



# Is the autonomy of the will a paradoxical idea?

Stefano Berte<sup>1</sup> 

Received: 10 November 2022 / Accepted: 23 February 2023 / Published online: 31 March 2023  
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

## Abstract

This essay tackles head on the argument that sees an inherent paradox in the autonomy of the will as the ground for the authority of the fundamental practical norms. It points out that only on reductive understandings of the autonomy of the will can this idea be qualified as paradoxical, thereby yielding outcomes that either contradict their premises or present autonomy under a false guise. With that done, it will proceed to offer a conception of the autonomy of the will which is not vulnerable to the paradox, and which may therefore be equipped to rest the fundamental practical norms on solid ground. Throughout this discussion, I will rely on constitutivism about practical reasons to specifically defend the twofold conclusion that (a) the paradox of autonomy can be avoided and that, relatedly, (b) if autonomy is properly conceptualised, it is fully equipped and well positioned to ground the authority of the fundamental practical norms.

**Keywords** Autonomy · Will · Constitutivism about practical reason · Agency · Action · Kant

## 1 Introduction

We routinely encounter practical norms that make demands on us. Most of us are expected, for instance, to act as good citizens, sympathetic partners, and loving parents. Each of those standards—good citizenship, sympathetic partnership, and loving parenthood—is associated with certain norms of behaviour. Because we ordinarily act under normative constraints, it is not surprising that philosophers have wondered where the *authority* of those constraints comes from, or why we are bound by them. In response to that authority question, one answer has commanded some consensus. It can be formulated as follows: the authority the *fundamental* practical norms have on us lies in the *autonomy* of our will. That is to say, certain practical demands are

---

✉ Stefano Berte<sup>1</sup>  
bertea.stefano@gmail.com

<sup>1</sup> University of Barcelona, Diagonal 684, 08034 Barcelona, Spain

authoritative because we ourselves confer authority on them. Relatedly, when we abide by the fundamental practical norms we are acting in conformity with demands that *we* have placed on ourselves.<sup>1</sup>

Grounding the authority of some practical constraints in the autonomous will of those to whom those constraints are addressed can well be understood as an appealing move, because it is highly coherent with, or even explicitly dictated by, modes of thinking that are widespread in our societies. These are the modes of thinking that underpin comprehensive worldviews such as the Enlightenment, liberalism, individualism, and modernist humanism, all of which deeply shape large sectors of the societies inhabiting the Western world today. While those modes of thinking and those worldviews are differentiated and hardly reducible to one another, they all share the basic belief that being able to make one's own choices, free from the interference of others, is essential not just to an individual agent's personal fulfilment but also to their sense of self-respect and, even more pointedly, to the recognition of their distinctive moral standing. After all, how can I have respect for myself and regard myself as a subject with my own standing within the moral domain if I am not even in a position to act in conformity with norms I myself come to value and I am instead led, or even forced, to defer my practical decisions to others?

As much as autonomy may seem appealing as a foundation for the authority of practical demands, however, it has hardly gone unchallenged. In fact, this thesis has been criticised on different grounds and sometimes outright rejected.<sup>2</sup> A powerful argument against grounding the authority of practical norms in autonomy (especially, once autonomy is understood in Kantian terms) stresses the inherently paradoxical nature of the idea of autonomy of the will, which is claimed to be either a self-contradictory concept or a deceptive one.<sup>3</sup> The paradox extends to the strategy of conceptualising autonomy as the foundation on which to rest the authority of the basic constraints governing our action, since grounding something in a concept enveloped in paradox can hardly count as providing a foundation of that thing.

This state of affairs makes for a pressing need to provide a justification for the thesis that the authority of the fundamental practical norms resides in autonomy. This essay tackles head on the argument that sees an inherent *paradox in the autonomy of the will as the ground for the authority of the fundamental practical norms*—a paradox hereinafter I will simply refer to as the “paradox of autonomy.” Thus, having first introduced the argument that the autonomy of the will as the foundation of the authority of practical norms gives rise to a paradox, I will proceed to offer a conception of the autonomy of the will which is not vulnerable to the paradox, and which is therefore equipped to rest the fundamental practical norms on solid ground. To that end,

---

<sup>1</sup> This strategy for the foundation of the authority of practical norms has been explored in particular (although not exclusively) within the Kantian tradition of practical philosophy. Influential contributions to this debate can be found in Allison (1990, pp. 85–106 and pp. 230–49), Schneewind (1998, pp. 3–13), Reath (1994; 2006, pp. 121–72; and 2013), Wood (2007, pp. 106–22), Engstrom (2009, pp. 149–83), Korsgaard (1996 and 2009), and Sensen (2013).

<sup>2</sup> In fact, some of the criticisms are premised on a broad and sweeping rejection of autonomy as a cardinal value. Figuring prominently among those who have taken this strong critical stance are Gilligan (1982), Sandel (1982), Taylor (1989), Benjamin (1998, pp. 183–218), Friedman (2005), and Oshana (2005).

<sup>3</sup> The paradoxical quality of autonomy (or, at least, of the Kantian understanding of autonomy) is discussed, for instance, in Pippin (2000, pp. 65–118), Pinkard (2002, pp. 45–65), and Stern (2012, pp. 7–99).

I will rely on a constitutivist meta-normative framework. In the contemporary meta-normative literature, “constitutivism about practical reasons” is the view that, if we are to make sense of practical normativity, we have to see it as conceptually connected with agency.<sup>4</sup> Constitutivism about practical reasons, thus, means that “certain normative claims apply to us merely by virtue of the fact that we are agents,” and that some normative reasons are accordingly grounded “in facts about the nature of agency” (Katsafanas, 2018, p. 367). The constitutivist approach commits us to the claim that we ought to follow the fundamental practical norms because to do otherwise would be tantamount to failing to engage in action proper, thereby ceasing to qualify as agents (or at least as nondefective agents). Conversely, some universal normative demands arise simply from the fact that we are agents. Throughout this discussion, I will rely on a Kantian version of constitutivism about practical reasons that owes much, and yet is not fully reducible, to Christine Korsgaard’s (2009) constitutivist approach. This will enable me to defend the twin conclusion that, if autonomy is properly conceptualised, (a) the paradox of autonomy can be avoided and, relatedly, (b) autonomy is fully equipped to ground the authority of the fundamental practical norms. In order to support this twofold claim, I will outline a notion of the autonomy of the will that, I intend to argue, is located *in* the individual—it is (partly and yet essentially) defined by an agent’s capacity for self-governance and self-rule—and yet, at the same time, transcends the will and behaviour of single agents (that is, it relates to features and properties that lie beyond what this or that particular person can conceivably will to do). Apparently, this conception of autonomy capitalises on the diagnosis of the paradox of autonomy, which is not only taken seriously but also acknowledged to be suitable to set certain requirements that any theoretically sound conception of autonomy needs to account for. Indeed, I take the claim that (some interpretations of) autonomy may give rise to a paradox to be a powerful reminder of the fact that, while autonomy is a capacity that agents exercise through their will, it is also norm-bound, and so is best characterised as a coherent combination of contingent self-determination and universal “lawfulness.” In a nutshell, then, it is in this combination of independent decision-making and subjection to law, I will argue, that we find not only the kernel of autonomy but also the key to avoiding the paradox of autonomy.

This introductory statement is also revealing of the specific *scope* of the argument to be introduced below. In what follows I will articulate a conception of autonomy that is anchored to Korsgaard’s constitutivist account of agency. A construction that takes its cue from Korsgaard’s practical philosophy, I will argue, is insightful for it has the potential to prevent the paradox of autonomy from arising. However, my argument does not have the ambition to specifically defend (Korsgaard’s version of) constitutivism about practical reasons from the criticisms made and concerns raised in the existing literature.<sup>5</sup> In fact, the scope of this contribution is more modest and

<sup>4</sup> Contemporary proponents of constitutivism about practical reasons (in different versions of it) include Gewirth (1978, pp. 21–42), Smith (1994, 2013, and 2015), Korsgaard (1996 and 2009), Schapiro (1999 and 2001), Velleman (2000), Ferrero (2009), Street (2012), Bertera (2013), Katsafanas (2013), and Walden (2012 and 2018). Insightful discussions of constitutivism about practical reasons can also be found in Silverstein (2012 and 2015), O’Hagan (2014), Lindeman (2017), Ferrero (2018 and 2019), Paakkunainen (2018), and Mayr (2019), among other places.

