



CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of Political Studies

Master thesis

2023

Martí Bridgewater Mateu

CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of Political Studies

Martí Bridgewater Mateu

Borderline duties and fuzzy values
An analysis of vagueness in ethics

Master thesis

Prague 2023

Author: Martí Bridgewater Mateu

Supervisor: Prof. Dr. Vuko Andrić, Ph.D.

Academic Year: 2022/2023

Bibliographic note

BRIDGEWATER-MATEU Martí. *Borderline duties and fuzzy values. An analysis of vagueness in ethics*. Mater thesis. Charles University, Faculty of Social Sciences, Institute of Political Studies.

Supervisor Prof. Dr. Vuko Andrić, Ph.D.

Abstract

The aim of this thesis is to argue that the problem of overdemanding obligation in consequentialism within normative ethics is fundamentally a problem of vagueness in the specific sense in which the issue is treated within the philosophy of logic and language. This realisation leads to a substantial reframing of the question with two main takeaways: the first is an undermining of the plausibility of maximising consequentialism, and the second is that certain fairly intricate positions taken on issues of the philosophy of logic bear heavily on commitments we may adopt at the level of normative ethics.

Abstrakt

The aim of this thesis is to argue that the problem of overdemanding obligation in consequentialism within normative ethics is fundamentally a problem of vagueness in the specific sense in which the issue is treated within the philosophy of logic and language. This realisation leads to a substantial reframing of the question with two main takeaways: the first is an undermining of the plausibility of maximising consequentialism, and the second is that certain fairly intricate positions taken on issues of the philosophy of logic bear heavily on commitments we may adopt at the level of normative ethics.

Klíčová slova

Overdemandingness, satisficing consequentialism, scalar consequentialism, vagueness, sorites paradox

Keywords

Overdemandingness, satisficing consequentialism, scalar consequentialism, vagueness, sorites paradox

Range of thesis: 115.550 characters (with spaces)

Declaration of Authorship

1. The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.
2. The author hereby declares that all the sources and literature used have been properly cited.
3. The author hereby declares that the thesis has not been used to obtain a different or the same degree.

Prague 27/07/2023

Martí Bridgewater Mateu

A handwritten signature in black ink, consisting of several overlapping, stylized loops and lines, positioned to the right of the author's name.

Institute of Political Studies
Master thesis proposal

Topic Characteristics / Research Question(s):

Vagueness is a pervasive and often unavoidable phenomenon involving well-studied (and ill-understood) logical anomalies. In the somewhat restricted sense relevant to this project, it is understood as the feature of concepts and expressions by which they clearly apply to some cases and not others, but there is no sharp boundary between the two (Oms & Zardini, 2019). Much research has been done on the adequate logical handling of vague predicates, which has led to the development of competing accounts of vagueness in general, and solutions to the closely related Sorites paradox in particular. While some such accounts remain within classical logic (epistemicism being the most relevant), most potential solutions involve the development of some alternative non-classical logic, with the most prominent (but not only) examples being sub- and especially supervaluationism, dialetheism and fuzzy logics.

These competing analyses of vagueness have mostly been developed with a focus on descriptive predicates like tallness and baldness, but vague predicates also abound in normative contexts, as with the basic notion of “good” and the standard deontic categories of permissibility and obligation. Research in this area exists but is scarcer, and rarely involves the more formal analytical tools that have been developed for the logical handling of vagueness in general. This is the gap this thesis would aim to address.

Different logical approaches to vagueness have also been taken to weigh on debates about moral realism, and connections are certainly interesting to explore, given how they involve different types of challenges to the notion of a straightforward truth value for ethical statements. Discussions usually focus on the issue of moral realism (Asgeirsson, 2019), but much could be gained from exploring more specifically how notions of naturalism in metaethics relate to distinctions between vagueness as a natural property (such as the personhood of a zygote-baby continuum), and as an intrinsic property of moral predicates bypassing natural properties, such as when wondering how bad actions need to be to merit impermissibility. In this regard, it would be fruitful to pursue a deeper examination of the moral intuitions behind attempts at formalising fuzzy deontic logic (Dellunde & Godo, 2008). At the level of normative ethics, several problems seem to display some of the conceptually puzzling features of vagueness. One such notable case concerns the definition of traditionally binary deontic categories in relation to scalar notions of value, which is an especially pressing issue for consequentialist approaches to ethics. A flexible logical understanding of vagueness could shed new light on discussions on the merits of and alternatives to maximising consequentialism, for example in relation to Pettit’s objections to satisficing consequentialism (Slote & Pettit, 1984) and Norcross’s defence of scalar consequentialism (2020).

Similar insights could be gained in relation to so-called collective harm problems, in which a very large number of agents find themselves equally contributing to a

tiny extent in bringing about a very bad outcome, which makes it difficult to understand the moral position of the individuals involved. Many situations of great contemporary relevance are of this kind, like individual decisions in relation to climate change and electoral participation. The crucial role of vagueness in these contexts is made especially clear by the emergence of context-specific versions of the Sorites paradox (Nefsky, 2019), the understanding of which could be greatly improved by a systematic analysis under the different candidate solutions for the paradox in general.

It has also been argued that vagueness also plays a prominent role in relation to the problem of incommensurability of values (Hsieh & Andersson, 2021). However, discussions about the extent to which incommensurability is a type of vagueness rarely delve into the specifics of how different formal and analytical approaches of vagueness might offer more or less plausible accounts of the problem.

Working hypotheses:

1. The logically perplexing features of vagueness confound many discussions in moral philosophy.
2. Different accounts of the semantics of vagueness are not neutral with respect to questions of ethics and metaethics.
3. Differences between natural and ethical notions of vagueness involve distinct logical forms for ethical statements.

Methodology:

In the first part of the thesis, I will provide an overview of the logical challenges posed by vagueness, with special emphasis on the issue of their unavoidability. After a short clarification on how vagueness is to be understood and how it relates to the Sorites paradox, I will provide a short historical contextualisation of the fit between vagueness and different logics, particularly covering the discomfort with vagueness that characterised the rise to prominence of standard logic. I will then provide a summarised account of the main candidate accounts of vagueness that feature in contemporary debates on the matter, especially focusing on the philosophical interpretation of the formalisms they involve, and the trade-offs in intuitions that they offer.

In the second part of the thesis, I will provide an overview of ethical problems in which vagueness seems to play a prominent part. On the one hand, I will produce a rough literature review of the literature explicitly relating problems in moral philosophy (broadly construed) with vagueness, with an eye on both portraying the current state of the art and assessing the degree of formal and analytical specificity of the connections that have been made. On the other, I will complement this with an attempt at identifying discussions in moral philosophy in which vagueness plays an important conceptual role, even if this is not explicitly acknowledged in the discussions in question. I will do this by showing

that certain concepts and patterns of reasoning display strongly analogous behaviour to that of cases that are generally acknowledged as vagueness.

In the third part of the thesis, I will explore how differences in the understanding of vagueness in general provide grounds for the reassessment of the discussions covered in the previous part. For this, I will resort to the formal resources of non-standard logics (and standard logic whenever relevant) in cases in which this clarifies how this ought to change the state of play. Again, special attention will be paid to the trade-offs in intuitions that different approaches may incur.

Contents

INTRODUCTION	2
WHAT IS THE PROBLEM WITH CONSEQUENTIALISM?	4
1.1. WHAT IS CONSEQUENTIALISM?	4
1.2. IS THE MAX TOO MUCH TO ASK?	11
1.3. SHOULD WE BE ASKING FOR ANYTHING IN THE FIRST PLACE?	18
2. WHAT IS THE PROBLEM WITH VAGUENESS?	23
2.1. WHAT IS VAGUENESS?	23
2.2. WHAT ARE MY OPTIONS?	26
3. ONE PROBLEM, SEVERAL STRATEGIES, ANY SOLUTIONS?	32
3.1. TWO PROBLEMS OR ONE?	32
3.2. WHAT IF THERE IS A SPECIFIC THRESHOLD?	35
3.3. WHAT IF THERE IS NO SPECIFIC THRESHOLD?	38
CONCLUSION	44
BIBLIOGRAPHY	46

Introduction

This is a paper mixing moral philosophy and the philosophy of logic and language. More specifically, it is a paper on the relationship between two problems, which I believe are in some sense the same problem: the problem of overdemanding obligation in normative ethics, and the problem of vagueness and the sorites paradox in the philosophy of logic, truth and semantics.

I do not particularly think these two approaches are in competition methodologically, and yet I think it is fair to say that authors in each of the two fields have generally looked more inward than outward in giving shape to their positions, arguments and assessments thereof. It goes without saying that this has nevertheless resulted in enormously insightful debate, and to the modest extent that I have been able to use these sources as the background for my reflections, they will be derivative to some considerable extent. As will be easy to observe throughout the paper, if my contribution should advance either discussion, this is not so much in the full reformulation of available philosophical positions, but more in the way they are assessed. In making this connection, I hope to undermine any implicit impressions that it would be wise to pursue the issues of vagueness and moral overdemandingness in complete separation. While this connection might be more fruitful in terms of raising issues than in terms of settling them, I take that to be well within the legitimate remit of a philosophical paper.

I will circumscribe my comments to consequentialism as a family of views in ethics, but I do not think the issues highlighted here only concern consequentialists. Deontologists are likely to find themselves similarly worrying over whether certain situations involve serious or relevant enough concerns to merit the establishing of duties, and all virtues discussed in virtue ethics seem to be defined along a continuum of traits of conduct and character that also contains vices, and so they will also face similar issues in relation to vagueness. With this in mind, I have attempted to construe consequentialism as broadly as possible, but I am sure that philosophical reflection on the interplay between theories of vagueness and non-consequentialist theories of ethics would be worth undertaking elsewhere in fuller detail.

Also, concerns beyond the strict remit of philosophy are not out of sight completely either, as consequentialism is arguably among the most influential philosophical ideas in the

modern world, perhaps especially in matters of politics and economics. From the point of view of intellectual history, all of contemporary economic analysis is deeply influenced by consequentialism and, conceptually, at least a lot of it is at present as well. It is debatable whether, for instance, normative recommendations issued in welfare economics somehow entail or imply certain forms of utilitarianism, but it is difficult to contest that they are often interpreted in this vein. Consequentialism seems to answer some questions that some rival moral theories do not (regardless of whether we take this to be a wrong answer), such as questions about which bridges to build, about which tax rate to set, which party to vote for, and so on, and if it is not the only framework one can adopt in answering them, it certainly seems to be a very influential one. The paper bears on these matters insofar as it discusses how consequentialism can be modified to retain its plausibility, and what kinds of conceptual commitments this would require.

In the first section of the paper, I will attempt to define consequentialism in a way that is as broad as possible, but still articulated enough that I can pinpoint which components of a consequentialist view specifically make it liable to the problem of overdemandingness. I will then proceed to discuss the gravity of this problem, and how different versions of consequentialism fare in addressing it. In the second section, I will seem to change tack by introducing the problem of vagueness and the sorites paradox, together with an overview of competing theories of this phenomenon. In the third section, I will more specifically flesh out how this problem features in the debate on overdemandingness, and I will discuss how different combinations of positions in both debates would fare overall.

What is the problem with consequentialism?

1.1. What is consequentialism?

Consequentialist views in ethics hold that moral normativity is fully determined by consequences (Sinnott-Armstrong, 2022). Consider the following scenario, illustrating the common issue of how strictly we are morally bound by our promises: I promised my friend Alex that, when faced with two options A and B at some point in the future, I would choose A. Suppose further that I find myself at that later point in time, and new information indicates that choice of B would actually result in the exact same outcome as A, with the only difference that Alex would gain something very nice that I know Alex likes, like a bag of cashews or ownership of a house. Should I now prioritise my having made a promise and choose A, or should I prioritise the gain in value associated with B?

Under a consequentialist view of ethics, the question becomes just a matter of which option involves most value overall¹. If we assign no substantial value to the keeping of promises *per se* (like most consequentialists), then we will choose option B, as any tiny amount of intrinsic disvalue associated with breaking promises will easily be offset by other gains of value in benefit to Alex. The main idea here is that morality is forward-looking and circumscribed to the value of what will actually occur if we make different choices. This is to be contrasted with deontological views, which would typically forbid the breaking of my promise implied by a choice of B, and require its keeping by a choice of A. The only thing determining which course of action we are obliged to pursue is the fulfilment of duties (in this case, a duty to uphold a promise made in the past), and not by the comparison of how options score on some measure of value (Alexander & Moore, 2021).

