress in moral theory we *do* care about the content of the aggregated theories and compare it to an external standard: we want to find the theory whose deontic content best fits our well-considered deontic judgments and thus (ideally) covers all and only those deontic judgments we have most reason to accept as true. As far as I can see, Arrowian aggregation of theories does not even guarantee that the aggregated theory is at least as deontically adequate as its most deontically adequate basic theory. I cannot elaborate on this point, but it may additionally limit the appeal of Peterson's case for the Gradualist Utilitarian Project.

⁶ At this point, one might question whether Independence of Irrelevant Alternatives is a reasonable constraint on aggregating moral theories. The condition is already disputed as a constraint on individual and collective rationality. However, when rankings do not reflect individual rational preferences but the deontic content of theories that have been carefully designed to increase deontic adequacy (see note 5), I see no reason to require either basic theories or the combined theory to satisfy Independence. *Pace* Peterson (2025, 9), many basic theories (especially in population ethics) *do* violate Independence because they aim to capture the holistic assessment of options that seems to be part of our practice of well-considered deontic judgments. Hence, another way to block the argument at this fourth step is to argue that giving up Independence is the better response to the tension.

guiding idea of Dutilitarianism as a genuine or meaningful compromise between utilitarianism and deontology. But in fact, it *is* compatible for two reasons that leverage the distinctions from section 1.

First, reconsider what Arrow's original Nondictatorship condition (Arrow, 2012, 30) requires of the social welfare function which aggregates individual preferences: that there is no individual i_d ("the dictator") such that for all pairwise choices, if i_d has a strict preference on that choice, then the collective has the very same preference on that choice — regardless of the other individuals' preferences on that choice. Formally, the conditional simply states a coincidence (cp. Hylland, 1986, 51n10): the (strict) collective decision result is always what i_d (strictly) wants. This can be read in different ways. The reading intended by Arrow — and suggested by the term "dictator" — is that i_d *determines* the social preference.⁷ On this reading, the collective decision result is always what the dictator wants *because* (i.e., in virtue of the fact that) she wants it. The dictator's preference is what *makes* it the case that the collective has the preference it has. But there are other, less ambitious readings of the conditional: The individual i_d is the lucky person who merely happens to be always right (in terms of what the collective prefers). Psephologists would be very happy to identify this individual, because in election forecasts, i_d would make the perfect predictor: the collective decision result is always what the predictor wants — but not *because* (i.e., not in virtue of the fact that) she wants it.

The Arrowian framework lacks the conceptual tools to distinguish between these two readings. This may be less problematic in social choice theory, but it becomes a problem in moral theory choice: it highlights that the Arrowian framework is a purely extensional framework and not sensitive to the *reasons why* (i.e., to the features in virtue of which) the aggregated collective result is the way it is. When applied to moral theories, it only concerns those parts of the (basic or combined) theories that tell us *what* is right or wrong, but not with those parts that tell us *why* — i.e., in virtue of what — it is right or wrong. Within this framework, we cannot distinguish between " φ is right if and only if F" and " φ is right if and only if *and because* F," and there is thus no room for individuating moral theories substantively.

This lacuna undermines the third step of Peterson's argument. Suppose that for a deontological basic theory T_d and a utilitarian basic theory T_{uti} , T_{duti} is a dutilitarian combination of T_d and T_{uti} that satisfies the first three Arrowian conditions. Then the theorem tells us that the deontic content of T_{duti} is co-extensional with the deontic content of one of its basic theories, say T_{uti} . This basic theory and the dutilitarian theory will always happen to assign the same deontic status to any given action. Nevertheless, it may still be true that what *makes* an action have this deontic status *sometimes* depends on those right- and wrong-making features specified by T_{uti} and *sometimes* on those right- and wrong-making features specified by T_d : Consider an arbitrary action φ that is morally right according to the dutilitarian theory T_{duti} . Then given deontic co-extensionality, φ must also be morally right according to the utilitarian basic theory T_{uti} . There are two cases. Either φ is also morally right according to the basic deontological theory T_d , or it is not. In the latter case (where T_{uti} and T_d

⁷ Cp. Arrow's motivation for the formal definition: "In its pure form, [dictatorship] means that social choices are to be based solely on the preferences of one man" (Arrow, 2012, 30).

deontically disagree), we can infer that it is only in virtue of the right- and wrong-making feature specified by T_{uti} that φ is right according to T_{duti} . But in the former case (where T_{uti} and T_d deontically agree), it may still be true that φ is right according to T_{duti} *in virtue of* the right- and wrong-making feature specified by the *deontological* theory T_d . The utilitarian theory would simply happen to give the correct verdict here (in agreement both with the basic deontological and the combined dutilitarian assessment), but it would do so for the wrong reason. Thus, over the whole set of possible actions there will be *two kinds* of right- and wrong-making features doing the work in T_{duti} : deontological features (in all cases in which T_{uti} and T_d deontically agree) and features relating to the consequences of actions (in the other cases). And this is precisely what Substantive Dutilitarianism asserts.