<sup>5</sup> This ambitious task is left to other advocates of constitutivism, some of whom have in fact already undertaken it. Sophisticated arguments in defence of constitutivism (including the variant of constitutivism

limited than that: the only claim that my argument can successfully support is the thesis that the paradox of autonomy does not arise insofar as, in order to conceptualise the autonomy of the will, we bank on (a certain variant of) constitutivism about practical reasons. This conclusion leaves it open the issue as to whether or not constitutivism as a distinct meta-ethical stance (or more specifically, the version of that stance I articulate in this work) is the philosophically soundest approach to practical philosophy that is currently available to us. What I intend to distinctively argue in this work, in other terms, is that, *conditional* on accepting a given (controversial) meta-ethical view and associated conceptions of agency and autonomy, the paradox of autonomy can be avoided. Correspondingly, I will abstain from providing independent backing to such controversial meta-ethical view and corresponding conceptions of agency. Which is why I referred to the project I am embarking here as modest and limited. Despite this qualification, the endeavour made in what follows remains philosophically important, since it promises to give us an understanding of autonomy that, being resistant to the critique of committing us to some paradox, can be safely relied on in order to ground the authority of the fundamental practical norms.<sup>6</sup>

## 2 The paradox of autonomy

The idea of autonomy, central as it may be to certain philosophical traditions, has come under serious scrutiny through a series of criticisms that have challenged its validity in some fundamental ways, especially when taken as the basis on which to ground the authority of practical norms. One of these criticisms consists in pointing out the paradoxical quality of autonomy. The thesis of the “paradox of autonomy,” in its most basic form, can be stated as follows. As the very etymology of the term indicates, autonomy is a Janus-faced notion, for on the one hand it describes what is self-directed—*autos*—while on the other it describes what is norm-governed—*nomos*. Autonomy thus combines two clearly distinct components: (a) independence of the will—an *autonomous* agent is a *self*-governing agent: one in control of their life and thus capable of being their own sovereign—and (b) lawfulness—an *autonomous* agent is someone who acts in accordance with some set of norms, rules, or principles, as against acting according to a normless or particularistic will.<sup>7</sup> This combination, however, is claimed to be unstable and even to give rise to a paradox, once autonomy is taken as the ground for the normative validity of practical standards.

For, on the one hand, insofar as agents are *autonomous*—that is, insofar as they are self-governing and self-legislating creatures—they must be able to set their own norms of conduct. Apparently, those who give themselves their own principles of action can also *change* these principles, thereby ridding themselves of these principles. This means, however, that an autonomous agent is ultimately subject to no norm whatsoever,

---

Footnote 5 continued

I embrace here) have been offered in the literature quoted in footnote 4 above, to which I refer the reader who may be interested in carrying out (or just assessing the value of) the ambitious task.

<sup>6</sup> I should like to thank an anonymous reviewer for this journal for inviting me to specifically clarify the scope of my argument in these terms.

<sup>7</sup> Following Korsgaard (2009, pp. 72–76), by “particularistic will” I mean to someone acting on grounds that are taken to only apply to the situation at hand and to have no bearing on other, even similar, cases.

since a standard that can be embraced or rejected at will can hardly be considered a genuinely normative standard. On this basis, the autonomy of the will has been found to be an *inconsistent* idea. For it commits one to hold that the will, in view of the principles by which it is guided, is under some constraint or normative necessitation—it is *autonomous*—while at the same time denying that the (autonomous) will is under any such a necessitation, for it is the author or chooser of its own principles, which it can change or abandon at its leisure (and this makes it normatively unconstrained).

On the other hand, insofar as agents are conceived as *autonomous*—and so as *norm-governed*, rather than as *norm-less*—they have to accept the authority of normative standards they are not in a position to alter or to rid themselves of. Which can be the case only if those are normative standards that autonomous agents lack control over, that is, only if they lack the power to legislate over the subject matter or areas of activity that come under the purview of those standards. In that case, however, agents cannot be claimed to have the capacity to determine themselves and govern their own conduct: they can hardly be considered self-legislating beings, capable of mastery over themselves. And, since those who do not legislate for themselves are not *autonomous*, autonomy appears to be a *deceptive* idea that is not what it claims to be.

In sum, autonomy is claimed to be an unstable notion, which is apt to fall either into lawlessness and arbitrariness or into subjection to externally sourced standards and requirements.<sup>8</sup> This dilemma manifests itself most clearly in combination with the thesis that the authority of the fundamental practical norms rests on autonomy. The claim here is that we are going to have a hard time trying to explain the authority of practical norms if we conceive of them as stemming from an agent's will, for those norms are going to appear to be completely contingent, and so at the discretion of a self-regulating will. On the other hand, if we insist on their authority being impersonal, we will be unable to present the fundamental practical norms as self-legislated. That autonomy can be seen to be entrapped between two opposite forces—the arbitrary self-determination of the will and the heteronomy of an independently imposed constraint—means that, when the authority of practical norms is grounded in autonomy, it poses the challenge of clarifying how an unstable, or even self-contradictory, notion can provide the foundation for the authority of the fundamental practical norms. For how can something as problematic as autonomy be thought to have the potential to ground something else, thereby securing the authority of the fundamental practical norms?

### 3 A constitutivist way out of the paradox of autonomy

In this section, I intend to argue that there is nothing about the *broad idea*, or *concept*, of autonomy that brings about the paradox discussed; rather, the paradox may or may not come out depending on the *specific conception* of autonomy one embraces. Namely, far from being inherent in the general idea of autonomy, the paradox is the result of a philosophically questionable declination, or specification, of that general

<sup>8</sup> The paradox of autonomy is presented in similar terms in Reath (2006, pp. 92–96), Khurana (2013, pp. 50–1 and (2017) pp. 68–85), O'Neill (2004, pp. 14–6 and (2013), pp. 282–5), Sensen (2013, pp. 267–70), and Kleingeld-Willaschek (2019, pp. 2–3), among other places.

idea. In order to support this conclusion, in what follows I will introduce a *conception* of autonomy that, I will argue, generates no paradox when it is appealed to ground the authority of the fundamental practical norms. This way, I intend to elicit the fact that the paradox is not located in autonomy in itself. The paradox, in sum, is not to be found in in the concept of autonomy, but rather it is triggered by certain understandings of autonomy and thus is contingently associated with some declinations of autonomy as a distinctive notion, as opposed to being conceptually connected to that very notion. Likewise important, the understandings of autonomy that produce the paradox are not only philosophically questionable but also far from necessary. In fact, there is at least one credible specification of the concept of autonomy—the conception that I will introduce in this section—under which the paradox of autonomy does not arise.

In my understanding, or conception, of (the broad idea of) autonomy, I aim to account for the fact that any credible view of autonomy must simultaneously (a) locate autonomy *within agency*—autonomy is an internal capacity we exercise qua agents and thus not a heteronomous imposition from the outside, but rather a necessary condition of agency itself—and (b) connect autonomy to something *fixed* that falls *outside* the range of what can be willed by this or that individual agent. This means that I do agree with the critics of autonomy that there are two structural conditions that (any view of) autonomy needs to meet. For one thing, autonomy needs to be acknowledged to be a power of agency that lies *in* the will, rather than being *superimposed on* the will. For another, autonomy needs to be conceptualised as lawlike, something that guides action on the basis of norms or principles that cannot simply be dismissed by fiat. If we do away with condition (a), we take out the *autos* in autonomy, and the result is “*auto-less autonomy*” and thus the very *opposite* of autonomy: heteronomy in disguise. If we do away with condition (b), we take out the *nomos* in autonomy, and the result is “*nomos-less autonomy*,” where autonomy resolves itself into the power of the will to arbitrarily choose any course of action under any criterion which can be swapped for any other criterion at any time. Both conceptions of autonomy are incoherent, because in taking out an element that is essential to autonomy, they each end up contradicting their premises, by thus yielding in one case a world of *heteronomous* agents who cannot be *their own* source of action, and in the other a world of *anomic* agents who can only point out the common source of their actions in their *particularistic and arbitrary* will. This suggests that if one can work both conditions (a) and (b) into the conception of autonomy, one can avoid the attendant paradox. Let me therefore take that suggestion as my starting point.