To clarify, it is not quite the specific first-order moral view that conceptually differentiates one kind of view from the other, but rather the way of framing the question on what we ought morally to do. As indicated earlier, consequentialists believe that what we should do is just a matter of what is more valuable, whereas deontologists believe it depends on

¹ Here and in the remainder of this paper, I will be discussing act-consequentialism as opposed to rule-consequentialism, because I find it simpler and more theoretically relevant. However, at least roughly analogous points to the ones made here would apply if rule-consequentialism were considered instead.

what our duties are and which courses of action would involve fulfilling or neglecting them. Although they would be deviant in the sense of being outliers and going against the grain of the philosophical insight or intuition typically defended by each kind of ethical theory, degenerate positions are possible: a deviant consequentialist view could be that nothing is as valuable as the keeping of promises, resulting in a choice of A, whereas a deviant deontological view might be that we have many duties but keeping our promises is not one of them, so we may choose B.

Like others (Scanlon, 1998), I suspect that this means deontological and consequentialist theories are in some sense mutually reducible, with philosophically interesting disagreements between them mostly translatable to disagreements about value within a sufficiently complex axiology. For instance, if a staunch deontologist chastised me for violating an individual's rights in order to protect the rights or interests of millions of others, I fail to see how they could avoid the commitment to the claim that they in some sense value my respect of that individual's rights more highly. Be that as it may, a more careful discussion of this rather controversial claim would be well beyond the scope of this paper, but it should suffice to indicate how, although in the following I will be speaking within the remit of a very broadly construed form of consequentialism, there might be ways the point I will be making could be reframed to apply to deontological views as well.

Within consequentialism, many relevant distinctions can be made in order to articulate different specific and philosophically well-motivated versions in competition (Brink, 2005; Sinnott-Armstrong, 2022). For the purposes of the present analysis, it ought to suffice to distinguish between two components of a consequentialist view, which I will call its theory of value and its decision criterion². In succinct terms, a theory of value assigns a degree of value to the outcome associated to each available option in a choice,

² I am departing a bit from common usage with this somewhat *ad hoc* distinction, as it is perhaps more common in the literature to distinguish between theories of value and theories of *rightness*. I hope that the following discussion in this section will show why I believe this would be too restrictive and imprecise, as it precludes reasonable philosophical concerns about normativity and deontic concepts.

and a decision criterion tells us how these evaluations are to normatively inform our choice.

A very important and confusing problem in ethics is whether we can always compare how valuable things are and, if so, how. Many consequentialists are monists about value, i.e., they believe that there is only one intrinsic or non-instrumental value, with all other apparently different values being reducible to one same overarching notion. Most notably, utilitarianism can be thought of as the position of being both a consequentialist and a value-monist, committed to the claim that all value ultimately stems from well-being – generally understood as something like overall aggregate welfare (Crisp, 2021; Sinnott-Armstrong, 2022). Other consequentialists are pluralists about value but believe trade-offs between values of different kinds can be resolved satisfactorily, whereas yet some other consequentialists believe certain different kinds of value are fundamentally incomparable (Hsieh & Andersson, 2021).

This is hopefully illustrated by the following example. Suppose that one of three policies X, Y and Z will be carried out among a society with one citizen of kind A and 99 citizens of kind B, all equally deserving of welfare. Each would result in the distribution of welfare for citizens of each kind indicated below (expressed in cardinal units of welfare w).

	X	Y	Z
A's welfare [1] (w_A)	1	101	0
B's welfare [99] (w_B)	1	0	0
Agg. Welfare (Σ_w)	100	101	0
Gini coefficient (σ_w)	0	0.99	0

Let “ $X > Y$ ” mean that X is better, i.e., more valuable overall than Y. If only aggregate welfare is intrinsically valuable, then $Y > X > Z$. On the other hand, it is also plausible (more so to my mind, although that is beside the point) that, compared to X, Y is lacking in some value that is not reducible to aggregate welfare, such as egalitarianism about the distribution of said welfare (this can be measured by some indicator like the inverse of

the Gini coefficient, for instance). However, if only equality were valuable, and not aggregate welfare, then $Z \approx X \succ Y$, which hardly seems plausible.

If we wish to commensurate both values at stake, we typically need to define some functional form for the overall value of each alternative. One way would be to introduce a lexicographical order, i.e., giving increases in one primary value absolute priority over increases in the other secondary value, which would effectively only function as a tie-breaker. This would entail that $X \succ Z \succ Y$ if equality were lexicographically placed above aggregate welfare, and $Y \succ X \succ Z$ if the order is reversed. Such an ordering would, perhaps implausibly, be “blind” to gains in the secondary value so large that they intuitively merit sacrificing tiny changes in the primary value. Other specifications of this comparability relation might be more intuitive, such as the definition of overall value as a weighted average of the indicators for different values, which would give the rather plausible $X \succ Y \succ Z$ given appropriately defined weights³. Subtler issues can be modelled by sophisticating this kind of functional form, for instance by defining weighing factors as functions of value levels, among others.

An alternative to these different ways of considering both or either value in ordering options would be to deny any such commensuration and claim that, although both aggregate welfare and equality are valuable, they are incomparable, which will make it impossible for us to rank alternatives by overall value. This means that comparisons can be established with respect to one value, but not across them, i.e., trade-offs cannot be resolved satisfactorily. The main confusion to dispel here concerns parity: incomparabilists do not claim that the difficulties in comparing X and Y should lead us to taking them to be equally valuable, but rather to refraining from making any statements concerning how they compare in terms of overall value, as no such notion is warranted under the view. Furthermore, it is worth noting that one may reasonably be a comparabilist about certain values and an incomparabilist about others: maybe trade-offs

³ This will not be discussed in the present paper, but the point of parity between comparable values has been modelled in terms of vagueness in the precise sense in which the notion is discussed in later sections (Asgeirsson, 2019; Hsieh & Andersson, 2021), which I believe further reinforce my emphasis on the importance of the interdependence between positions we adopt in ethics and metaethics, and those we adopt in the philosophy of language and logic.

between equality and aggregate welfare can be resolved satisfactorily, but trade-offs between them and beauty cannot.

These are some of the main issues we face when defining a theory of value, i.e., when systematising how different states of affairs are to be described and perhaps compared in terms of their worth. Suppose now that we settle on a certain way of evaluating the outcomes of the implementation of policies X, Y and Z, and that some agent is actually faced with a choice between them. As explained earlier, consequentialism requires that moral normativity be defined in terms of the value of alternatives, but even while agreeing on the evaluation of alternatives, different people might reasonably take different views on what exactly is the moral bearing of this evaluation on the agent.

By far the most common way of defining a decision criterion is to use deontic operators, which are often taken to be the main or only kind of action-guiding concepts involved in ethics (McConnell, 2022), although as we will see in later subsections, this is a controversial claim. By deontic operators I mean terms that we ascribe to alternatives which are available to the decision-maker, like “permissible”, “obligatory” and “forbidden”, and also “right” and “wrong”⁴, in contrast with evaluative terms for states of affairs, such as “good” and “bad”. Some quick examples might help to clarify the distinction. In normal conditions, choosing to gratuitously shout in a public library is a wrong alternative to pick, and stegosauruses being extinct is an example of a bad state of affairs. Since other agents’ choices amount to states of affairs for us, another person’s choice to gratuitously shout at a library where we are trying to work will generally result in a bad outcome, but rightness and wrongness are circumscribed to the point of view of

⁴ The most important pairs of terminological contraries involved in deontic statements would seem to be *permissible/impermissible*, *allowed/forbidden*, *obligatory/optional*, and *right/wrong*. The easiest relationship to establish is probably between the first two pairs, which are synonymous respectively. The relationship between these and the third pair is the following: actions are either obligatory to not do, obligatory to do or optional, with the first being thereby forbidden, and the other two being permissible, as well as obligatory or optional, respectively (McNamara & Van De Putte, 2022). The relationship of these concepts with the fourth and final pair is the least obvious: while it is common to equate wrongness with impermissibility, with rightness synonymous with permissibility, this latter term is sometimes more narrowly reserved for obligatory actions exclusively, with optional actions labelled neither right nor wrong (Timmons, 2001). For the remainder of this paper, I will generally follow the first broader characterisation, but at times philosophically motivated distinctions will be made explicit.

agency. This is to say that it is neither linguistically natural nor conceptually rigorous to say that it is wrong that there are no more stegosauruses, because it was never the outcome of a choice made by a morally responsible agent.

Given this distinction between the moral consideration of states of affairs and choice alternatives, it might seem that consequentialism simply requires that we evaluate the latter in proportion to the former, i.e., that we consider available actions “right”, “permissible”, “obligatory” and the like to the extent that their outcomes are good. This kind of take, however, is overwhelmingly rejected in contemporary forms of consequentialism because, unlike degrees of value or goodness/badness, these deontic operators are generally understood as binary, dichotomous pairs, also called “all-or-nothing” properties of choice alternatives (Brink, 2005; Howard-Snyder & Norcross, 1993; McNamara & Van De Putte, 2022; Norcross, 2020; Sinnott-Armstrong, 2022)⁵. The idea, which will be revisited in later sections of this paper, is the following: although the gratuitous death of two people is worse than the gratuitous death of one person, in normal circumstances it is no more permissible to bring about the latter than the former, as both are equally fully impermissible. This means that comparisons between the permissibility of choices are substantially simpler than comparisons between the goodness and badness of outcomes or situations: permissible actions are more permissible than impermissible actions but given any two permissible actions none is more permissible than the other, and exactly analogous points hold for other deontic operators.

This conceptual behaviour of deontic statements is meant to allow for some kind of connection between normative statements susceptible of truth-evaluation and more immediately action-guiding imperative statements. Deontic statements like “It is morally forbidden to evade taxes” or “It is morally permissible to jaywalk carefully” seem to be *prima facie* susceptible to truth-evaluation: they seem to mean something insofar as they have successful usage in communication, and it is fairly clear that some claims of this sort

⁵ It is very interesting to note that, despite its ill fate in contemporary philosophy, this notion of gradual right and wrong seems to be very explicitly what Mill meant by his highly influential definition of utilitarianism as the view that “holds that actions are right *in proportion* as they tend to promote happiness, wrong *as they tend to produce* the reverse of happiness [emphases added]” (Mill, 1861).

are more plausible than others. Something similar seems to happen with modal statements that avoid standard deontic operators such as “one must not avoid taxes” and “one may jaywalk carefully”, which are also apparently susceptible of being true or false, and indeed seem to follow respectively from the former two deontic statements (perhaps with a clause expressing something like “morally speaking”). In a similar vein, “one should not avoid taxes” and “one should be careful when jaywalking” seem also to be somehow entailed although, unlike in the previous case, the implication seems to go only in one direction: it seems possible in principle that one should take a bath despite not having an obligation to do so.

The situation is a lot more complicated when we consider imperative statements like “do not avoid your taxes” and “be careful when jaywalking”. They also in a sense seem to “follow” from the previous statements, but we are prone to philosophically denying them a truth value, and so this would not seem to be a common notion of implication. Furthermore, they are sometimes apparently susceptible of being operated on with truth-functional connectives, both in connecting imperatives between them (“bring coffee or do not come”), and in connecting imperatives with declarative sentences expressing propositions (“stop singing or it will start to rain”). On the other hand, as is familiar to anyone who has ever done any coding, “if stop singing, then it will start to rain” does not seem to make any sense, whereas “if it starts to rain, stop singing” does. This suggests an asymmetry between imperatives and should-statements in relation to the applicability of Convention T (Tarskian biconditionals) as a basis on which to argue that they are truth-bearers expressing something like propositions.