Hence, the existence of a single basic theory that “dictates” (or, as we could also say, perfectly predicts or mimicks) the dutilitarian verdicts is fully compatible with the guiding idea of Dutilitarianism if we individuate moral theories substantively. Consequently, for the Dutilitarian Impossibility Theorem to become a genuine problem for Dutilitarianism that must be solved by embracing Gradualism (which is the essence of the Conditional Rescue Claim), the argument *must* assume extensional individuation of theories (cp. the covert but explicit acknowledgement in Peterson, 2025, 4n8). Yet, as shown in section 1, the Gradualist Dutilitarian Project is *committed* to substantive individuation. This makes the case for the Conditional Rescue Claim inconsistent.⁸ Vice versa, if dutilitarians consistently adopt substantive individuation, the argument will fail, because then the violation of Nondictatorship *is* compatible with the guiding idea of Dutilitarianism.

A second reason leads to the same conclusion. Recall that Dutilitarianism — like its contenders, e.g., utilitarianism or deontology — is a *class* of moral theories, namely the class \mathcal{T}_{DUTI} . The explication (developed above at the end of section 1) of what makes a theory dutilitarian, i.e., an element in \mathcal{T}_{DUTI} , is permissive and does not rule out the possibility that, for a given set of deontological theories $\{T_d^i\} \subseteq \mathcal{T}_D$ and a given set of utilitarian theories $\{T_{uti}^i\} \subseteq \mathcal{T}_{UTI}$, there are *several* dutilitarian combinations $T_{duti}^1, \dots, T_{duti}^n \in \mathcal{T}_{DUTI}$ with distinct deontic contents. The Arrowian Dutilitarian Impossibility Theorem shows that for each single dutilitarian combination $T_{duti}^k \in \mathcal{T}_{DUTI}$ that satisfies the first three Arrowian conditions, there will be one basic theory (either from the deontological set $\{T_d^i\} \subseteq \mathcal{T}_D$ or from the utilitarian set $\{T_{uti}^i\} \subseteq \mathcal{T}_{UTI}$) that is decisive. But this is quite compatible with the possibility that, for different dutilitarian combinations T_{duti}^k and T_{duti}^l , the basic theories that are decisive in each case come from different *classes* of basic theories; e.g., for T_{duti}^k , the decisive basic

⁸ In response, Peterson might reply that the relevant co-extensionality of two substantively different theories must be co-extensionality *for all possible rankings* and that, according to Hume’s Dictum, necessarily co-extensional properties are numerically identical. Then the two theories would ultimately specify the very same right- and wrong-making features and the case for the Conditional Rescue Claim would not be inconsistent. However, I do not think that appealing to Hume’s Dictum dialectically strengthens Peterson’s argument: It seems to me that *being an equilateral triangle in Euclidean space* and *being an equiangular triangle in Euclidean space* are distinct properties, even though they are necessarily co-extensional. Similarly, *being the smallest prime number*, *being the only even prime number*, *being the first even number*, and *being the positive square root of four* are all distinct, albeit necessarily co-extensional, properties. I’d like to thank an anonymous referee for helpful comments that prompted me to discuss this possible line of response.

theory might be a utilitarian theory T_{uti}^3 , while for T_{duti}^l , the decisive basic theory might be a deontological theory T_d^7 .

But we can now see that even with the established impossibility result, there still *is* a meaningful sense of the claim Dutilitarianism is a genuine compromise between utilitarianism and deontology (remembering that just like the Conditional Rescue Claim and the claim about the Gradualist Dutilitarian Project, this claim must be about the *classes* of dutilitarian, utilitarian and deontological theories): If we consider the entire set of dutilitarian theories (i.e., consider Dutilitarianism properly, as a class), then the set of its decisive basic theories — i.e., the set of theories T for which it is true that for some dutilitarian combination $T_{\text{duti}}^k \in \mathcal{T}_{\text{DUTI}}$ satisfying the first three Arrowian conditions, T is a basic theory of T_{duti}^k and T is decisive for T_{duti}^k — may consist of theories from *both* the class of deontological theories \mathcal{T}_D and the class of utilitarian theories \mathcal{T}_{UTI} . If this is the case, the set of decisive basic theories — and thus *the set of theories determining the deontic content of Dutilitarianism (as a class)* — will be a mixture of deontological and utilitarian theories; and in this sense, Dutilitarianism (as a class) will be a compromise between deontology and utilitarianism.

Hence, there is a second way to render the guiding idea of Dutilitarianism compatible with the violation of Nondictatorship. This way is available as long as there is *no single class* of theories \mathcal{T} (either \mathcal{T}_D or \mathcal{T}_{UTI}) such that each basic theory that is decisive for some dutilitarian combination comes from \mathcal{T} . Peterson's Dutilitarian Impossibility Theorem does not establish this. It only shows that

- (A) for each dutilitarian combination, there is a basic theory which is decisive for that combination and which is a member of one of two classes of theories.

The impossibility result that is needed would have to make a claim with a different order of the relevant quantifiers, viz. that

- (B) there is one of two classes of theories such that for each dutilitarian combination, there is a basic theory which is decisive for that combination and which is a member of that class of theories.