### 3.1 The conception of autonomy as constitutive will

As a point of entry into a conception of the autonomy of the will that leads to no paradox in the attempt to ground the authority of the fundamental practical norms, I first define the domain to which autonomous beings and the norms they are governed by pertain. This is the domain of action—that is, the practical domain—where a key normative question is, How ought we to act? Engaging with this domain and question requires us to think about our action, reason about it, make choices on that basis, and execute them. In this domain, our distinctive existence is thus fundamentally and

necessarily defined by the capacity to undertake action. For without that capacity we could not count as inhabitants of the practical territory.

From this premise, we get the further idea that the *capacity for action* figures as an essential condition absent which an individual cannot be recognised as a practical entity and so cannot be referred as an agent. The set of dispositional features enabling us to embark on action, in other words, establishes our minimal constitution in the practical domain. To this arrangement of dispositional properties I will refer here as the *constitutive will*. The constitutive will, as it is introduced here, provides both the precondition of action and the kernel of agency. Indeed, it should be seen as the *structure* around which agency revolves, meaning the fundamental *constitution* we possess within the domain of action. More specifically, such a structure, or constitution, consists of a constellation of basic dispositional properties: the dispositional properties enabling one to perform actions and thus making someone an agent. That is, the constitutive will is best understood as the necessary structural, or formal, precondition for an individual to be capable of acting and so to both exist in the practical domain and be placed under norms of action. As such, the constitutive will is not only the essential building block of the practical domain—the shaping hand of our acting self—but also the assemblage of dispositional traits singling out the mode of being we qua agents cannot do without. Hence, it accounts for a practical structure that is fundamental to, and non-disposable by, us. The constitutive will is in this sense expressive of the characteristic mode of being of an(y) agent: it describes what as a matter of constitutive necessity an agent *essentially* is, that is, someone with a complex of capacities or dispositions (a) enabling them to engage in action, and (b) lacking which or dismissing which they would not be recognisable as an agent to begin with.<sup>9</sup>

If that all sounds a bit abstract, or even somehow vague—removed from the space where people actually make the decisions that shape their lives in ways both big and small—it is because the constitutive will, as I constructed it here, forms the core part of agency and, thus, is in fact a *minimal* self. Apparently, describing the constitutive will as one's fundamental practical structure—the structure defining what one is qua agent *simpliciter*—is not to deny that in their practical exchanges single agents (can and ordinarily do) characterise themselves in richer, thicker, and more specific terms, by thus nurturing countless self-representations that are not entirely reducible to the mode of being dictated by the constitutive will.<sup>10</sup> When compared to those additional self-representations, the picture of the acting self that stems from one's constitutive will stands apart, though. For the constitutive will sets forth in outline what an agent necessarily is qua agent period, that is, it determines who someone capable of acting and capable of being subject to practical norms inherently is. Accordingly, only the constitutive will individuates the broadest, minimal, and necessary constitution that

<sup>9</sup> This statement is revealing of the fact that the identification of an agent with the constitutive will is neither a *metaphysical* necessity nor a *natural* necessity. It is, as the very phrase—constitutive will—indicates, a *constitutive* necessity. By constitutive necessity, I mean the *pragmatic* kind of necessity, namely, the kind of necessity that no agent can fail to acknowledge without being entangled in a *performative* contradiction.

<sup>10</sup> These additional self-representations, however, are less than fundamental to an agent, since they heavily depend on a number of contingent factors, such as an individual's interests, preferences, upbringings, and social backgrounds. Therefore, whereas those views of one's self may add to the self-representation shaped by the constitutive will, they substantially differ from the latter, since they are largely modifiable and disposable by each individual agent.

defines the acting selves, qua “bare” agents. In other terms, the constitutive will provides individual agents with a non-optional organising structure, by thus shaping a view of the acting self that, minimal and non-exclusive as such a view may be, everyone cannot coherently fail to endorse within the practical territory—this is a view that no agent can intelligibly refute without ceasing to be an agent.<sup>11</sup> Indeed, any failure on the part of an agent to conceive of themselves as a creature accounted for by the constitutive will would prevent such an agent from conceiving of themselves as who they basically are: (just) agents or beings capable of action.

To recapitulate the steps undertaken so far, I have defined the constitutive will as the arrangement of dispositional properties practically defining us at the broadest, minimal, and most fundamental level. Relatedly, I have claimed that the self-representation issuing from the constitutive will cannot be neglected by anyone who stays within the boundaries of the practical domain and so keeps their agential status, no matter what their subjective proclivities, individual inclinations, and personal variabilities are. Once it is so described, the mode of being the constitutive will leads to us is both “ours”—it defines who *we* are at a basic and deep level—and *non-disposable*—it is something each of us cannot discretionally dodge (at least as long as we do not forfeit our agency). That is, on the one hand, the constitutive will is a dimension internal to the agent—to any individual agent; on the other, it is not a dimension that the agent—any individual agent—is in the position to ditch, since the constitutive will does not define one of the several contingent modes of being agents may or may not embrace, but rather it defines the necessary and non-alterable framework agents incorporate just by virtue of inhabiting the practical domain. The constitutive will, in other terms, has a twofold dimension: it simultaneously features a self-centred component and a norm-governed component.

The further suggestion that I am putting forward here is *to define the autonomy of the will as the distinctive set of dispositional properties making up our fundamental practical structure, namely, as (what I have referred to as) the constitutive will*. In this proposal, in which the constitutive will establishes what an autonomous creature is—call it the *conception of autonomy-as-constitutive-will* for short—we grasp the idea of autonomy once we describe it as the product, or emergent outcome, of the constellation of traits our constitutive will consists in. Conversely, autonomy is best conceived as constitutive wilfulness, namely, as the structure essentially associated with our agency, or the framework establishing the basic and non-forfeitable layer of our practical existence.

Incidentally, this means that the relation between constitutive will and autonomy I introduce here is *definitional*, as distinct of either *teleological* or *functional*. Unlike other champions of constitutivism about practical reasons, who regard autonomy as an essential *aim* we, qua agents, constitutively have,<sup>12</sup> I do not claim that autonomy is the

<sup>11</sup> This conclusion has been challenged by David Enoch, who in a couple of influential works (Enoch, 2006, 2011) has devised (what is known in the literature by the name of) the shmagency objection. I take the shmagency objection to be by no means fatal to the approach I endorse here (and ask the reader to do the same, at least for the sake of argument), since it has been convincingly refuted in Ferrero (2009 and 2018), O’Hagan (2014), Silverstein (2015), Smith (2015). Rosati (2016), and Paakkunainen (2018), or so I think.

<sup>12</sup> I am referring to Velleman (2000 and 2009) and Smith (2013 and 2015), for instance.



fundamental *goal* of the constitutive will and so the constitutive will necessarily *aims for* autonomy. Nor does my version of constitutivism constructs autonomy as some paradigmatic *function* of the constitutive will. Rather, the variant of constitutivism I am committed to here takes the constitutive will to be the notion in terms of which autonomy is *defined*. In my view, then, the constitutive will is the essential *condition of existence* of autonomy. Related to this, in the version of constitutivism I endorse autonomy is conceived and constructed as a predicate of the *will*, as opposed to being a feature that certain *norms* of our agency—the constitutive norms of our agency—have.