General philosophical interest aside, these issues seem relevant because imperatives are perhaps the clearest instance of action-guiding statements. On the one hand, this suggests that it might be fruitful to somehow analyse how different deontic formulations of decision criteria imply different imperative statements, in order to clarify comparisons. On the other, it seems to bear on the necessity of deontic statements in defining a decision criterion within a specific consequentialist theory. Indeed, it seems that one could simply define an algorithm using evaluations as input and decisions as output, such as “Choose the alternative with the highest overall value”, and indeed holding a view in which such an imperative obtains need not commit one to any specific deontic claims such as views on obligation and the like.

Perhaps the most puzzling connection between a deontic statement and its normative implications concerns the concept of supererogation, which is common usage in moral philosophy to indicate that some action is above and beyond the call of duty, and indeed is central to the issue that will be discussed in the following subsection. A statement like “It is supererogatory to donate all of one’s money to ending world hunger” seems to entail that it is false that one must do so, and that one may morally do so or not, but there is a strange implication that one in some sense also should. This “should”-entailment, however, seems difficult to translate directly in imperative mode, despite its clearly moral kind of compulsion. One can program for the decision of a permissible action by randomising choice among permissible options, but it does not seem possible to convey to a machine an instruction like “you must do one thing and it may be any of these, although you should take their different value into account, without these considerations entailing any specific compulsion”. This kind of non-obligatory, defeasible normative tilting towards an option, however, seems to be exactly what we mean when we discuss the supererogatory nature of certain options, and so it is necessary to tread with caution when comparing the normative commitments of different ways of specifying deontic claims.

1.2. Is the max too much to ask?

In the previous subsection, we have seen that a consequentialist view in ethics tells us how to make moral choices, which involves telling us (i) how to assign value to the outcomes of different available options (ideally, in some comparable order), and then (ii) how to act (generally, but not necessarily, in terms of deontic operators). We have also seen, albeit briefly, the philosophical motivation behind different theories of value, and the basic conceptual apparatus with which decision criteria are expressed. In this subsection and the next, I will describe a problematic feature within debates on consequentialist decision criteria. This problem is known in the relevant literature under the heading of “overdemandingness” (Sinnott-Armstrong, 2022). I will try to explain its philosophical motivation and summarise the state of the art.

The most common decision criterion among consequentialist theories is maximising, and indeed it is so common that it is sometimes built into definitions of consequentialism in

general (Shafer-Landau, 2018). Under a maximising consequentialist view, making an optimistic, i.e., value-maximising choice is obligatory (and therefore permissible and right), and making a non-optimistic choice is forbidden (and therefore impermissible and wrong). This is a very clear way of distributing deontic operators based on an evaluation of outcomes, and it is fairly intuitive: given the choice between two actions, one better than the other, it might seem reasonable that there is an obligation to choose the former over the latter. Maximising consequentialism simply obtains from iterating this reasoning for any number of options.

It would seem in principle that if incomparabilism about value is true, then maximising consequentialism is false. Maximisation implies an order relation in value, and incomparabilism about value specifically denies that there is an ordered relation of overall value, with ordered degrees existing only within each value but not across them. Suppose that we are in a situation like the earlier choice between policies, and that equality and aggregate welfare are indeed incomparable values in the sense described above. Since there is no way to resolve the trade-off, although we can and must choose between X, Y and Z, we will not be able to maximise. Since we cannot be obliged to do what it is impossible for us to do, it cannot be obligatory for the agent to maximise.

It has been suggested that maximising consequentialism can be made compatible with incomparabilism (Carlson, 1995). The idea is that we can weaken the maximising claim by either taking choice of X and Y to be neither right nor wrong (with only Z being wrong), or by making both X and Y right insofar as no overall better alternative is at hand. I am unconvinced by this line of argument, as it just seems a more stubborn fact that you can only have a maximum in an order, as indicated earlier, and so something must give. Clearly there is an obligation to avoid the wrong option Z, so it is either obligatory to choose X, or Y, or either. If neither of the two is right in the narrow sense of being obligatory, then it is obligatory to choose either, whereas if neither is right in the broader sense of being permissible, then we are committing ourselves to the puzzling claim, to be commented further in future sections, that an available option may be neither permissible nor impermissible. At any rate, I fail to see what is maximising about the view, as we would be neither requiring nor perhaps even allowing the obtaining of a maximum. If we take both to be right because of the inexistence of overall superior alternatives, then it must be in the sense of permissibility (since they are mutually exclusive choices). In that

case, I additionally fail to see the incomparabilism of the view, as it becomes indistinguishable in its normative implications from a comparabilist view claiming parity in effect between differently valued alternatives. Be that as it may, if this kind of maximise-any-value incomparabilist consequentialism turns out to be workable, then the following points on comparabilist maximising consequentialism would seem to apply.

So let us instead suppose that there is some way to compare the overall value of alternatives, i.e., that comparabilism is true. In short, the problem of overdemandingness for maximising consequentialism stems from the intuition that some non-optimific actions are not wrong (not impermissible), because they fall short of the maximum by a margin that is morally relevant *in kind* (it concerns differences in value), but not *in degree* (these differences in value seem too small to matter). Suppose that your theory of value is monism about aggregate welfare and your decision criterion is maximising consequentialism. Then it seems that, in normal conditions, perfectly everyday actions like buying a birthday present for a family member or going to a nice restaurant are morally impermissible on the same grounds whereby murder sprees are impermissible: since more human wellbeing would be achieved by dedicating that time and money to, say, alleviating world hunger, they fail to maximise and are plain wrong. Indeed, it seems reasonably plausible that no individual person has ever performed an exactly optimific action, which would imply the rather outlandish claim that no permissible action has ever been performed. This is clearly a very implausible view, and it has received a lot of criticism, but it is difficult to pinpoint exactly what is wrong with the view, and what ought to replace it, which is why many authors remain maximising consequentialists *by default*, even while accepting that it has strange implications (Wolf, 1982; Slote and Pettit, 1984; Sobel, 2007; Bykvist, 2009; Schaler, 2009).

It is important to distinguish between two different levels at which maximising consequentialism might be overdemanding. One would be this: since it often seems there are indefinitely many things one could potentially do, for whichever course of action we consider undertaking, it is generally easy enough to think up some marginally improved alternative that would establish the impermissibility of the original alternative, to the point that we may be unable to stop this discarding sequence and identify an exact maximum which would be allowed. This would constitute *cognitive* overdemandingness, insofar as it stems from our limited cognitive capacity to adequately compute often severely

imperfect information, which would be put under undue moral pressure. It has been plausibly argued that this is a very real problem for maximising consequentialism, but that it can be addressed by switching from an “objective” to a “subjective” formulation of our decision criterion, restraining consideration to rationally expected alternatives and consequences, as opposed to actual ones (Andrić, 2021). However, there is still a non-cognitive way in which maximising consequentialism is still plain morally overdemanding, as it still seems excessive to just prohibit alternatives that are only marginally worse than the best alternative I can reasonably conceive of, as shown in the examples discussed below.

One way of addressing these deeply counterintuitive implications might be to retain maximising consequentialism but introduce modifications to our theory of value. On the one hand, we may challenge welfarism in favour of (perhaps most plausibly), a comparabilist pluralist view on values that would recognise the worth of the kinds of things maximising utilitarians might forbid. For instance, if there is some value in having a meaningful connection with friends and family that does not fully reduce to its welfare effects, then it is less clear that more cost-effective ways of promoting welfare than cultivating such connections (such as fully devoting one’s time and money to alleviating world poverty) will always generate more value overall, and therefore permissible conduct might (plausibly) include both kinds of activities. On the other hand, it may be necessary for our theory to allow us to value our own welfare, connections, etc. more highly than those of others, so as to avoid awkward implications such as our having to refrain from buying our mother a birthday present if that money could be spent in a way that resulted in more mothers receiving presents, and so on. These seem to be the underlying issues at the heart of many objections to maximising utilitarianism, based on its supposed incompatibility with very reasonable everyday notions of a good life (Wolf, 1982; Sobel, 2007; Sinnott-Armstrong, 2022).

These are philosophically relevant points, but they concern a consequentialist view’s theory of value and not its decision criterion, whereas I believe that the overdemandingness problem is principally about the maximising component of a consequentialist view, and not any specific commitments on value assignments. I believe this because a version of the problem of overdemandingness seems to recur after each *ad hoc* axiological modification is introduced in order to mitigate the problematic

counterintuitive implications that result from staying within maximising consequentialism⁶.

These confounding factors are hopefully abstracted in the following hypothetical situation: suppose that an agent finds herself having to choose one out of N arbitrarily many alternative courses of action X_i , differing only in the number of times a certain constant positive cardinal amount of value α is generated, be it the only kind of value there actually is if value-monism is true, or some single kind of value with all other kinds being fixed if comparabilist value-pluralism is.

Alternatives	Cardinal value of outcome	Ordinal ranking
X_1	$N\alpha$	1
X_2	$(N - 1)\alpha$	2
...
X_{N-1}	2α	$N - 1$
X_N	α	N

The two columns on the right represent the assumption that $X_1 > X_2 > \dots > X_{N-1} > X_N$ i.e., that choice of X_1 would result in the best outcome, that its outcome is better than that of any other alternative, that conversely the outcome of X_N is and the like. If the true decision criterion is maximising, then, because its ordinal ranking is 1, X_1 is obligatory and the only right and permissible choice, with all other alternatives being wrong and impermissible to choose (and therefore obligatory not to do).

⁶ This is not to say that the problem of overdemandingness is unrelated to issues of axiology, as in particular a maximising decision criterion aggravates the lack of appeal of one's theory of value for those who disagree with it. For example, in the previous welfare distribution problem, both non-maximising welfare-monists and welfare-equality-pluralists might agree that it could be permissible to bring about either X or Y but not Z, whereas their maximising counterparts will disagree. In other words, the deontic commitments of non-maximising views are more robust to changes in the theory of value.

The previous discussion on the implausible implications of maximising consequentialism hopefully shows how this is not the only way of distributing deontic operators worthy of philosophical consideration, which brings us to the second decision criterion we will discuss here: satisficing. Satisficing consequentialism is the claim that there is some threshold below optimality for the distribution of the deontic operators of permissibility, obligation and rightness. The view is meant to alleviate the overdemandingness objection by making space for a more intuitively plausible standard for permissible action, and for the *prima facie* reasonable possibility that actions may be supererogatory, i.e., above and beyond the call of duty (Heyd, 2019). Put in terms of our abstract list of ranked alternatives, this amounts to the claim that for some X_K such that $X_1 \succ X_K$, any X_i such that $X_i \succcurlyeq X_K$ is a morally right and permissible choice, with all other alternatives being impermissible and our obligation being simply to choose a permissible rather than an impermissible choice.

The term *satisficing* was coined in the context of management and decision theory by Herbert Simon, as a portmanteau of *satisfying* and *sufficing*. It refers to a decision-making strategy by which decision-makers choose an alternative meeting some acceptability threshold rather than applying optimisation procedures, generally due to problems such as imperfect information or high decision-making costs (Simon, 1956). The most common example is the following: I am trying to sell a house and I receive a very generous offer, so I decide to sell rather than to make sure to find out exactly how much each potential buyer would pay in order to make sure that I sell for maximum price. The theory is meant to both descriptively explain and normatively justify the rationality of this kind of behaviour, within the approach of bounded rationality, which studies how rationality standards do and ought to modulate in response to limitations of different kinds, such as cognitive and environmental (Wheeler, 2020).

The normative justification for the rationality of satisficing behaviour is that it can actually amount to an expected optimisation strategy once all costs have been factored in: in the example of selling a house, the idea is that one wants to maximise net gain, and the cost of extending the search is expected to lead to a worse outcome than the acceptance of the offer. The view that it is rational to choose an alternative even when a better one is just as available is sometimes called “radical” satisficing (as opposed to rational satisficing or “suboptimising”), and it has been criticised severely and plausibly for its

normative indifference to normatively relevant information (Tucker, 2017). As noted by Slote himself in first using the concept to define satisficing consequentialism, satisficing in ethics entails precisely this kind of stronger and difficult claim involved in “radical” satisficing, as the point of the idea is that one may sometimes have better alternatives at hand but still forgo them for worse yet still permissible ones (Slote and Pettit, 1984). The hope for the view is that, even if “radical” satisficing might be normatively unjustified for matters of instrumental rationality (concerned with the choice of means for given ends), it could still hold for the purposes of ethics, which is also concerned with non-instrumental choices of better or worse ends.