Claim (B) would indeed undermine the idea that Dutilitarianism (as a class) is a genuine compromise between two other classes, deontology and utilitarianism, and it would really threaten the internal coherence of Dutilitarianism. However, (B) is not implied by (A). So once again, Peterson's Dutilitarian Impossibility Theorem does not raise a challenge that gives dutilitarians compelling reason to resort to Gradualism.

3 Lessons

I have argued that Peterson's case for the Conditional Rescue Claim — that Gradualism rescues Dutilitarianism's internal coherence — is far from conclusive. This is because, in order to present Dutilitarianism as a distinct contender to monistic unification projects, Peterson must individuate moral theories substantively; yet the challenge that Gradualism is meant to solve poses a real challenge only under extensional individuation. Moreover, even if the Conditional Rescue Claim could be made conclusive in some way I yet fail to see, we should not forget the conditional nature of the claim:

to the extent that Gradualism is hard to defend, the conditional claim can also be read as a *reductio* of Dutilitarianism. Since the distinctive point of Dutilitarianism is to propose an *alternative* to monistic unification projects in moral theory, we might wonder whether that route is promising at all. Perhaps we should not abandon monistic unification attempts too early, but instead try harder.

Nevertheless, I believe that at the heart of Peterson's case for the Gradualist Dutilitarian Project lies another rationale that deserves serious consideration when we aim to make progress in moral theory: Beyond a Sidgwickian-Parfitian rational systematization of our moral practices, we might also aim at *conceptually improving* these practices by making room for expressing more fine-grained, more differentiated overall moral assessments, or, in Peterson's words, "more nuanced verdicts" (Peterson, 2025, 16; cp. p. 13).

Arguably, this may require amendments to our conceptual framework. But at which level should we make these amendments? Peterson's Gradualism is a revisionary claim about the very first-order concepts *in which* we make those verdicts we wish to make sense of by rational systematization — concepts like MORALLY RIGHT or MORALLY WRONG. While Sidgwickian-Parfitian monistic unification projects share the commitment to making room for more differentiated overall moral assessments, they have taken different routes — without abandoning binary rightness: they have, e.g., separated blameworthiness verdicts from deontic verdicts and from moral assessments of the value of a person's character or motive; they have differentiated between what we morally ought to do, what we rationally ought to do and what we ought to do in the evidence-relative sense; and explored the relation of these various senses of OUGHT to what we have sufficient or most (moral) reason to do.

I do not claim that these attempts to make room for more nuanced verdicts are more likely to succeed than Peterson's Gradualism. I merely want to highlight that they operate at a different level: Rather than revising or re-engineering the first-order concepts *in which* we make those moral assessments that we want to make sense of by rational systematization, monistic unification projects typically make distinctions in the meta-theoretical, second-order vocabulary *with which* we rationally systematize moral assessments (and with which we thereby make sense of the first-order concepts). This is yet another respect in which the Gradualist Dutilitarian Project differs from competing attempts to make progress in moral theory.

If we attempt to tentatively draw a more general lesson from Peterson's "Dutilitarianism," it might be this: In order to make progress in moral theory by making room for more nuanced moral assessments, we may either tinker with the first-order terms in which we make those assessment, or tinker with the meta-theoretical terms with which we want to make sense of those assessments. Perhaps we should pursue the potentially more revisionary first route — i.e., Peterson's route — only after we have exhausted the second and after we have ensured that the meta-theoretical framework is fine-grained enough to allow for *all* the nuances in our more nuanced moral assessments — such as the difference between " φ is right if and only if F" vs. " φ is right if and only if *and because* F" that we need in order to individuate theories substantively.

Thus, although the Gradualist Dutilitarian Project ultimately fails, it does so for an interesting reason — which seems to be the most common way to make progress in moral theory.

Funding Open Access funding enabled and organized by Projekt DEAL. No funds, grants, or other support was received.

Data availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Competing interests The author declares no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arrow, K. J. (2012). *Social choice and individual values* (3rd ed.). New Haven/London: Yale University Press.
- Carlson, E. (2025). Peterson's gradualist hypothesis. *Asian Journal of Philosophy*, 4(94), 1–8.
- Hylland, A. (1986). The purpose and significance of social choice theory: Some general remarks and an application to the 'lady chatterley problem'. In J. Elster, & A. Hylland (Eds.), *Foundations of Social Choice Theory* (pp. 45–73). Cambridge: Cambridge University Press.
- Parfit, D. (1987). *Reasons and persons*. Oxford: Clarendon Press. Reprinted with further corrections.
- Parfit, D. (2011). *On what matters* (Vol. 1). Oxford/New York: Oxford University Press.
- Peterson, M. (2025). Dutilitarianism. *Asian Journal of Philosophy*, 4(60), 1–22.
- Putnam, H. (1994). *Words and life*. Cambridge, MA/London: Harvard University Press.
- Seidel, C. (2019). New wave consequentialism: An introduction. In C. Seidel (Ed.), *Consequentialism: New directions, new problems* (pp. 1–28). Oxford/New York: Oxford University Press.
- Sidgwick, H. (1981). *The methods of ethics* (7th ed.). Indianapolis/Cambridge: Hackett.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.