Now, to go back to the main point I am arguing for in this essay, the conception of autonomy-as-constitutive-will has the resources to prevent the paradox of autonomy from arising. For by construction the traits the constitutive will is made of are properties abstracting from personal differences and yet sourced in us, who accordingly cannot be claimed to be determined by alien factors and external forces when implementing our constitutive will. This is key, because it means that, insofar as autonomy is conceptualised as constitutive wilfulness—namely, insofar as we take the constitutive will to define what it means for us to be autonomous agents—the paradox of autonomy does not show up. For the constitutive will is an arrangement of dispositional properties that are at once *internal* to the self and *impersonal*. As long as they are *internal* to the self, they cannot be considered heteronomous in quality, since they are not imposed on the self from outside forces. As long as those traits are *impersonal*, they are non-alterable by a single self, which accordingly cannot modify those (inner) properties as it deems fit and so even in a completely arbitrary fashion. For these reasons, the features defining the constitutive will, as I introduce it, are neither heteronomous nor arbitrary. Accordingly, autonomy, once autonomy is conceived in terms of the constitutive self, cannot be claimed to be either inconsistent or deceptive. Related to this, the conception of autonomy-as-constitutive-will can be claimed to be able to fully accommodate both the autos-component and the nomos-component that inherently characterise the general concept of autonomy.

To unpack these dense statements, on the one hand, the conception of autonomy-as-constitutive-will duly accounts for the internality of autonomy—its being self-related—since on this conception autonomy is claimed to consist of dispositional properties the *self* in its practical instantiation possesses. Far from being something imposed on the agent by some *external force*, the constitutive will is a construct internal to, and distinctive of, the acting self. Relatedly, when it is defined as constitutive wilfulness, autonomy stands for a set of features all agents (necessarily) identify themselves with and acknowledge as their own insofar as they think clearly. Therefore, autonomy, as I propose we should conceive of it, stands for an arrangement of traits that can be claimed to belong to, or be owned by, the agent in a genuine and deep sense. And, the fact that the constitutive will can hardly be said to be an external, or even *heteronomous*, construct forced upon us by some foreign powerhouse means that autonomy-as-constitutive-will is hardly a deceptive notion susceptible of dissolving into heteronomy.

On the other hand, the conception of autonomy-as-constitutive-will makes room for the *norm-related* component that is conceptually associated with the idea of autonomy. The conception of autonomy-as-constitutive-will secures our independence from the normative influences that are external to, and so are not conceptually linked with,

our agency, such as social constraints, conventional bounds, traditions, psychological pressures, etc. Nonetheless, autonomous beings, when they are autonomous in the sense fixed by the idea of autonomy-as-constitutive-will, are *hardly unbound, unconstrained, and normless* individuals. On the contrary, they are placed under the fundamental norms that structure and make possible the constitutive will as well as govern its workings. For, qua a distinctive structure, constitution, or mode of being, the constitutive will is governed by certain principles, norms, and standards: the principles, norms, and standards that distinctively characterise, and derive from, such a structure, constitution, or mode of being. Indeed, having a structure, constitution, or mode of being means to be subjected to certain normative constraints too. More specifically, it means to be subjected to the principles, norms, and standards that define and govern such a structure, constitution, or mode of being. Accordingly, autonomy, as it is conceived here, is inseparable from the idea that individual agents are under certain normative constraints—those defining the constitutive will. And those constraints are practically necessary in the sense that no agent is in a position to disregard them. Thus, the conception of autonomy-as-constitutive-will fully accounts for the fact that a distinctively normative dimension is built into the very idea of autonomy and thus autonomous agents are subjected to principles, norms, and standards of action that enjoy a measure of independence from this or that agent. On this basis, it can be claimed that the constitutive will does not merely describe who we are but also it incorporates standards we ought to follow. Correspondingly, autonomy, when defined as the constitutive will, includes a normative dimension equipping it to work as the source of the fundamental practical norms.

Crucially, as long as both a self-referential component and a norm-related dimension are inscribed in autonomy-as-constitutive-will—the constitutive will delimits a space that originates within agency and yet cannot be modified as an individual agent wishes—the paradox of autonomy is conceptually prevented from arising. This is the case because the conception of autonomy-as-constitutive-will eases the tension between self-determination and independence of the will, on the one hand, and subjection to normative standards and absence of arbitrariness, on the other. That is to say, insofar as it is defined as (the combination of dispositional properties defining) the constitutive will, autonomy incorporates both the demand of being norm-related and the requirement of being an expression of the self—indeed, autonomy so conceived corresponds to our basic and generic self. For this reason, the conception of autonomy-as-constitutive-will, as it will be further characterised and elaborated on in the next section, has the potential to block the very possibility that the paradox of autonomy arises. Relatedly, the conception of autonomy-as-constitutive-will removes the main obstacle to the recognition of autonomy as the ground of the authority of the fundamental practical norms. Indeed, when it is defined by the constitutive will, autonomy has the resources to ground the authority of the fundamental practical norms, since its contents are neither established by individual acts of will—and so are not at the mercy of specific agents—nor fixed by self-unrelated and thus heteronomous sources. These points will be articulated in further detail and expanded on in the next section, where I spell out and defend the implications of the philosophical framework introduced so far.

### 3.2 Autonomy-as-constitutive-will, action, and agency

In the theoretical framework introduced in this work, the possibility of regarding autonomy as the ground of the authority of normative standards heavily depends on how the constitutive will is conceptualised. Therefore, at this stage of the argument, one may well wonder what the constitutive will and so autonomy-as-constitutive-will ultimately amount to.

A caveat needs to be introduced before I proceed to addressing this question. As already mentioned in the introductory section, the conception of constitutive will and, more generally, the variant of constitutivism about practical reasons underpinning that conception are deeply indebted to Korsgaard's constitutivist reinterpretation of these that are central to Kant's practical philosophy.<sup>13</sup> This debt to Korsgaard's philosophical framework cannot be overstated and will become even more apparent in this section, once my understanding of action and agency will be spelled out. Accordingly, it is fair to say that there is no *significant* divergence between my approach to practical issues and Korsgaard's. Nonetheless I think that the views defended here are *not reducible* to Korsgaard's constitutivist account of action, agency, will, and autonomy at least under an important respect, which will be elaborated on below.<sup>14</sup>

Now, the question, what do the constitutive will and so autonomy-as-constitutive-will ultimately amount to? can in turn be answered by going back to the statement that in the proposed construction the constitutive will is best understood as the combination of dispositions that make action possible—the constitutive will being the precondition, or presupposition, of action. Conceptualising the constitutive will as an arrangement of action-enabling conditions (and thus as a set of dispositional properties in the absence of which no action would take place) means that, if we want to get a sense of what the constitutive will and (so) what autonomy-as-constitutive-will consists of, we need to have a closer look at the very idea the constitutive will is a presupposition of, namely, action. So, in my theoretical construction, an exploration of the distinctiveness of action has the potential to provide us with a direct insight into autonomy(-as-constitutive-will) itself. Crucially, insofar as it bears some conceptual connection with autonomy, *action* is not to be understood as a broad, indiscriminating, and generic idea singling out anything that is done with a purpose.<sup>15</sup> By contrast, action, as it relates to autonomy, is best conceived as a distinctive form of conduct, namely, conduct that is *non-necessitated* and, thus, belongs to the acting self. And, a sound strategy to elicit the defining traits of action in this specific and distinctiveness sense—action qua non-necessitated behaviour that belongs with the individual performing it—consists in differentiating it from another idea that can be located in the proximity of, and yet should not be conflated with, action, namely, reaction.

<sup>13</sup> I am referring in particular to Korsgaard (1996 and 2009). My variant of constitutivism owes much to Reath (2006, 2013, and 2019), too.

<sup>14</sup> Cf. pp. 18–19 in this section.