An example adapted from Slote’s original presentation might help to illustrate the problem that satisficing consequentialism finds in maximising consequentialism’s commitments on moral permissibility. Suppose that an agent finds herself in her cabin in the woods in the middle of a snowstorm, when a good friend of hers who got lost starts knocking at the door asking for shelter. Suppose that the agent faces a choice between the following three options: A) let her friend outside to freeze, which would result in a terrible outcome, B) let her friend in and offer her some warm soup the agent was just having, and C) let her friend in and offer her a choice between the soup and some moussaka that is in the fridge. It is implausible that, even if the agent knew her friend actually prefers moussaka over soup, A and B are impermissible with only C being permissible (and therefore obligatory). Instead, common sense intuitions seem to be that A is impermissible and B and C are both permissible. Notice, however, that both B and C are equally available to the agent, and C is better, if only slightly, than B. Furthermore, C seems to be better than B with respect to the same values of hospitality and welfare-promotion that make A much worse than the other two. The specific intuition defended by satisficing consequentialism is that value need not be promoted up to the maximum, but up to some reasonable minimum, with value promotion beyond that point deserving merit, but not being required.

Although it has some defenders (Rogers, 2010; McKay, 2021), satisficing consequentialism is broadly speaking an unpopular view (Bradley, 2006; Tannsjo, 2019; Slater, 2020). Some critics contend that, unlike maximising consequentialists, satisficing consequentialists cannot produce any well-articulated means of identifying a salient point in any hypothetical distribution of value across choice alternatives that would constitute

the plausible threshold for permissibility. In addition, some complain that in its formulation of a decision criterion, satisficing consequentialism remains normatively indifferent about something that seems morally relevant, namely, the excess value that is forsaken by a choice of non-supererogatory minimum permissible value alternatives. They generally conclude that even though maximising consequentialism is counterintuitively stringent, there is also something to its consistently pointing in the right direction, so although it is highly revisionary of everyday intuition, such revision may be due after all. The few remaining defenders of satisficing consequentialists believe that this is too high a cost in intuition, and there must be some reasonable boundary that does not make us all permanent wrong-doers. And thus, the debate appears to be a difficult intuition trade-off, leading to an unsatisfactory standstill.

The problem here, I believe, is a common perplexing feature of trying to fold any conceptual continuum in two. As mentioned earlier, any plausible theory of value treats it as scalar or gradual, and any clearly understandable notion of deontic operators treats them as dual, but since under consequentialism changes in value ought seemingly to have some sort of normative implication, we seem to have insufficient means to express or conceptually articulate this. This issue will be developed in richer detail in later sections, but before that it will be useful to comment on a third solution that has been proposed along these lines in order to settle the debate, which has even smaller purchase than satisficing consequentialism.

1.3. Should we be asking for anything in the first place?

As mentioned before when discussing the extinction of stegosauruses, it is conceptually inadequate to say that bad situations not brought about by morally responsible agents could in any sense have been wrong for anyone to do. However, it seems perfectly natural to say that some things are better to do than others, in proportion to the relevant comparison between the value of their outcomes. Scalar consequentialism is a third alternative to maximising and satisficing consequentialism, and it is committed to the view that not only can we morally evaluate available choices in this way, but that this all a rigorous theory of ethics should aim at doing, without any significant non-rhetorical use

for notions of “rightness”, “wrongness”, “permissibility”, “obligation” and the like (Norcross, 2006b, 2006a, 2020).

The bulk of the widespread rejection of this view can probably be attributed to its radically revisionary character: whereas many authors would quite unremarkably define deontic categories into their very notion of ethics (Skorupski, 2010; Crisp, 2011; Singer, 2023), Norcross seems to be fundamentally denying them any serious role as concepts in moral philosophy. Still, the appeal of the view is hopefully clear in light of the problem of infinitely many degrees of value for binary deontic statuses: in scalar consequentialism, we can only say that certain actions are better to do than others, with some being best and worst, i.e., symmetrical concepts to the ones used to evaluate situations and choices. Norcross even goes as far as claiming that, just like with deontic categories, the notions of “good” and “bad” simpliciter are flawed insofar as they are a binary pair (Norcross, 1997).

I, for one, have sympathy for the view and its radical implications. In particular, I suspect the notions of moral permissibility and obligation may in some sense be more-or-less dead theological metaphors about penal arrangements in the afterlife. A suggestive historical comparison might be that of the concept of “sin”: whereas it once was at the very core of moral philosophy (Dawson, 2019), most philosophers today would likely treat it as interesting historically, but as conceptual deadweight for lucid and truth-conducive ethical analysis: any legitimate theoretical functions that the concept may have carried out in the past are translatable into more secular and clearly defined concepts, and more specifically theological implications can be rightly cast aside. Although such accounts might contribute to motivate and mitigate the cost of getting rid of traditional deontic categories, they are unfortunately well beyond the scope of this paper. Still, it is clear to see that, despite its interesting results and motivations, the view undeniably has important weaknesses: its cost to intuition is very high, and even though we may not be too clear about what we mean exactly with attributions of moral permissibility, it is a stretch from there to claim that we therefore don’t really mean anything by them. Differently put, it is difficult to conceive of non-*ad hoc*, well-articulated justifications of philosophical attitudes that are realist concerning value statements and anti-realist concerning deontic statements, which is exactly where scalar consequentialism seems to land.

Beyond these issues, the view has also been accused of falling short of a fully articulated guide to moral action. As mentioned earlier in relation to shouting in a public library, actions over which an agent has no control are states of affairs for that agent. One can claim that an action may be better than another in the sense that it is better for it to be carried out than another but, if this amounts to an evaluation in the sense of a description of certain value-related properties of states of affairs (as occurs when we accurately refer to our loss of stegosauruses), then it would not be action-guiding statement. The worry is that such comparative assessments of hypothetical situations involving one's own actions in the immediate future would not constitute an answer to a question about what we ought morally to do, but about how things would stand from the point of view of value.

Put more succinctly, for opponents of scalar consequentialism, evaluative statements *are not* normative statements, however closely related they may need to be within any consequentialist theory of ethics. Norcross's main defence from this line of argument involves a substantial shift from the mainstream regarding normativity and deontic statements, by claiming that morality, strictly speaking, demands nothing of us, but merely gives us moral reasons to act which are both normative and gradual in nature, with their strength proportional to the value of the outcomes under consideration. He takes these kinds of reasons to constitute a strong moral guide to action, without being formulated in terms of allowing and prohibiting choices (Norcross, 2020).

Neil Sinhababu proposes a version of scalar consequentialism that is meant to be slightly less revisionistic, which also does away with hard deontic categories like permissibility and obligation, but retains a notion of right and wrong action, which is reformed to take a self-admittedly linguistically awkward, but supposedly conceptually clear scalar use (Sinhababu, 2018). This is meant to both mitigate the costs of the view, and better account for its capacity to guide action as opposed to only assessing it. In fact, Sinhababu argues that, insofar as they contain not less, but more information than discrete usages of "right" and "wrong", scalar ones are better guides to action. His example concerns an agent asking advice from a friend on how to get to the airport, which is provided as the following qualified ranking of options, each "more right" than the next: the subway is best, a cab is fine, walking is bad, and stealing a car is terrible. This is said to be more detailed than lumping differently valuable alternatives into one same degree of compulsion, and so it is not lacking in any information that discrete guiding may have contained.

I suspect that this supposed gain in information cuts both ways. Take Sinhababu's own airport example on action-guiding scalar rightness. In normal circumstances, it is okay for the agent to either take the subway, a cab or walk, even while granting that each results in less and less wellbeing, which may in turn easily be granted morally relevant value. However, there does seem to be a strong sense in which it is not morally okay for the agent to steal a car. Intuitions about why this is so seem to stem, perhaps among others, from the difference in undermining one's own welfare as opposed to others', as well as the existence of comparatively large cardinal jumps in the degree of aggregate value involved in each alternative, but this is largely beside the point on the meaningfulness of deontic categories. It seems we may have actually lost information on the strength with which certain options are to be morally discouraged, and this problem is, of course, no smaller for Norcross's original view.

We seem to be back to circling around trade-offs involved in asking incompatible things from the same theory, but we are now in a better position to assess how different decision criteria fare in responding to the overdemandingness objection. As we saw earlier, maximising consequentialism fails to meet it because it could not possibly be more demanding. This means that, if it is possible to be morally overdemanding, then maximising consequentialism is, and the possibility of moral overdemandingness is strongly supported by overwhelmingly plausible intuitions that, unlike optimific actions, morally permissible actions are fairly ordinary, as are supererogatory ones. It seems that avoiding these undesirable implications requires leaving space for suboptimal but permissible options, which is exactly what satisficing consequentialism is designed to allow for. But, as we saw, maximising and scalar consequentialists will rightly complain that satisficing consequentialists are both rather slippery about how and where the crucial distinction between permissible and impermissible action should be drawn, and also normatively indifferent about the kinds of differences in value that a consequentialist ought to consider relevant: the superiority of the supererogatory optimum over the permissible minimum seems to be merely reported in evaluation, but it does not translate into any kind of requirement on action. As for scalar consequentialism, Sinhababu's airport example seems, in fact, strongly analogous to the previous example with snowstorms, soup and moussaka, and it therefore likewise calls for room for supererogation. However, of course, supererogation is not captured by scalar consequentialism, because it is a derivative concept with respect to obligation: one cannot

go above and beyond the call of duty if there are no duties. The view therefore does not quite solve the problem of overdemandingness, instead stipulating it out of the question by denying any degree, appropriate or otherwise, of demandingness.

2. What is the problem with vagueness?

2.1. *What is vagueness?*

In this section, I will present a brief state of the art concerning vagueness in the philosophy of language and logic. If the explanation is clear, it may be possible to anticipate why I believe this bears heavily on the issues discussed in the previous section, although more specific consequences will be spelled out in the following section.

It will come as news to no one that language is sometimes vague, but it is not uncommon to have a less than precise notion of what we mean by “vagueness”. In the philosophy of language and logic, the term “vagueness” is reserved for a more or less specific problem which selects only one of the many different semantic problems generally meant by the term in its everyday usage. In this somewhat narrow sense relevant for our purposes, vagueness can be understood as the very pervasive feature of concepts whereby they clearly apply to some cases and clearly fail to apply to others, but there is no sharp boundary separating the two. Although research on vagueness has exploded since the 1970s and is currently a lively field of discussion in the philosophy of language and logic, there is no precise consensus definition of the notion, and it is often instead characterised in terms of paradigmatic examples, and discussed in terms of certain key features, generally without major costs to mutual understandability. I will rely on the common example of human tallness and the following three closely interrelated (apparent) key features of vague predicates: they present borderline cases, they present tolerance and they are susceptible to sorites paradoxes (Wright, 1975; Keefe, 2000; Hyde and Raffman, 2018; Oms and Zardini, 2019; Sorensen, 2022)⁷.

It is clear that, at a rough 211cm, Catalan basketball player Marc Gasol is tall, whereas former Catalan president Jordi Pujol, at a rough 165cm, is not tall. It is unclear whether British monarch Charles III, at a rough 178cm, is tall, and it does not seem that data about his tallness (or that of others, in fact), would provide much clarification, and so he

⁷ Boundarylessness is sometimes treated as a separate feature from tolerance (Dietz and Moruzzi, 2010), generally in order to distinguish commitments about the existence of specific boundaries from commitments about the possibility of indefinitely spreading vague predication, and the like. I will omit this because I do not believe it has much of a specific bearing on the following discussion.

constitutes a borderline case of tallness. Tolerance expresses the idea that there are differences in human height too tiny to merit changing whether we should predicate tallness, so that if somebody is tall, so is somebody who is only minimally taller, and vice versa for the negation of this predicate. Let us see this in the context of the sorites paradox.