<sup>15</sup> Insofar as it is taken in its broad and generic meaning, action stands for a naturalistic term that is coextensive with *behaviour* at large, meaning any *movement* that is *causally guided* by mental states of one kind or another. This conception of action is theorised in Davidson (1980), Bratman (1987, 1999, and 2007), and Velleman (1989 and 2000), among others.

*Reaction* is a kind of conduct that, unlike action, can be qualified as *necessitated*, since it is dictated by the combination of external circumstances and states operating within and through the agent, such as instincts, desires, drives, inclinations, and so forth—henceforth I will refer to these states generically as “impulses”. Under the broad rubric of reaction, we can place both reflex deeds, like ducking in order to dodge a ball that has been hurled at us, and adaptive behaviours, like changing one’s path when seeing someone approaching in order to minimise the risk of infection during a coronavirus pandemic. Whilst reflex deeds can be described as controlled and purposive, adaptive behaviours are additionally intentional. Both of these kinds of doings, however, should be understood as reactive in quality—and so instances of reaction, as distinct of action (when action is used in its specific and distinctive sense). For they both are extemporaneous adaptations to the circumstances we find ourselves in and so are best described as ways of responding off-the-cuff to pressures, whether internal or external—that is, they constitute a reaction to stimuli of some kind. In contrast to reacting, as it is involved in reflex and adaptive courses of conduct, *acting* requires one to engage in a kind of behaviour that is not only controlled, purposive, and intentional but also based on normative standards one sets for himself by way of deliberating about what to do—call it *principled* conduct for short.<sup>16</sup> Only in principled conduct, we have *action*, as distinct of reaction, since only in the presence of principled conduct one can be said to play a fully active and genuinely independent role in setting the relevant course of conduct. Related to this, exclusively in principled conduct one’s movements can be qualified as more than just sub personal, or quasi-event-like, occurrences happening through someone to be loosely qualifiable as the author of the resulting behaviour.<sup>17</sup> Which is also why, only in principled conduct *agency* is distinguishable from subjectivity.

*Subjectivity*, as I use the term here, stands for the condition of providing the space within which some deed occurs that can hardly be associated with, and attributed to, the acting self as a whole—the subject merely being the conduit through which some behaviour, as it is dictated by impulses and circumstances, materialises. By contrast, *agency* refers to the capacity for embarking in a performance that is actively undertaken and can be fully ascribed to those who perform it (with those who perform it being, in turn, understood as inherently unified creatures, as distinct from arrays of disparate parts pushing one to move around). For, unless we associate agency with a type of doing featuring among its essential traits not just control, purposiveness, and intentionality, but also guidance by principles we are not in the position to secure the conclusion that action is a “movement attributable to an agent, considered as an integrated whole” and, on this basis, is distinguishable from mere reaction.<sup>18</sup> This is the reason why only conduct resulting from deliberation (and so only the kind of deed to which I referred above as principled conduct) can be regarded as a performance that is owned by, and attributable to, the agent—it is an exercise of agency. Correspondingly, the

<sup>16</sup> For an idea of action that, whilst it cannot be fully equated to the one put forward here, is likewise discriminating and conceptually akin to it, see Korsgaard (2009, pp. 8–14).

<sup>17</sup> This notion of action is further discussed in Berteaux (2009, pp. 191–6).

<sup>18</sup> Korsgaard (2009, p. 45). On the idea of action as an activity undertaken by a whole agent, understood as a unified practical structure over and above her specific parts, see also Korsgaard (2009, pp. 18–26 and pp. 72–76; and (2014,) pp. 96–9).

other kinds of deeds are best conceptualised as undertakings that are brought about by some possibly unendorsed dimension at work within the agent or forced on the agent by external factors—that is, they are the expressions of one’s mere subjectivity.

This stance justifies the claim that the doer can be qualified as an *agent*, as distinct of a *subject*, only when a course of conduct can be imputed to a unified acting self, as distinct of being caused by an impulse within the acting unit or to some subpart of an individual pulling such an individual into different directions or even tearing an individual apart and turning it into a recipient where events merely occur. Otherwise stated, an agent is not just someone who is capable of settling one’s course of conduct at the time of action in a non-purely-deterministic fashion—this merely being a subject. An agent is best conceived as someone who takes a leading role towards the resulting practical output, which can thus be qualified as fully authored by, and entirely attributed to, someone. This statement, though, requires the understanding of action as principled conduct, and thus as a form of deliberated activity essentially depending on one’s adoption of general norms. For only those who undertake principled conduct can be said to have a controlling influence on the resulting behaviour. Namely, only those who proceed from a standard mediating between them, on the one hand, and their internal components and the surrounding environment, on the other—a standard that enables them to do more than contingently reacting, or just adapting, to the internal pressures they experience and the situation they happen to be in—can be qualified as authors, or owners, of their conduct.<sup>19</sup>

Now, in the theoretical framework I am relying on in this work, where the meaning of autonomy is established by the properties enabling one to act (with acting to be kept distinct from reacting), the *constitutive will*, as it defines autonomy, is fixed by the distinctive set of dispositional properties making one the kind of creature that is able to carry out deliberate conduct. And, insofar as the constitutive will refers to the combination of dispositional properties one needs to possess in order to embark on action (in turn, understood as an undertaking shaped by a principle one has deliberately adopted), the question “What does autonomy-as-constitutive-will amount to?” finds the following answer. *Autonomy-as-constitutive-will* is the constellation of dispositional properties, or capacities, which one needs to possess in order for them to be able to deliberate on practical issues and embark in principled conduct in consequence of such a deliberation.

This characterisation of autonomy-as-constitutive-will may be regarded by some as insufficiently specific, since it only points towards the direction one has to look at in order to establish the content of autonomy-as-constitutive-will, whilst failing to indicate what such a content *in concreto* is. Yet, the formulation of autonomy-as-constitutive-will just introduced is far from uninformative or empty. For the discussion carried out so far at least indicates that, in my approach, the constitutive will and so autonomy-as-constitutive-will are defined by no less than two distinctive capacities: the capacity to reflect on our conduct—to which capacity I would refer as “reflectivity—and the capacity to anchor our practical choices on rational considerations—call it “rationality”. In turn, these fundamental capacities are conceptually associated with

---

<sup>19</sup> I thank an anonymous reviewer for this journal for recommending that I specifically address the distinction between subjectivity and agency in this context.

the norms and constraints that are essential to their exercise. Qua essential dimensions of reflectivity and rationality, those norms and constraints are the very conditions of existence, or possibility, of reflectivity and rationality, qua distinctive and essential agential capacities. And, insofar as autonomous agency is defined by certain capacities, the basic standards setting out the conditions of possibility of those capacities provide the barebones of the arrangement of the fundamental practical norms the authority of which can be grounded in the constitutive will. It is with a view of further spelling out what autonomy-as-constitutive-will amounts to, then, that in this conclusive part of the argument I will elucidate those essential capacities and fundamental norms in some detail.

I should quickly add, though, that the purpose of the treatment that follows does not consist in identifying all the fundamental practical norms autonomous agents are subjected to. For determining the detailed contents of autonomy-as-constitutive-will and specifying the fundamental practical norms the authority of which is grounded in autonomy-as-constitutive-will go well beyond the scope of the argument deployed in this work. Indeed, the aim of the argument devised in this contribution is (confined to) establishing and consolidating the link between autonomy and constitutive will, by thus (just) laying the foundations of an inquiry into the contents of autonomy-as-constitutive-will and its fundamental norms. Relatedly, carrying out such an inquiry in full is a challenge to be confronted on another occasion. Hence the largely unsystematic, exploratory, and incomplete quality of the remarks that follow.<sup>20</sup>

As to the first capacity defining autonomy-as-constitutive-will—reflectivity—it stands for the general capacity to think before acting. Reflectivity enables the agent to take a critical standpoint from which to assess a situation and hence not be forced into a certain line of conduct just because this is what the present circumstances or the agent's impulses urge someone to do. Reflectivity, thus, is the mediating element that makes it possible for agents to step back and take a detached perspective on the situation before them, in such a way that agents are not confined to the blind mechanism of stimulus and reaction. It is this capacity that enables an agent to respond to the promptings of the environment not simply by yielding to their momentary impulses but by taking into account the principles they endorse.<sup>21</sup> So, instead of doing what seems most appealing (because most likely to satisfy an extemporaneous impulse) or most natural (because the situation makes it seem that way), an agent can create a situation in which these forces are not necessarily constraining. As a constitutive feature of agency, therefore, reflectivity enables one to respond non-reactively to the promptings of the environment by bringing into play a standard on the basis of which the immediate givens of circumstance and impulses can be critically assessed. Reflectivity, or at least its exercise by agents, is in turn made possible by the existence of certain norms. Among those norms, which are thus constitutive of reflectivity, they figure

---

<sup>20</sup> What follows is meant to address a point explicitly raised by an anonymous reviewer for this journal, whom I thank most warmly for giving me the opportunity to further expand on the contents of my conception of autonomy.

<sup>21</sup> As Connie Rosati (1995: p. 61) puts it, a reflecting self “comes to dissociate herself from some of her desires, motivations, and traits, while identifying with others. This capacity to step back, to engage in self-reflection, gives persons a kind of freedom from identity with their immediate activities or their immediate motivational tendencies”.

some formal demands, such as those to not necessarily yield to one's instincts, to be alert to the necessity of dissociating oneself at least occasionally from the desires and urges one may experience, and to establish a distinction between oneself and one's impromptu motives, immediate impulses, extemporaneous wishes, and off-the-cuff reactions. Those formal requirements should then be included in the set of the fundamental practical norms that can be grounded in autonomy, once the latter is conceived in terms of the properties defining the constitutive will.

But, it takes more than reflectivity and its defining norms to account for the whole of what is distinctive about agency. Indeed, reflectivity conceptually calls for at least another fundamental capacity, which I take to be constitutive of our agency too: rationality. This is the case because reflectivity, in enabling one to reflect on one's conduct, gives one as well the ability to reflect on oneself. In turn, self-reflection gives one self-consciousness, a turning inward on oneself and one's inner states. And self-consciousness enables one to see that even though the self functions as a whole, in unity, it does not form a single, undifferentiated substance but is rather a compound—its components being one's different impulses. Through this discovery, an agent comes to appreciate that the unity of the self consists not in a state given to someone but in a composition and synthesis by which someone's multiple elements work together. These elements exert forces that drive an individual's action in different directions not necessarily compatible or coherent with each another—hence the synthesis needed to achieve continuity of action. For an acting self will have no coherent conduct unless the forces operating within itself cannot be reduced to unity (or at least to a working whole). And in order to achieve unity one must prioritise; and do so consciously, by thus establishing among the elements of the self a ranking on which basis to determine what course of action one should take on any particular occasion. This task requires that one choose the grounds from which to act; which choice in turn requires reasoning. For in order to prioritise the constituents of the self, agents must take into account not only their raw instinctual powers but also the role played by rational considerations. Thus, constitutive of agency is the capacity to grasp, work with, and process reasons. This is what I call “rationality”.

Rationality describes an agent's essential responsiveness to reasons. It enables an agent to work out principles and to take them into account as determinants of conduct, by building them into our general framework of action and giving them the proper practical weight. It is therefore by virtue of rationality that some of the forces bearing on an agent are endowed with a distinctive significance and are prioritised over other forces or influences. And though rationality does not stay the forces of impulses, for these still operate within the self and can therefore tilt action this way or that, it does free agents from a strict dependence on such forces by giving agents the power to choose whether or not to be guided by their impulses.<sup>22</sup> The demands of these inhabitants of the self are, thus, filtered by rationality and fashioned into principles that synthesise

---

<sup>22</sup> In this function, rationality can be said to achieve what Kant referred to as negative freedom, meaning our independence from determination by desire and inclination: rationality is what enables principled conduct, by which expression I mean precisely the human ability not to be determined by such alien causes as desires and inclinations. In the words of Schneewind (1998: p. 20), “desires are among the necessary conditions of human action, but they are not sufficient conditions”. On negative freedom, see Kant (1930: pp. 121–2), and Kant (1956: pp. 28–30). For critical remarks, see Allison (1998), Hill (1998), and Guyer (2005: pp. 115–45).

and unify them. This process involves weighing reasons when several come to bear on the matter as well as working several reasons together into a coherent whole, or into a wider scheme, and considering the consequences of action in choosing between alternative courses of action. That way principles, as distinct of impulses, become the ultimate basis of action. Rationality, in sum, is key to making it possible for us to have action as distinct from mere reaction. Which is why the ability to appreciate reasons and act on this basis specifically and distinctively describes anyone capable of action, and so anyone who is an agent, namely, it is an essential element in the definition, or constitutive, of agency. Rationality too is a capacity the existence and use of which is associated with certain norms, which thus provide the very conditions of possibility of rationality, as an agent's distinctive dispositional property. These norms include the requirements to act consistently, to devise and pursue plans of action that are mutually coherent and not self-defeating over time, to take the means that are necessary in relation to the ends one has deliberated to pursue, and to give some degree of continuity and implementation to the outcome of one's deliberative processes. These requirements too are part of what I consider the wider arrangement of fundamental practical norms that can be grounded in autonomy-as-constitutive-will.

Apparently, the conceptions of action, agency, constitutive will, and, thus, autonomy just introduced are rooted in Korsgaard's practical philosophy. Yet, I believe that the views defended here as well as the idea of autonomy-as-constitutive-will are not entirely reducible to Korsgaard's constitutivist account of action, agency, will, and autonomy for the following main reason. Korsgaard (2009, pp. 81–108) singles out two fundamental principles as constitutive of *action*: Kant's hypothetical imperative and categorical imperative. This means that for Korsgaard at its core action is an efficacious and autonomous movement. Action is an efficacious movement, since it is informed to the hypothetical imperative, which directs one to choose the essential means to one's ends. Action is an autonomous movement, because it conforms to the categorical imperative, which instructs one to act on rational considerations, as distinct of impulses. From this, it follows the further conclusion that *agency* too is shaped by the two main Kantian imperatives. For Korsgaard, an agent is someone capable of undertaking efficacious and autonomous courses of conduct (in the specific, hypothetical-and-categorical-imperatives-related, sense in which Korsgaard describes efficacy and autonomy). The account defended here belongs to the same family of approaches to action theory as Korsgaard's (these are the so-called "authorship views" of action).<sup>23</sup> Moreover, I take from Korsgaard the idea that agency makes us unified wholes who own, or author, our courses of conduct, by thus producing them in more than a mere reactive sense. However, my version of constitutivism parts its way from Korsgaard's version of constitutivism, since it does not establish any constitutive connection between action and agency, on the one hand, and hypothetical imperative and categorical imperative, on the other hand. By contrast, in my constitutivist approach I specifically understand action as principled conduct and, related to this, agency as the set of capacities one needs to possess in order to be able to carry out courses of conduct that are grounded in principle. In the version of constitutivism about practical reasons underlying my argument, then, the traits that are taken to be

<sup>23</sup> I borrow this terminology from Millgram (2020).



constitutive, or defining, of our practical domain—reflectivity and rationality, as these dispositional properties have just been introduced—cannot be exhaustively accounted for in terms of the hypothetical imperative and the categorical imperative. Which is why the resulting conception of action and agency should be acknowledged to be at least partly divergent from, and on the whole irreducible to, Korsgaard's.