The principle of bivalence holds that all propositions are either true or false, which means that for any individual X_i with height i cm, “ X_i is tall” will be either true or false. As mentioned already, ordinary common-sensical intuitions about tallness, Gasol and Pujol establish (rather plausibly) that “ X_{211} is tall” is true and “ X_{165} is tall” is false. Tallness cannot decrease with increases in height, so “ X_i is tall” will be true for all X_i such that $i \geq 211$, and false for all X_i such that $i \leq 165$. This furthermore implies monotonicity, and since the truth-value of the predication is different between the cases of X_{211} and X_{165} , it cannot have changed more than once, but it must have changed at least once. This would entail that there is some real number i^* such that $165 \leq i^* < 211$, and “ X_{i^*} is tall” is false, whereas “ X_i is tall” is true for all $i > i^*$. But this is exactly what the aforementioned statement of tolerance denies.

The following observation might illustrate the profound implausibility of denying that tallness presents the feature of tolerance. A friend who is a trusted physicist tells me that atomic vibration in one’s body causes one’s height to oscillate on a scale no larger than ångströms (10^{-10} m) at a frequency rate on the order of 10^{12} to 10^{15} oscillations per second. If my friend is right, then, if some individual X_{i^*} ’s average height through this cycle of atomic vibration is i^* , X_{i^*} will have gone from being tall to not being tall some 10^{12} to 10^{15} times in a given second. That this could happen seems clearly bizarre and false, and so the paradox is served. It will be useful to consider it in the following more succinct schematic form:

E1. A 1650mm tall person is not tall.

T. If an x mm tall person is not tall, neither is an $x+1$ mm tall one.

E2. A 2110mm tall person is not tall.

Succinctly put, the major philosophical problem here is that this seems to be a valid argument, its premises seem to be true, and its conclusion seems to be false, which is exactly what by definition cannot occur in a valid argument. Call (E) (for “everyday use” or “evident”, as preferred) the claim “ $E_1 \ \& \ \neg E_2$ ”, i.e., the conjunction of (E1) with the negation of (E2), which was our previous predication of tallness to Gasol but not to Pujol. (T) is for “tolerance”. If we want to avoid the paradox, we can either deny (E), deny (T), or deny the validity of the argument. Neither strategy looks particularly promising, and the challenge is even more daunting when one considers the fact that we are interested in vagueness in general rather than tallness in particular, which means that we would ideally need an account that works for the entire array of cases of vague concepts, be it with the same strategy, or with some appropriately described combination of them (Keefe, 2000).

Put more generally for any vague term, the anomaly consists in a tension between three fundamental philosophical claims. One concerns the robust applicability of classical logic, and more specifically the principle of bivalence, i.e., the claim that all propositions are either true or false (neither both nor neither). The other two have to do with strong intuitions about vague concepts: that some attributions of vague properties are true and false (E), and that some changes are too small to merit a change in the attribution of vague properties (T). Before discussing the different strategies that challenge some of these assumptions, however, it might be worthwhile to comment very briefly on the philosophical significance of this problem in general, and in particular for practical philosophy.

At first glance, this problem might seem rather trivial and far-fetched, and most would agree that there is a level at which it is, since there are indeed more pressing issues than the classification of Catalan celebrities as tall or not. But little philosophical reflection is needed to see that it poses a significant and pervasive problem, as it has central implications for our capacity to assess inferences involving vague notions, or just understanding their usage. This is made all the more inescapable when one notices how vague notions are not only widespread and centrally positioned in both theory and practice, but difficult to differentiate from supposedly well-behaved non-vague notions (Kennedy, 2007).

As well as tallness, common examples of vague concepts in the literature are “heap of sand”, “bald”, “tadpole”, and “green”. The following is just a cursory glance of normatively important concepts displaying the logical and semantic anomalies associated with vagueness: sufficient evidence for conviction, progressive/conservative, populist, major enough scandal to require the resignation of the PM, preference of cream over chocolate pastry or time over money, gain in economic output large enough to be worth a small loss in equality, human being on the zygote-baby continuum, generous, child, and friend. When the attribution of these concepts is up for debate in their natural context, they tend to stir a lot of heated discussion, which can become sterile if the specifically puzzling features of vague concepts are not taken into account (arguments about whether a 13-year-old can ride a rollercoaster that is not meant for children can easily spiral).

Some such issues of practical importance have been analysed with explicit reference to contemporary philosophical debates on vagueness, although this connection is examined in varying degrees of detail (Mcbrayer, 2007; Endicott, 2011; Constantinescu, 2014; Dougherty, 2014; Asgeirsson, 2019; Nefsky, 2019). In particular, it is hopefully already plain to see that this is the problem affecting the debate presented in the previous section, on overdemanding obligation and the relationship between scalar value and discrete deontic operators in consequentialism, and yet to the best of my knowledge, this particular connection has not been made explicitly. In the following subsection, I will briefly summarise the main contenders in contemporary philosophical discussion on the adequate handling of vagueness and the solving of the paradox, in order to spell out its implications for normative ethics in the following section.

2.2. What are my options?

In this section, I will present a succinct overview of current debates between the most plausible strategies in solving the problems posed by vagueness in general and the sorites paradox in particular. These are generally classified in terms of whether and to what extent they challenge classical logic and semantics.

One way to retain classical logic and semantics would be to adopt a nihilist strategy, i.e., to deny the truth or falsehood of statements involving vague terms, based on the notion

that their lack of defined extensions makes them semantically defective to the point that they cannot express propositions. Put in terms of the paradox scheme set out previously, this would allow us to take the argument scheme to be valid, while avoiding commitment to the problematic claim E2 that a 210cm tall person is not tall, not only because E1's failure to be a true premiss makes the argument unsound, but because E2 itself is not taken to be true or false at all, but rather some kind of pseudo-proposition like "All pfeffs are pfeffs". The view is currently very unpopular not only because of its very directly counterintuitive claims (it seems bizarre to deny that 140cm tall people are short, that 210cm tall people are tall and the like), but also because, in the light of previous comments about the large spread of vagueness within language and the blurriness of distinctions between vague and non-vague terms, adopting this account would render unacceptably large swathes of both everyday and technical language semantically null and void. In this light, one can trace the rejection of such an approach to widespread philosophical attitudes against ideal language philosophy in the latter half of the 20th century, and indeed it would constitute a very clearly problematic example of a semantic theory so restrictive that it rejects rather than explain the features of our actual use of language.

Rather than nihilism, epistemicism is the main strategy within classical logic and semantics, and it involves the straightforward denial of the truth of the statement of tolerance (T in the scheme), by commitment to the claim that there actually is a threshold for the applicability of vague concepts as a matter of fact. This preserves both the truth of (E) and the normal usage of classical logic insofar as the argument can be considered valid but simply unsound, given how it has a false premise. The label "epistemicism" is due to the way the view deals with the pressing problem of explaining why the statement of tolerance appears to be true, and why the truth-value of attributions of vague properties to borderline cases seems indeterminate. In general, epistemicists claim that, although applicability thresholds are real, they are somehow deeply unknowable (Williamson and Simons, 1992; Keefe, 2000).

Among the main difficulties of such an account is, perhaps, that it is difficult to produce a well-articulated, understandable, non-*ad hoc* and plausible account of the supposed unknowability involved. Questions about the threshold of vague terms are inquiry-resistant in a way that is difficult to find elsewhere. It is certainly difficult to know whether Socrates burped an odd or an even number of times, to the point that it may turn

out to be unknowable for us now, and yet one can imagine exactly the kind of data one would need to settle an inquiry into the issue, which would be perfectly achievable if there were no material impediments to perfect information. If we knew every single fact knowable about the height of all possible humans, we would plausibly find ourselves knowing a lot more about the phenomenon of human tallness for sure, but in the exact same situation concerning where the supposed threshold for attribution is.

Most explanations of the deep unknowability of thresholds for vagueness rely heavily on some account of how we use a semantics for vague terms that is sharply defined for everyone but impossible to know. The idea seems to be that given (1) all possible data about tallness, (2) all possible data about the uses of the discrete predicate “tall” and (3) a correct semantic theory about how the uses of this predicate determine its meaning, then we might be in a position to know where the specific threshold lies for each vague term (Williamson, 1994; Graff, 2000; Sorensen, 2022).

Perhaps the main problem with this move is that this seems to require making a very *ad hoc* requirement on an as of yet unknown theory of semantics, and perhaps a very implausible *ad hoc* claim at that, since, to the extent that such an ideal situation is imaginable, it would still seem well within the realm of possibility that a specific threshold may still be lacking under perfect semantics. This suspicion is backed by the impression that it seems plausible that the term “tall” is often (and perhaps always) used in such a way that its meaning never involves the kind of hyper-specific threshold within the possibilities of human height needed for the theory to work. This is only stressed by the fact that height is a continuous property, since it is unclear that thresholds at specific points would even make conceptual sense in these kinds of cases. It seems more natural to claim that, when we say that Marc Gasol is tall and Jordi Pujol is not, we know well enough what we mean, but we just don’t mean it in a way that involves any very particular threshold.

The following observation might seem a reason to favour some kind of realism about specific thresholds: even if all cut-off points seem to lack any saliency over the rest that could lead us to suspect they are *it*, some cut-off points are actually a lot more plausible than others. For instance, 177cm seems a lot more plausible a threshold than 166cm or 210cm, so maybe there is some maximally plausible threshold value. But I believe this is

just tolerance with a vengeance: if we start to consider cases from 165cm upward, then our optimism will begin to grow... in a Sorites series: we will end up reaching points too high for plausibility, without any previous point having been a specifically good point to stop. The initial suspicion seemed to be that there is indeed some straw that breaks the camel's back, but it turns out one may be looking for a needle in a haystack where one has no reason to think there is a needle. The underlying problem is that any proposed cut-off does not only seem arbitrary, but just plain false.

The remaining strategies are those that challenge, if not classical logic directly, at least the way it is usually applied. Despite the ease with which the tolerance principle can be put in conditional form and the paradox derived by *modus ponens*, it has been noted since the early days of the contemporary resurgence of the debate that the sorites paradox does not rely on any very specific logical formulations, connectives or complex rules of inference, and so those look like ill-fated places to go looking for reform (Dummett, 1975; Keefe, 2000). This is not to say that the reforms that are due will not affect connectives or rules and, in fact, they are likely to, given how they tend to require substantial revisions of the formalisation of truth, and since connectives and rules of inference are defined truth-functionally, they will generally need to be redefined in some way.

Truth as a property of propositions seems to behave strangely in cases of vagueness. Attributions of vague properties to borderline cases may seem to be both true and false (truth glut), or neither true nor false (truth gap). The main threat here is, respectively, to the principle of non-contradiction (which states that no proposition is both true and false), and to the aforementioned principle of bivalence (which states that no proposition is neither true nor false).

Subvaluationism is, roughly, the strategy that focuses on the former of the two. The idea would be that, whereas "Marc Gasol is tall" is true and "Jordi Pujol is tall" is false, "Charles III is tall" is both true and false. This strategy is related to the dialetheist position in the more general philosophy of logic, which claims that some contradictions are true, while avoiding a slide into trivialism by means of the principle of explosion, whereby anything follows from the truth of a contradiction. Formal problems with this avoidance of explosion aside, affirming contradictions is for many authors the archetype of philosophical unacceptability, which explains the view's widespread rejection.

Supervaluationism is a model of the truth gap explanation, and it is generally taken to be a much less revisionary approach to classical logic. The theory relies on a notion of “precisification”, which embodies the notion of a possible reasonable sharp threshold for the applicability of a vague term. Within each precisification, all propositions are either true or false in the standard sense, but since some attributions of vague terms are true or false for all precisifications (the ones stated in claim (E), for instance), we can take them to be “super-true” or “super-false”. Super-truth and super-falsehood model truth and falsehood *simpliciter* for claims like “Marc Gasol is tall” and “Jordi Pujol is tall”, respectively, and since borderline cases like “Charles III is tall” are neither super-true nor super-false (i.e., they are true in some precisifications and false in some others), they are neither true nor false in the standard sense. This accounts for the appearance that tolerance is true, since there is indeed no pair of contiguous cases such that the attribution of the vague predicate is true (super-true) for one case but false (super-false) for the other, but this does not entail that the predication must therefore be true for both (Keefe, 2000; Cobreros and Tranchini, 2019).