## 4 Conclusion

The central question explored in this work was whether autonomy can ground the authority of practical norms in such a way as to avoid the paradox of autonomy, meaning the dilemma under which we start out from a conception of autonomy only to end up with something that contradicts it—either an arbitrary agent (one incapable of acting on the basis of any norm) or a heteronomous agent (on whom a form of action is imposed that originates from an outside source). I argued that the possibility of the paradox sets out positive demands that a sound conception of autonomy should account for: any convincing view of autonomy needs to avoid two pitfalls—arbitrariness and heteronomy—by thus reconciling two basic components of autonomy qua a distinctive notion—lawfulness and independence of the will. This argument led me to defending the claim that the paradox is not inherent to the concept of autonomy itself but rather it is generated by the way in which autonomy may be constructed—its conception. On this basis, I moved to introduce a conception of autonomy, at whose core lies the idea of the constitutive will, which not only can be understood as an outcome of the will that single agents cannot reshape as they wish, but also (and relatedly) can be claimed to be equipped to ground the authority of the fundamental practical norms.

In a nutshell, the main strategy I used to defend autonomy as a coherent and powerful ground of normative authority consisted in redefining the *autonomy of the will*—the arrangement of dispositional features associated with, and distinctive of, free and rational selves—as *agential autonomy*—the set of dispositions enabling someone to be the author of one's action and, relatedly, such an action to be attributable to them. That is, I framed autonomy as the assemblage of dispositional properties an(y) agent needs to possess in order for them to be able to undertake action. In this process, I also accounted for action as deliberated conduct—this is not a merely naturalistic and conceptually indiscriminating meaning of action, but an understanding of action that was made to emerge through a reflection on what distinguishes action, qua distinctive undertaking, from both movement, or behaviour in a generic sense, and mere reaction. My argumentative strategy, which is best interpreted as a variant of constitutivism about practical reasons, therefore, led me to introduce a specific meaning of autonomy, action and agency.

By way of concluding my treatment of the autonomy of the will, I should notice that the variant of constitutivism about practical reasons I relied on in setting up the argument denying the paradox of autonomy is to be doubly qualified. This double qualification contributes to the originality of my approach and keeps it apart from other versions of constitutivism, as they are theorised in the literature. First, my constitutivist stance is not reducible to the claim that our whole *agency* necessarily amounts to principled and reflective individuality. The thesis I am committed to is rather that

the *core of agency*, as it is defined by the *central meaning* of action, needs to be located in such a form of individuality. This means, among other things, that the form of constitutivism I am committed to is introduced to vindicate the authority of (just) the *fundamental* practical norms, as opposed to *all* practical norms. Namely, my constitutivism supports the conclusion that (only) the standards which define the kernel of our practical existence, as it is established by our capacity to undertake action (where action is understood in its central meaning), can make necessary normative claims on us. By contrast, the authority of the other practical norms does not (need to) rest (directly) in autonomy-as-constitutive-will.

The second qualification that is worth adding in regard to this conception is that the form of constitutivism I upheld in my treatment can be characterised as Kantian, since it is based on, and shaped by, an idea of autonomy originally introduced and consistently defended in Immanuel Kant's practical philosophy.<sup>24</sup> Indeed, my constitutivist conception is deeply indebted to the influential contributions that some contemporary philosophers working within the Kantian tradition have made in their reinterpreting the central theses of Kant's construction through a constitutivist lens.<sup>25</sup> But that is as far as I am willing to stretch the Kantian analogy. For, while the idea of autonomy-as-constitutive-will is Kantian in spirit, I am neither claiming that for Kant in autonomy lies the foundation on which rests the authority of the fundamental practical norms nor am I committing to the view that Kant's account of autonomy is constitutivist. In fact, on the one hand, what Kant's philosophy seems to indicate is that this authority, at least as far as the categorical imperative is concerned, is grounded not in autonomy but in an *a priori* principle of practical reason.<sup>26</sup> On the other hand, in Kant's philosophical project meta-ethical interests, which in recent years have figured prominently in various constructivist and realist approaches to Kant's practical thought, do not occupy as central a place.<sup>27</sup> Accordingly, while the constitutivist redefinition of autonomy argued for in this paper is Kantian in spirit, its intent is neither to offer an exegesis of Kant's philosophy nor to contribute to a novel understanding of it.

**Acknowledgements** This research has been supported by a Maria Zambrano Fellowship at the University of Barcelona. I would like to thank Claudia Blöser, Fabian Burt, Gabriele Gava, Alexander Kaib, and the other participants in the weekly Colloquium held at the Chair of Modern Philosophy of the Goethe Universität zu Frankfurt for insightful comments on the argument presented in this article. I am especially

<sup>24</sup> Kantian constitutivism is just one of the possible variants of constitutivism about practical reasons. For a concise critical overview of the most influential versions of constitutivism see Katsafanas (2018).

<sup>25</sup> Here I am referring in particular to Korsgaard (1996 and 2009) and Reath (2006, 2013, and 2019). Implicit in my choice to proceed on a constitutivist basis in addressing the question of the authority of the fundamental practical norms, therefore, is an assumption about the superiority of the constitutivist reading of Kant's practical philosophy in comparison with the realist and constructivist readings as a basis on which to develop Kant's original ideas or to use them as building blocks in developing new conceptions through which to address various issues in practical philosophy. Among the most influential realist readings of Kant's moral philosophy are Ameriks (2000), Guyer (2000 and 2007), Wood (2007, pp. 106–22), Kain (2004 and 2010), Stern (2012), and Schönecker (2013). Constructivist (albeit not specifically constitutivist) readings of Kant's project in practical philosophy can instead be found in Rawls (1980), O'Neill (1989), Hill (1992), and Bagnoli (2013, 2014, and 2017). And while it falls well beyond the scope of this paper to explain that assumption, it should at least be noted here that the argument presented in this paper is best understood bearing such an assumption in mind.

<sup>26</sup> This argument is convincingly deployed in Kleingeld and Willaschek (2019).

<sup>27</sup> This point is made in Bacin (2018).

grateful to Achim Brosch, Andrews Reath, Filippo Valente, and Marcus Willaschek for their feedback to, and constructive criticism of, an earlier draft of this work. Needless to say, responsibility for the views expressed herein, as well as for any error of form or content, rests solely with me.

## Declarations

**Conflict of interest** I formally declare that I have no financial or non-financial interest that is directly or indirectly related to the work submitted for publication.

## References

- Allison, H. (1990). *Kant's theory of freedom*. Cambridge University Press.
- Allison, H. (1998). Morality and freedom: Kant's reciprocity thesis. In P. Guyer (Ed.), *Kant's groundwork of the metaphysics of morals* (pp. 273–302). Rowan & Littlefield.
- Ameriks, K. (2000). *Kant and the fate of autonomy: Problems in the appropriation of the critical philosophy*. Cambridge University Press.
- Bacin, S. (2018). Moral realism by other means: The hybrid nature of Kant's practical rationalism. In E. Schmidt & R. dos Santos (Eds.), *Realism and anti-realism in Kant's moral philosophy* (pp. 153–176). De Gruyter.
- Bagnoli, C. (2013). Constructivism about practical knowledge. In C. Bagnoli (Ed.), *Constructivism in ethics* (pp. 153–182). Cambridge University Press.
- Bagnoli, C. (2014). Starting points: Kantian constructivism reassessed. *Ratio Juris*, 27, 311–329.
- Bagnoli, C. (2017). Kant in metaethics: The paradox of moral autonomy, solved by publicity. In M. Altman (Ed.), *Kant handbook* (pp. 355–377). Palgrave.
- Berteau, S. (2009). *The normative claim of law*. Oxford.
- Berteau, S. (2013). Constitutivism and normativity: A qualified defence. *Philosophical Explorations*, 16, 81–95.
- Benjamin, J. (1998). *The bonds of love*. Pantheon.
- Bratman, M. (1987). *Intentions, plans, and practical reason*. Harvard University Press.
- Bratman, M. (1999). *Faces of intention: Selected essays on intention and agency*. Cambridge University Press.
- Bratman, M. (2007). *Structure of agency: Essays*. Oxford University Press.
- Davidson, D. (1980). *Essays on actions and events*. Oxford University Press.
- Engstrom, S. (2009). *The form of practical knowledge*. Cambridge University Press.
- Enoch, D. (2006). Agency, shmagency: Why Normativity won't come from what is constitutive of action. *Philosophical Review*, 115, 169–198.
- Enoch, D. (2011). Shmagency revisited. In M. Brady (Ed.), *New waves in metaethics* (pp. 208–233). Palgrave Macmillan.
- Ferrero, L. (2009). Constitutivism and the inescapability of agency. *Oxford Studies in Metaethics*, 2, 303–333.
- Ferrero, L. (2018). Inescapability Revisited. *Manuscrito*, 41, 113–158.
- Ferrero, L. (2019). The simple Constitutivist Move. *Philosophical Explorations*, 22, 146–162.
- Friedman, M. (2005). Autonomy and male dominance. In J. Christman & J. Anderson (Eds.), *Autonomy and the challenges to liberalism: New essays* (pp. 150–173). Cambridge University Press.
- Gewirth, A. (1978). *Reason and morality*. University of Chicago Press.
- Gilligan, C. (1982). *In a different voice*. Harvard University Press.
- Guyer, P. (2000). *Kant on Freedom, Law and Happiness*. Cambridge: Cambridge University Press.
- Guyer, P. (2005). *Kant's system of nature and freedom*. Oxford University Press.
- Guyer, P. (2007). *Kant's Groundwork for the Metaphysics of Morals: A Reader's Guide*. London: Continuum.
- Hill, T. (1992). Kantian constructivism in ethics. In T. Hill (Ed.), *Dignity and practical reason*. Cornell University Press.
- Hill, T. E. (1998). Kant's argument for the rationality of moral conduct. In P. Guyer (Ed.), *Kant's groundwork of the metaphysics of morals. Critical essays* (pp. 249–272). Rowan & Littlefield.

- Kain, P. (2004). Self-legislation in Kant's moral philosophy. *Archiv für Geschichte der Philosophie*, 86, 257–306.
- Kain, P. (2010). Practical cognition, intuition, and the fact of reason. In B. Lipscomb & J. Krueger (Eds.), *Kant's moral metaphysics: God, freedom and immortality* (pp. 211–230). de Gruyter.
- Kant, I. (1930). *Lectures on ethics (1775–1780)*. Methuen.
- Kant, I. (1956). *Critique of practical reason (1788)*. Bobbs Merrill.
- Katsafanas, P. (2018). Constitutivism about practical reason. In D. Star (Ed.), *Reasons and normativity* (pp. 367–391). Oxford University Press.
- Khurana, T. (2013). Paradoxes of autonomy. *Symposium*, 17, 50–74.
- Khurana, T. (2017). *Das Leben der Freiheit*. Suhrkamp.
- Kleingeld, P., & Willaschek, M. (2019). Autonomy without paradox: Kant, self-legislation and the moral law. *Philosophers' Imprint*, 19, 1–18.
- Korsgaard, C. (1996). *The sources of normativity*. Cambridge University Press.
- Korsgaard, C. (2009). *Self-constitution*. Oxford University Press.
- Korsgaard, C. (2014). The normative constitution of agency. In M. Vargas & G. Yaffe (Eds.), *Rational and social agency* (pp. 190–214). Oxford University Press.
- Lindeman, M. (2017). Constitutivism without normative thresholds. *Journal of Ethics and Social Philosophy*, 12, 231–258.
- Mayr, E. (2019). Blame for constitutivists: Kantian constitutivism and the victim's special standing to complain. *Philosophical Explorations*, 22, 117–129.
- Millgram, E. (2020). "Practical reason and the structure of actions", *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2020/entries/practical-reason-action/>
- O'Hagan, E. (2014). Shmagents, realism and constitutivism about rational norms. *The Journal of Value Inquiry*, 48, 17–31.
- O'Neill, O. (1989). *Constructions of reason*. Cambridge University Press.
- O'Neill, O. (2004). Self-legislation, autonomy and the form of law. In H. Nagl-Docekal & R. Langthaler (Eds.), *Recht—geschichte—religion* (pp. 13–26). Akademie Verlag.
- O'Neill, O. (2013). Postscript: Heteronomy as the clue to Kantian autonomy. In O. Sensen (Ed.), *Kant on moral autonomy*. Cambridge University Press, 282–88.
- Oshana, M. (2005). Autonomy and self-identity. In J. Christman & J. Anderson (Eds.), *Autonomy and the challenges to liberalism: New essays* (pp. 77–97). Cambridge University Press.
- Paakkunainen, H. (2018). Doing away with the 'Shmagency' objection to constitutivism. *Manuscript*, 41, 431–480.
- Pinkard, T. (2002). *German philosophers 1760–1860: The legacy of idealism*. Cambridge University Press.
- Pippin, R. (2000). *Hegel's practical philosophy: The realization of freedom*. Cambridge University Press.
- Rawls, J. (1980). Kantian constructivism in moral theory. *Journal of Philosophy*, 77, 515–572.
- Reath, A. (1994). Legialsting the moral law. *Nous*, 28, 435–464.
- Reath, A. (2006). *Agency and autonomy in Kant's moral theory*. Cambridge University Press.
- Reath, A. (2013). Kant's conception of the autonomy of the Will. In O. Sensen (Ed.), *Kant on moral autonomy* (pp. 32–52). Cambridge University Press.
- Reath, A. (2019). What emerged: Autonomy and heteronomy in the groundwork and second critique. In S. Bacin & O. Sensen (Eds.), *The emergence of autonomy in Kant's moral philosophy* (pp. 176–195). Cambridge University Press.
- Rosati, C. (1995). Normativity, and the open question argument. *Nous*, 29, 46–70.
- Rosati, C. (2016). 'Agents and shmagents: An essay on agency and normativity' *Oxford Studies of Metaethics*
- Sandel, M. (1982). *Liberalism and the limits of justice*. Cambridge University Press.
- Schapiro, T. (1999). What is a child. *Ethics*, 109, 715–738.
- Schapiro, T. (2001). Three conceptions of action in moral theory. *Nous*, 35, 93–117.
- Schneewind, J. B. (1998). Natural law, skepticism, and methods of ethics. In P. Guyer (Ed.), *Kant's groundwork of the metaphysics of morals. Critical essays* (pp. 3–25). Rowan & Littlefield.
- Schönecker, D. (2013). Kant's moral intuitionism: The fact of reason and moral predispositions. *Kant-Studien Online*, 1, 1–38.
- Sensen, O. (2013). The moral importance of autonomy. In O. Sensen (Ed.), *Kant on moral autonomy* (pp. 262–281). Cambridge University Press.
- Silverstein, M. (2012). Inescapability and normativity. *Journal of Ethics and Social Philosophy*, 6, 1–26.
- Silverstein, M. (2015). Theshmagency question. *Philosophical Studies*, 172, 1127–1142.

- Smith, M. (1994). *The moral problem*. Wiley-Blackwell.
- Smith, M. (2013). A constitutivist theory of reasons: its promises and parts. *Law Ethics and Philosophy*, 1, 9–30.
- Smith, M. (2015). The magic of constitutivism. *American Philosophy Quarterly*, 52, 187–200.
- Stern, R. (2012). *Understanding moral obligation*. Cambridge University Press.
- Street, S. (2012). Coming to terms with contingency: Human constructivism about practical reason. In J. Lenman & Y. Shemmer (Eds.), *Constructivism in practical philosophy* (pp. 40–59). Oxford University Press.
- Taylor, C. (1989). *Sources of the self*. Harvard University Press.
- Velleman, D. (1989). *Practical reflection*. Princeton University Press.
- Velleman, D. (2000). *The possibility of practical reason*. Oxford University Press.
- Velleman, D. (2009). *How we get along*. Cambridge University Press.
- Walden, K. (2012). Laws of nature, laws of freedom, and the social construction of normativity. *Oxford studies in metaethics*, 7, 37–79.
- Walden, K. (2018). Morality, agency, and other people. *Ergo*, 5, 69–101.
- Wood, A. (2007). *Kantian ethics*. Cambridge University Press.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.