In degree theory, the key thought is that rather than only having 0 or 1 as potential true values for true and false, propositions can take a real number in the interval $[0,1]$, representing degrees of truth. The idea is that some propositions are truer than others, which actually has a lot going for it in terms of ordinary language uses of the notion, although it is also easy to confuse with other notions when formulated with precision. The easiest confusion is probably with claims about probability or claims about semantic disambiguation, which are to be distinguished from the kind of degree of truth involved in modelling vague properties with degree theory. The truth of a proposition attributing a vague property to a case is supposed to go from zero in clearly false cases, then increase as the property intensifies, and reach 1 when it is clearly true. For instance, the claim that a given square is black is truer for a dark grey square than for a white one. If we imagine a white square sliding through grey until reaching pitch black, then we should imagine the degree of truth of the statement “the square is black” gradually increasing from 0 to 1 in the process (Paoli, 2019).

Roughly sketched, the account of how degree theory helps in solving the sorites paradox is the following. A conditional statement is true to the extent that the consequent is no falsier than the antecedent, and likewise valid arguments are those guaranteeing that the

conclusion is no falsier than the premises (truth-preservation). In the paradox scheme described earlier, we begin by giving (E1) full truth, with the statement of tolerance (T) then taken to be *almost* true, which allows for the assertion of the first iteration possible (concerning whether two grains constitute a heap) still be mostly false, but perhaps a *tiny* bit true (marginally truer than E1, at least). With each iteration of this, we chip away at the truth of the statement, and by the point we reach E2, we have lost all truth and obtain the desired preservation of E while producing a plausible (if difficult to understand) assessment of (T) and our willingness to affirm it.

In sum, these are mostly highly revisionary approaches, and indeed willingness to consider revisions of classical logic and the way it ought to be applied, while polemical, is a common feature of this field of discussion. As theories, they often struggle to make sense of how they can harmoniously coexist with instances in which classical logic seems to behave well, and sometimes their costs in intuition are clearer than their gains in explanatory power. Additionally, large parts of the discussion on assessing these theories are closely linked with the extent to which changes in logic are deemed philosophically desirable for reasons unrelated to vagueness, as well as on attitudes toward logical pluralism.

3. One problem, several strategies, any solutions?

3.1. *Two problems or one?*

The previous section may have at times seemed like it came out of another text entirely, but if it was clearly explained, then my point is hopefully already reasonably apparent: the distribution of discrete deontic operators based on scalar notions of value is vague in the exact same way that the attribution of tallness based on scalar height is. In this section, I will discuss the interplay between different theories of vagueness and consequentialism. Before delving into it with any detail, however, I believe we are now in a position to put the issue a bit more clearly.

In the presentation of consequentialism in section 1, I mostly discussed how consequentialists, having arranged choice options by value (with any two being comparable in terms of “better than” and “worse than”), need to provide a decision criterion, i.e., an account of how that evaluation translates into normative commitments about what ought to be done in practice. I also mentioned that a common and philosophically important way is by means of deontic operators, by which I meant the conceptual pairs right and wrong, permissibility and impermissibility, and the like. We can now formulate a specific version of the Sorites paradox for a range of options along a value continuum:

E1*. Bringing about the best outcome is right.

T*. If bringing about an outcome of value x is right, then so is one of value $x-1$.

E2*. Bringing about the worst outcome is right.

This formulation in terms of rightness is, of course, just one of many possible formulations, and the same scheme can be used analogously for permissibility, or indeed impermissibility and wrongness (by adding rather than subtracting value in the corresponding tolerance statement T*). In fact, the same kind of sorites series can be

formulated for the attribution of goodness simpliciter to states of affairs (and, again, the exact same thing holds for the reverse version for badness):

E1**. The best state of affairs conceivable would be good.

T**. If a state of affairs with value x is good, then so is one with value $x-1$.

E2**. The worst state of affairs conceivable is good.

This seems to have important implications for many debates in ethics, in particular on axiological debates concerning the definition and meaningfulness of a notion of 0 value as a threshold separating positive or proper value from disvalue, good from bad, which seems to be a vague distinction in the sense discussed here. More specifically yet, this probably goes some way in explaining the puzzling features of the debate on the so-called Repugnant Conclusion, concerning whether it would always be better to have some enormous amount of only minimally pleased people rather than any more reduced (but still decently sized) group of happy people (Arrhenius, Ryberg and Tännsjö, 2022), since the notion plays on a long addition of small amounts of positive value, which would suddenly turn to a long subtraction if this paradoxically vague threshold were crossed.

Although these issues on the evaluation of states of affairs are not the focus of the present discussion, one tangent does seem relevant for our purposes, namely the relationship between evaluations of states of affairs, evaluations of choice alternatives and action-guiding normative commitments such as deontic statements. It seems that the first two are to be closely identified within any consequentialist theory of ethics, but as discussed already the relationship between the second and third is a bit thornier.

Consequentialism is particularly welcoming of the notion that it might be a good thing to do to bring about a bad outcome, if indeed no available outcome is good enough to be worthy of being labelled “good”. The classical example of the trolley illustrates just this, as most consequentialists would readily grant that even if five people dying is worse than one person dying, neither is good and indeed both are bad states of affairs. Indeed, I believe it is common for consequentialists to fault deontology precisely because of a certain kind wishful thinking concerning whether a morally scrupulous option must

always be at hand, with inconvenient examples presenting disproving exceptions for the entirely reasonable yet too broadly stated prohibitions involved in the formulation of specific duties, such as when we forbid stealing, killing, lying and the like. In sum, the relevant idea here is that consequentialists base their evaluations of choice alternatives on the values of the states of affairs, but bad states of affairs need not be bad to bring about, and conversely good states of affairs need not be good to bring about, as the evaluation of choice alternatives needs to take into account which other outcomes are available.

Now, it is less clear in which sense there is any such wiggle room between the evaluation of choice alternatives and normative statements concerning choice. Whereas, as just explained, it is natural enough to claim within a consequentialist framework that a bad outcome might be good to bring about, I think it would be rather puzzling and bizarre to say that some option A is a good option to choose and yet it would be wrong to choose it, and the same goes for claiming that choosing B would be a bad choice and yet it is right. And yet, as discussed in the first section of this paper, critics of scalar consequentialism may contend that when wondering what I should do I am somehow concerned with whether my options are right or wrong, and not just whether it would be good or bad for me to choose each of them, and more specifically *how* good or bad. At any rate, if we accept rightness as defined on the basis of the goodness of choices, then it is common and fairly reasonable to simply equate rightness with permissibility. Although I have my doubts that this identification is warranted (on which more in the following subsection), if this relationship between rightness and permissibility is granted or otherwise explained, then the rest of deontic commitments seem to follow as a matter of fairly simple deontic logic (McNamara and Van De Putte, 2022).

In sum, then, two different yet related distinctions of crucial importance for consequentialists seem to be vague in the specific sense explained in the previous section. One concerns axiology and the evaluations of states of affairs, as the distinction between good and bad states of affairs (and therefore good and bad outcomes) is vague. The other concerns the distinction between deontic categories for choices, namely whether they are right or wrong, permissible or impermissible choices. My claim is that the vagueness of this latter distinction is behind the paradoxical and puzzling nature of the debate on overdemandingness, and so the problem of how well we are morally obliged to act

(described in section 1) is, perhaps surprisingly, in essence the same problem as the one about how many hairs make or break being a bald person (described in section 2).

In the remainder of this section, I will attempt to succinctly discuss what, to my mind, would be the most fruitful interpretations of the vagueness involved in this context, and their implications in terms of moral philosophy. First, I will discuss epistemicism within classical logic and semantics, focusing specifically on its fit with satisficing consequentialism. I will then focus on the somewhat more logically revisionistic strategies of supervaluationism and degree theory, and especially how they relate to satisficing and scalar consequentialism, as well as their broader implications for the debate on overdemandingness in general.

3.2. What if there is a specific threshold?

Epistemicism seems largely incompatible with maximising consequentialism. Both views involve a commitment to the claim that the threshold is somewhere specific, but whereas maximising consequentialism entails a commitment to the claim that it is precisely the maximum, epistemicists are committed to the view that the threshold is deeply unknowable in some way, so it would be strange to claim that we do not know where the threshold for permissibility is, but that it is at the maximum.

There is one rather convoluted sense in which one may be tempted to draw a connection between the two views. Slote's original defence of satisficing consequentialism was meant to address two rather different kinds of problems faced by maximising consequentialism (Slote and Pettit, 1984). One is the notion that, in many scenarios (such as the one illustrated by the snowstorm, soup and moussaka example in subsection 1.2), the maximum is clearly available and supererogatory, i.e., not also the minimum permissible. This is the kind of case on which we have been focusing, and the one in which the maximising consequentialist claim seems directly incompatible with the epistemicist claim. Still, the other type of case concerns situations in which the optimistic choice is somehow undefined or indeterminate. Slote envisions a hero being offered a wish by a genie, so as to suggest such an openness of choice that, for any option one considers, a better an improvement could always be thought of. In such a scenario, one

could claim that we are at a loss as to where the maximum is in the same way that we are at a loss as to where the threshold for permissibility is, so they might therefore be the same. It is hopefully clear to see, however, that on the one hand this is not a tight argument, as the two mysterious points (for a permissible option and for the maximum) need not coincide, and on the other this claim would go exactly against the grain of Slote's rather plausible point, that the inexistence or unidentifiability of a maximum point does not preclude the existence of a permissibility threshold somewhere reasonable below it.

Epistemicism is in fact a lot more at home with satisficing consequentialism, and indeed both views seem to share a way of modelling the slipperiness of claims about where exactly something is which however *must be somewhere around that region*. One can generally treat this as a good model of some intuitions about moral behaviour, allowing us to express the thought that it is difficult to know whether certain morally iffy actions are really permissible, and providing arbitrary range-like estimations of thresholds that trade off length of interval against plausibility (Hawthorne, 2022). In the same way that we can be confident that the threshold for tallness is somewhere between 166 and 211cm for tallness, the threshold for permissibility is somewhere between walking and stealing a car in Sinhababu's earlier airport example, and between leaving a friend to freeze outside and letting them in for gastronomically unremarkable soup in the snowstorm example.

A thornier issue is how all questions of moral permissibility seem to be therefore modelled under epistemicist satisficing consequentialism as problems of risk. In a continuum of ranked alternatives, each choice will be either permissible or impermissible, and some will be riskier than others, although we will never know which were actually one or the other for borderline cases, even after the fact. I suppose this is in some ways plausible and in other ways implausible, given how something like instrumental rationality does seem to play a role when discussing the permissibility of iffy actions. A thought like "I don't want to risk doing the wrong thing" seems like a natural enough way of reasoning morally sometimes, but it is not clear that it would be the same thought as "I wonder whether acting like this would *really* be wrong", and this latter kind of worry seems to be the properly central bone of contention here.

It might help to compare the behaviour of two hypothetical moral agents facing the same sorts of circumstances. Petra, let us say, is always selflessly striving to do what is best, whereas Martha systematically acts in the kinds of ways that satisficing consequentialists would consider permissible but not supererogatory. If the combination of epistemicism and satisficing is right, then what we are justified in saying is that Martha's behaviour is more morally risky than that of Petra, but not that it is necessarily any worse from a deontic point of view (even if the outcome of her behaviour is comparatively lower in value). Whether Martha does in fact behave impermissibly will turn out to be a matter of luck based on the unknowable matter of fact concerning the actual threshold for permissibility.

Perhaps a more dangerous related philosophical liability of satisficing epistemicist approaches involves the implausibility of affirming that unknown (and unknowable) obligations can be morally binding. In a somewhat simplified way, we may argue that since I do not know exactly what my obligation is, there is some sense in which I cannot fulfil it and, therefore, because ought implies can, it is false that I was obliged to do that by modus tollens. But at a more general philosophical level it seems odd that, despite having full information about an action's consequences and their value, we would be certain that it is either strictly morally forbidden or perfectly permissible even, but completely at a loss about which one it is. And yet, this is what would follow from applying epistemicism to the issue of value and consequentialism.

Still, the combination of epistemicism and satisficing consequentialism clearly illustrates the aforementioned issue of disproportionate deontic significance for tiny changes in value, which here takes the form of the denial of tolerance for permissibility (T* in the permissibility-specific sorites scheme formulated in the previous subsection). In the same way that it seems deeply implausible that any single ångström of height could make the difference between a person's being tall or not, it seems deeply implausible that any infinitesimally small difference in value could mark the difference between permissible and impermissible conduct. This is, after all, what was illustrated by the example on moussaka and soup, which the satisficing consequentialist took to be a defeating counterexample for maximising consequentialist. The same problem afflicts the epistemicist satisficing consequentialist, even if she by definition does not know exactly where it does. In claiming that there is some suboptimal specific degree of value marking

the sharp threshold for moral permissibility but that it cannot be known, she is not avoiding commitment to some such problematic counterexample, but rather committing herself to one in some unknowable range of value below the top.

3.3. What if there is no specific threshold?

Subvaluationism involves claiming that both the attribution and the negation of vague concepts is true for borderline cases. Counterintuitive as it is to affirm something like someone's simultaneous tallness and non-tallness, this only seems to get worse for deontic statements. If some course of action is and is not permissible at the same time in the same sense, then it is just completely unclear if one may or should go for it, and so it seems to be the opposite of action-guiding to be given the kind of contradictory order that would follow from such a deontic statement.

Supervaluationism, on the other hand, seems like a substantially better fit, and indeed it might even be taken as a nice model of how the existence of different reasonable moral standards may in some sense meaningfully inform our moral agency. The story might go something like this: one reasonable precisification of "permissible" might draw the line at value x , whereas another might do so at value $x + 1$, and so on. This is not to say that all possible precisifications are reasonable: clearly the distinction is not between (A) spending your entire life maximising your capacity to alleviate the suffering of others, and (B) doing exactly the same but once foregoing three seconds of work to dry a drop of sweat from your forehead when it was not strictly necessary, nor is it between (C) gratuitously murdering one person and (D) gratuitously murdering two people. But at some point with value between B and C in the previous list some ways of cashing out the distinction will begin to be reasonable. We might then say that x is impermissible if, and only if, it is not permissible under any reasonable precisification, and vice versa for permissibility.

Such a view seems to have a nice pluralistic ring to it, but unfortunately it involves baffling commitments, and its viability is dubious. If we simply apply the supervaluationist framework for each deontic concept like permissibility and the like in the way just described, then what we obtain is a model for which a truth gap obtains in

the middle for all cases involving some disagreement between precisifications. In this case, this involves the idea that choice alternatives which could reasonably be considered either permissible or impermissible are *neither*, and this is a very odd claim indeed.

When someone says that it is neither true nor false that Charles III is tall, it is reasonably clear why they are saying that, even if it is controversial, and under a charitable reading it does seem *prima facie* compatible with Charles's being a human being with height. The idea is that there is something not too different from some form of category mistake here, like we would incur if we claimed that Seneca was unemployed, or if we claimed that *The Last Emperor* is a beige film: the concept does not really work well with this case, although it does with closely related ones, and its use is related to features present in the problematic case as well. This kind of story works a lot worse for permissibility, as it seems bizarre to say that some available options for an agent are neither permissible nor impermissible for that agent to choose, and it is certainly not an action-guiding statement concerning their moral status as choices. In principle, all available choices are either permissible or not permissible (i.e., impermissible), and that ought to be the end of the story. Category mistakes involve predicating permissibility of things that are not any moral agent's choice, such as would be committed by anyone saying strange things like "Charles III is an impermissible man", or "It is impermissible that stegosauruses went extinct several million years ago". But the issue with available choices is nothing like this, and as was mentioned earlier, if a course of action is available to me and I wonder whether I should take it, it is hardly informative to say that I neither may nor may not take it from the point of view of morality.

Still, it seems very reasonable to undermine the conventional identification between rightness and permissibility and say that, whereas a truth-gap concerning rightness seems *prima facie* reasonable, a truth-gap concerning permissibility seems a lot less so: some perfectly ordinary actions (perhaps most of them) seem entirely permissible and neither right nor wrong from the point of view of morality. This could be modelled via a range of reasonable precisifications for the pair right/wrong, with some actions being wrong by any reasonable account (and therefore wrong simpliciter) and therefore impermissible to carry out, and then one could claim that actions which are right by some reasonable account are permissible, with actions that are right by all reasonable accounts also being supererogatory. This kind of distinction between rightness and permissibility might seem

to allow us to say that what someone is doing might be in some sense wrong, but maybe not impermissibly wrong, such as, for instance, not ceding one's not-especially-reserved seat to an older person on the bus.

Still, this would really seem to be more of a play on words rather than a satisfactory supervenient account of the vagueness of "right", "permissible", "supererogatory" and their negations and complements. If supervenientism is true and rightness is vague, then actions which are right by some reasonable standards but wrong by some others are neither right nor wrong, and so no action could really be permissibly wrong or anything like that: since truth *simpliciter* is modelled as supertruth for both attributions and negations of vague concepts we will simply again find a strange truth gap concerning deontic statements about borderline cases, undermining moral guidance concerning the available choices involved.

In relation to the fit of such supervenientist approaches with maximising and satisficing consequentialism, the discussion would boil down to a disagreement about whether a threshold for rightness at the optimum level would constitute a reasonable precisification or not, and indeed a lot would hang on where we draw the line for reasonable interpretations of concepts. One reason to include the optimal threshold among these reasonable precisifications might be the fact that so many people claim to take the optimum as the only reasonable threshold, whereas the counterexamples provided by critics of maximising consequentialism, like the earlier soup and moussaka case, could be counted as reasons against this. Either way, the incorporation of the supervenientist framework seems to leave both satisficing and maximising consequentialists with the same respective problems and one further common problem concerning neither permissible nor impermissible alternatives.

In relation to scalar consequentialism, it is worth noting that at least a large part of the motivation and strategy behind Norcross's outright rejection of deontic categories seems to, in effect, mirror the aforementioned nihilist strategy for the sorites paradox in general. Since sudden jumps in deontic status cannot plausibly be justified on the basis of tiny changes in value but their common usage implies that some such jump does in fact occur, it is the common usage that is flawed, and although the terms might allow for some sort of communication about the rough picture, they do not truly convey the way things are in

the way meaningful true propositions do. Of course, there is a considerable overlap in the liabilities of these positions: scalar consequentialism implausibly denies meaningfulness and truth to *prima facie* obviously true statements like “gratuitously murdering thousands of people is impermissibly wrong” and “it is not obligatory to call everybody Steve”, in the same way that nihilism about other vague concepts would lead us to deny that any true statements can be affirmed about frogs, the colour red, and so on. I think this line of counterargument is overwhelmingly plausible.

A much more interesting connection can be drawn between scalar consequentialism and degree-theoretical approaches to the sorites paradox. In essence, this would amount to ascribing degrees of truth to deontic statements about choice alternatives, in some proportion with their associated value, along the lines suggested by Sinhababu on the “scalarisation” of the concept of rightness, extended perhaps to other deontic concepts. There seems to be some fairly deep congeniality here in the emphasis on the gradual nature of matters of truth and morality, as opposed to an oversimplified black-and-white view. Put in terms of the paradox scheme for rightness earlier, the degree-theoretical scalar consequentialist might hold E1* and deny E2* by denying T* full truth, insofar as, for at least some actions, another action with only slightly less associated value would be only a bit less right, yet still be largely right and not wrong enough to merit being labelled “wrong”. The resulting view would constitute a clarifying formalisation of Sinhababu’s take on a revised gradual notion of rightness that reasonably addresses the sorites paradox problem:.

Indeed, it might be possible to use degree-theoretical notions to avoid some of the potential problems faced by Sinhababu’s scalar rightness theory. In its original formulation, the view was committed to the following claim: the rightest choice is the one with most associated value, the wrongest is the one with least associated value, and the degree of rightness of other options is distributed proportionally according to their relative associated value. But as we saw earlier, it seems that some differences in value between options seem like they ought to matter a lot more than others. Consider a variation on his airport problem in which I must choose one of the following four ways of getting there, by order of associated value: riding my bicycle, driving my car, stealing and riding someone else’s bicycle, or stealing and driving someone else’s car. Sure enough, it’s worse to steal the car than the bicycle, but it seems like there is some salient line of

wrongness between the two middle options, rather than just increments in rightness or wrongness between them.

In incorporating a degree-theoretical account of the vagueness of “wrong” and “right”, we can now claim that wrongness and rightness saturate at full wrongness or rightness somewhere before the minimum and the maximum but do so smoothly (along a continuum), much in the same way that, in the example on tallness, the truth of claims “ X_N is tall” saturates at 1 smoothly but considerably before $N = 210\text{cm}$, so that both “ X_{209} is tall” and X_{211} is tall can be modelled as fully true as well. For the airport problem variation just mentioned, we could say, for instance, that either stealing a bicycle or a car would be fully wrong, even though riding my own bicycle would be fully right, and driving my own car would be less right, but still more right than wrong.

Despite its interesting features, this view mostly shares the limitations of scalar consequentialism, especially with respect to how rightness and wrongness relate to other deontic operators like permissibility and obligation, and whether statements involving them could be made to acquire meaning within the framework. If we construe rightness broadly as meaning permissibility, then saying an option is mostly but not fully right would entail saying that it is mostly but not fully permissible, whereas if the notion is construed narrowly as obligatoriness, it would entail the claim that something is mostly but not fully obligatory.

Although these kinds of statements might seem to have some kind of meaning when used in daily life (something about the intensity with which we would bestow praise or blame, or something along those lines), it is not clear what exactly they would mean, or how they would be action-guiding. Indeed, if somebody tells me I am a little bit obliged to do something, then I could understandably ask them if I may refrain from doing that or not. If they say no, then apparently I was fully and not partially obligated, and if they say yes, then it seems they did not really mean obligation, or at least not in the usual sense. Analogous problems occur for the notion of an action’s being partially permissible and partially impermissible.

When we say that the statement “A 178cm tall person is tall” is in part true and in part false, it may seem mysterious what follows from having this kind of logic and semantics in a philosophical sense, but it is not too difficult to see why somebody would say

something like this, and indeed if we all we are told about someone's height is that neither calling them tall nor not-tall would be completely true or false, we can probably imagine a reasonable (if fuzzy) range. On the other hand, it is a lot more mysterious what one would actually be saying if they say something like that about an action's permissibility since, unlike for the case with height, we would not really know whether we may morally take a certain choice if all we are told is it is a bit permissible and a bit impermissible, since there is no choice we may do a bit and avoid doing another bit. The problem in this regard seems to be common with the aforementioned issue with subvaluationism and claiming that 178cm tall people are somehow both tall and not-tall: whereas maybe something like it works for tallness, we would not really know what to do if someone simultaneously told us "press the button" and "don't press the button" about the same button.

Still, perhaps something like this view could hold, by either commitment to degrees of rightness without degrees of permissibility (perhaps without any kind of permissibility at all as in the original scalar consequentialist view), or perhaps there is even some subsidiary sense of permissibility and obligation that can be put in gradual terms. I suspect that there is something to be said about the notion of degrees of moral compellingness to action, which does seem to be reflected in Norcross's notion of evaluations as moral reasons, and common usage might be a good guide here, rather than philosophically suspect: if someone tells us that something is in part obligatory and in part not, they might mean that I may either do it or not (and so it is not true that I must), and yet in some sense I should, and that is probably the puzzling sense of supererogation discussed earlier in section 1, which involves a different kind of relation with imperatives than the one involved in programs and algorithms.

Conclusion

If my arguments in section 1 were compelling enough, we have hopefully seen that the overdemandingness problem might seem trivial but is in fact a serious threat to the plausibility of consequentialism, and a deeply puzzling philosophical riddle that has not been settled very satisfactorily within moral philosophy. As was mentioned in the introduction, the plausibility of consequentialism is no small matter, insofar as consequentialism is an enormously relevant framework both within and outside of ethics. If my explanation in section 2 was reasonably clear, we may be able to see that the problem is in fact, in a sense, larger than ethics, as it involves deep logical and more broadly philosophical problems about the notion of truth. My main objective in the paper was indeed to articulate this connection and give a glimpse of how I believe it ought to influence the debate in ethics. My attempt has not been, in this way, to “reduce” the ethical debate to a logical or linguistic one by denying that the case of deontic statements is affected by specific philosophical complexities that are due to their moral content, and indeed I have tried precisely to highlight and develop some of these specifically normative issues involved.

At a more specific level, I hope it is clear to see that maximising consequentialism is, by far, the view in ethics that is most clearly undermined if this reframing in terms of vagueness is taken onboard. In a nutshell, the best thing one can do is not the only thing one may morally do in the same way that not only the tallest person that is tall. There is good reason why no theory of vagueness has been attempted along these lines, whereas richer parallels do seem to exist between the main approaches to vagueness and non-maximising views on permissibility that have been endogenously developed in moral philosophy.

With respect to these non-maximising consequentialist views, namely, satisficing and scalar consequentialism, debating the issue in terms of vagueness does not settle debate, but it does not leave it unaffected either. The most congenial associations do seem to be between epistemicism and satisficing consequentialism on the one hand, and between , but I have tried to show that they are not the only possible combinations. Connections between the views are also somewhat ambivalent: as has been discussed to some extent,

there are ways in which one or another logic and semantics might mitigate criticisms of ethical views, but also new problems emerge, or at least new ways of looking at them.

I therefore wish it is clear that I do not take any specific combination of views to have settled either debate, on vagueness or permissibility. I don't suppose it would be philosophically appropriate to look to the philosophy of vagueness merely as an armoury for any specific take on moral philosophy, as I think it is fair to say that some issues about truth and logic somehow underlie rather than follow from commitments we may want to adopt about permissibility or any other ethical topic. For instance, knowledge of how I ought to distribute deontic operators like rightness and wrongness between choice alternatives requires having some more or less articulated notion of whether each one will be either fully right or wrong or whether they can be distributed in degrees, etc., and in reasoning clearly and systematically about and with these claims, I will either think of them as following classical bivalent logic and semantics, or some other system with truth gaps, gluts, degrees, and whatnot. In fact, moral philosophers looking for support rather than discussion in the philosophy of vagueness might easily despair to find themselves having to commit to some of a series of very complex minority views that few if any authors take to be particularly satisfying overall accounts of the phenomenon, even by common philosophical standards of what it is to settle an issue. Still, I believe both fields would benefit greatly from a joint thinking about the issues highlighted in this paper, and I don't see how it wouldn't be a theoretical necessity to some extent, and perhaps a practical to some other.

Bibliography

Alexander, L. and Moore, M. (2021) ‘Deontological Ethics’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Winter 2021. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/win2021/entries/ethics-deontological/>.

Andrić, V. (2021) *From Value to Rightness: Consequentialism, Action-Guidance, and the Perspective-Dependence of Moral Duties*. New York, NY, USA: Routledge.

Arrhenius, G., Ryberg, J. and Tännsjö, T. (2022) ‘The Repugnant Conclusion’, in E.N. Zalta and U. Nodelman (eds) *The Stanford Encyclopedia of Philosophy*. Winter 2022. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/win2022/entries/repugnant-conclusion/>.

Asgeirsson, H. (2019) ‘The Sorites Paradox in Practical Philosophy’, in S. Oms and E. Zardini (eds) *The Sorites Paradox*. Cambridge University Press, pp. 229–245.

Bradley, B. (2006) ‘Against Satisficing Consequentialism’, *Utilitas*, 18(2), pp. 97–108. Available at: <https://doi.org/10.1017/s0953820806001877>.

Brink, D.O. (2005) ‘Some Forms and Limits of Consequentialism’, in D. Copp (ed.) *The Oxford Handbook of Ethical Theory*. Oup Usa.

Bykvist, K. (2009) *Utilitarianism: A Guide for the Perplexed*. Continuum.

Carlson, E. (1995) *Consequentialism Reconsidered*. Dordrecht, Netherland: Springer.

Cobrerros, P. and Tranchini, L. (2019) ‘Supervaluationism, Subvaluationism and the Sorites Paradox’, in E. Zardini and S. Oms (eds) *The Sorites Paradox*. Cambridge University Press, pp. 38–62.

Constantinescu, C. (2014) ‘Moral Vagueness: A Dilemma for Non-Naturalism’, in R. Shafer-Landau (ed.) *Oxford Studies in Metaethics: Volume 9*. Oxford University Press, pp. 152–185.

Crisp, R. (2011) ‘Ethics’, *Routledge Encyclopedia of Philosophy*. Routledge. Available at: 10.4324/9780415249126-L132-2.

Crisp, R. (2021) ‘Well-Being’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Winter 2021. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/win2021/entries/well-being/>.

Dawson, H. (2019) ‘Shame in Early Modern Thought: From Sin to Sociability’, *History of European Ideas*, 45(3), pp. 377–398. Available at: <https://doi.org/10.1080/01916599.2018.1534447>.

Dietz, R. and Moruzzi, S. (2010) *Cuts and clouds: vagueness, its nature, and its logic*. Oxford University Press.

Dougherty, T. (2014) ‘Vague Value’, *Philosophy and Phenomenological Research*, 89(2), pp. 352–372.

Dummett, M. (1975) ‘Wang’s paradox’, *Synthese*, 30(3–4), pp. 201–32.

Endicott, T. (2011) ‘Vagueness and Law’, in G. Ronzitti (ed.) *Vagueness: A Guide*. Springer Verlag, pp. 171–191.

Graff, D. (2000) ‘Shifting Sands: An Interest-Relative Theory of Vagueness’, *Philosophical Topics*, 28(1), pp. 45–81.

Hawthorne, J. (2022) ‘Moral Vagueness and Epistemicism’, *Journal of Ethics and Social Philosophy*, 23(2).

Heyd, D. (2019) ‘Supererogation’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Winter 2019. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/win2019/entries/supererogation/>.

Howard-Snyder, F. and Norcross, A. (1993) ‘A Consequentialist Case for Rejecting the Right’, *Journal of Philosophical Research*, 18, pp. 109–125. Available at: <https://doi.org/10.5840/jpr\emph199317>.

Hsieh, N. and Andersson, H. (2021) ‘Incommensurable Values’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Fall 2021. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/fall2021/entries/value-incommensurable/>.

- Hyde, D. and Raffman, D. (2018) 'Sorites Paradox', in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Summer 2018. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/sum2018/entries/sorites-paradox/>.
- Keefe, R. (2000) *Theories of Vagueness*. New York: Cambridge University Press.
- Kennedy, C. (2007) 'Vagueness and grammar: The semantics of relative and absolute gradable adjectives', *Linguistics and Philosophy*, 30(1), pp. 1–45.
- Mcbrayer, J. (2007) 'Personhood, Vagueness and Abortion', *Australian Journal of Professional and Applied Ethics*, 9(1).
- McConnell, T. (2022) 'Moral Dilemmas', in E.N. Zalta and U. Nodelman (eds) *The Stanford Encyclopedia of Philosophy*. Fall 2022. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/fall2022/entries/moral-dilemmas/>.
- McKay, D. (2021) 'Solving Satisficing Consequentialism', *Philosophia*, 50(1), pp. 149–157. Available at: <https://doi.org/10.1007/s11406-021-00382-y>.
- McNamara, P. and Van De Putte, F. (2022) 'Deontic Logic', in E.N. Zalta and U. Nodelman (eds) *The Stanford Encyclopedia of Philosophy*. Fall 2022. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/fall2022/entries/logic-deontic/>.
- Mill, J.S. (1861) *Utilitarianism*. Edited by R. Crisp. Oxford University Press UK.
- Nefsky, J. (2019) 'Collective Harm and the Inefficacy Problem', *Philosophy Compass*, 14(4), p. e12587. Available at: <https://doi.org/10.1111/phc3.12587>.
- Norcross, A. (1997) 'Good and Bad Actions', *Philosophical Review*, 106(1), pp. 1–34. Available at: <https://doi.org/10.2307/2998340>.
- Norcross, A. (2006a) 'Reasons Without Demands: Rethinking Rightness', in J.L. Dreier (ed.) *Contemporary Debates in Moral Theory*. Blackwell, pp. 38–54.

Norcross, A. (2006b) 'The Scalar Approach to Utilitarianism', in H. West (ed.) *The Blackwell Guide to Mill's Utilitarianism*. Wiley-Blackwell, pp. 217–32.

Norcross, A. (2020) *Morality by Degrees: Reasons Without Demands*. Oxford University Press.

Oms, S. and Zardini, E. (eds) (2019) 'An Introduction to the Sorites Paradox Sergi Oms and Elia Zardini', in *The Sorites Paradox*. Cambridge: Cambridge University Press (Classic Philosophical Arguments), pp. 3–18. Available at: <https://doi.org/10.1017/9781316683064.001>.

Paoli, F. (2019) 'Degree theory and the Sorites paradox', in S. Oms and E. Zardini (eds) *The Sorites Paradox*. Cambridge University Press.

Rogers, J. (2010) 'In Defense of a Version of Satisficing Consequentialism', *Utilitas*, 22(2), pp. 198–221. Available at: <https://doi.org/10.1017/s0953820810000099>.

Scanlon, T. (1998) *What We Owe to Each Other*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Schaler, J.A. (2009) *Peter Singer Under Fire: The Moral Iconoclast Faces His Critics*. Open Court (Under fire series).

Shafer-Landau, R. (2018) *Living Ethics: An Introduction with Readings*. New York: Oxford University Press.

Simon, H.A. (1956) 'Rational choice and the structure of the environment.', *Psychological Review*, 63, pp. 129–138. Available at: <https://doi.org/10.1037/h0042769>.

Singer, P. (2023) 'Ethics', *Encyclopedia Britannica*. Available at: <https://www.britannica.com/topic/ethics-philosophy>.

Sinhababu, N. (2018) 'Scalar Consequentialism the Right Way', *Philosophical Studies*, 175(12), pp. 3131–3144. Available at: <https://doi.org/10.1007/s11098-017-0998-y>.

Sinnott-Armstrong, W. (2022) 'Consequentialism', in E.N. Zalta and U. Nodelman (eds) *The Stanford Encyclopedia of Philosophy*. Winter 2022. Metaphysics Research Lab,

- Stanford University. Available at: <https://plato.stanford.edu/archives/win2022/entries/consequentialism/>.
- Skorupski, J. (ed.) (2010) ‘The Routledge Companion to Ethics’, in *The Routledge Companion to Ethics*. Routledge. Available at: <https://doi.org/10.4324/9780203850701>.
- Slater, J. (2020) ‘Satisficing Consequentialism Still Doesn’t Satisfy’, *Utilitas*, 32(1), pp. 108–117. Available at: <https://doi.org/10.1017/s0953820819000402>.
- Slote, M. and Pettit, P. (1984) ‘Satisficing Consequentialism’, *Aristotelian Society Supplementary Volume*, 58(1), pp. 139–176. Available at: <https://doi.org/10.1093/aristoteliansupp/58.1.139>.
- Sobel, D. (2007) ‘The Impotence of the Demandingness Objection’, *Philosophers’ Imprint*, 7, pp. 1–17.
- Sorensen, R. (2022) ‘Vagueness’, in E.N. Zalta and U. Nodelman (eds) *The Stanford Encyclopedia of Philosophy*. Winter 2022. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/win2022/entries/vagueness/>.
- Tannsjo, T. (2019) *Hedonistic utilitarianism*. Edinburgh University Press.
- Timmons, M. (2001) *Moral Theory: An Introduction*. Lanham, Md.: Rowman & Littlefield Publishers.
- Tucker, C. (2017) ‘How to think about satisficing’, *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 174(6), pp. 1365–1384.
- Wheeler, G. (2020) ‘Bounded Rationality’, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Fall 2020. Metaphysics Research Lab, Stanford University. Available at: <https://plato.stanford.edu/archives/fall2020/entries/bounded-rationality/>.
- Williamson, T. (1994) *Vagueness*. Routledge, pp. 921–928.
- Williamson, T. and Simons, P. (1992) ‘Vagueness and Ignorance’, *Aristotelian Society Supplementary Volume*, 66(1), pp. 145–178. Available at: <https://doi.org/10.1093/aristoteliansupp/66.1.145>.

Wolf, S. (1982) 'Moral Saints', *The Journal of Philosophy*, 79(8), pp. 419–439.

Wright, C. (1975) 'On the Coherence of Vague Predicates', *Synthese*, 30(3/4), pp. 325–365.