Ontological-Transcendental Defence of Metanormative Realism

Michael Kowalik

Abstract

If there is something (P) that every possible agent is committed to value *above all else*, and certain actions or attitudes either enhance or diminish P, then normative claims about a range of intentional actions can be objectively and non-trivially evaluated. I argue that the degree of existence as an agent depends on the consistency of reflexive-relating with other individuals of the agent-kind: the ontological thesis. I then show that in intending to act on a reason, every agent is rationally committed to value *above all else* being an agent, what consists in exercising the capacity to act and having the freedom to discriminate between more or less valuable actions: the transcendental thesis. Since the degree of possession of this personal but non-contingent good depends on relating to other agents in a special way, certain actions and attitudes may be objectively right or wrong for all agents.

This is a post-peer-review, pre-copyedit version of an article published in Philosophia. The final authenticated version is available online at: http://doi.org/10.1007/s11406-019-00149-6

Introduction

Suppose that some normative claims are not contingent or socially constructed but more or less reflect something fundamental to our existence as individual agents. This would commit us to realism about properties that determine the right thing to do. Taking the notion of realism in a liberal sense, *metanormative realism* could mean that normative truths supervene on real properties that are entirely independent of mental states - the strongest (classical) realist thesis - or, according to Brink (1989, 54), depend on mental states in some general sense but are independent of "contingent *and variable* desires" - the weakest realist thesis. In any case, it may be that only some normative claims satisfy the relevant thesis of realism while others do not.[1] Here I defend an intermediate thesis that the minimal constitutive conditions of agency together with the universal value-commitments associated with reasons for action, despite being in some way dependent on agency and on mental states of agents, have practically normative (but not necessarily moral) implications that are independent of contingent and variable desires, conventions, or beliefs.[2]

The general desideratum of metanormative realism that would satisfy my preferred sense of *practical normativity* is that it should allow for resolution of intersubjective conflict by resorting, in addition to contingent material facts, to reasons that *ought to be* accepted as objective or universally binding by all rational and sufficiently informed agents. In other words, practical metanormative realism must be fit to ground universal normative claims about acting in a particular way.

My argument consists of *ontological* and *transcendental* theses working together. These are presented in two respective parts. First, I develop a formal account of the relationship between the degree of individual existence as an acting subject (henceforth *agent*) and the degree of reflexive kinship with a community of individuals of the same ontological kind. Second, I argue that all agents value *above all else* exercising the capacity to act and their freedom to discriminate between more or less valuable actions. By linking the purely ontological condition of agency with the universal value-commitments associated with reasons for action I conclude that preservation of 'my' fundamental good - the capacity to act and the freedom to discriminate between more or less valuable actions - is conditional on preservation of the same kind of good for *all* agents.

Logic of Coexistence

Working on the premise that agency entails nothing in excess of being an acting subject, I intend to formalise the relevant sense of what it means to be a subject. I begin with a hopefully uncontroversial proposition that x is a subject only if x is an individual (is absolutely identical only to itself)[3] and is conscious of being an individual (is reflexively relating in excess of the abstract reflexivity of being identical to itself)[4]. I do not claim that identity and self-consciousness are *sufficient* conditions of being a subject, or that self-consciousness is *just* reflexive-relating. I do nonetheless commit to the

---

[1] "...the realist need not maintain even that all genuine moral disputes are resolvable. He can maintain that some moral disputes have no uniquely correct answers." (Brink 1984)

[2] My definition of metanormative realism is therefore weaker than the 'robust metanormative realism' defended by Enoch (2010, 414-415).

[3] The law of identity is not only that, at time t, 'x is absolutely identical to itself' ($x=x$), but that 'x is absolutely identical *only* to itself': $\forall m(\exists! x=m)$ or $\exists! x$ for short. Every individual, by virtue of being identifiable, implies either constitutive or contextual uniqueness, which is to say, the quality of being a *one* that is differentiable from every other *one*, "for it is impossible to think of anything if we do not think of one thing; but if this *is* possible, one name might be assigned to this thing" (Aristotle 1984b, 1006b). Cf. "To single x out is to isolate x in experience; to determine or fix upon x in particular by drawing spatio-temporal boundaries and distinguishing it in its environment from other things and unlike kinds..." (Wiggins 2001, 6)

[4] "To be an I, a self, is to have the capacity for reflexive self-reference." (Nozick 1981, 78); "Reflexiveness (...) is the essential condition, within the social process, for the development of mind." (Mead 1934, 134);

well established thesis that subjectivity entails a first-person perspective as well as self-identification.[5] In this sense, to be a subject is to identify as 'I', as 'myself'.

A consistent function expressing subjectivity must not, on pain of circular reasoning, presuppose subjectivity of the individual (or a part of the individual) with respect to which the condition of reflexive-relating is to be satisfied.[6] It follows that *direct* (unary) self-relating of a subject fails to satisfy the relevant condition because it presupposes existence of the same subject. By analogy, a finger cannot point at itself. Conversely, were subjectivity not presupposed by the unary reflexive function it would only express identity of the individual to itself ($x=x$) which holds for any object. It obviously does not express reflexive-relating *in excess* of 'being identical to itself' that the notion of self-consciousness implies.[7]

An alternative refutation of direct reflexive-relating obtains from Wittgenstein's account of self-reference:

> "If, for example, we suppose that the function $F(p)$ could be its own argument, then there would be a proposition $"F(F(p))"$, and in this the outer function F and the inner function F must have different meanings; for the inner has the form $\phi(p)$, the outer the form $\psi(\phi(p))$. Common to both functions is only the letter $"F"$, which by itself signifies nothing."[8]

A function nested within a function is consistent with how we normally think of self-awareness, in that the self is aware of being aware of something. More generally, the function of awareness implies a subject who is internally aware of an object presenting itself or submitting to awareness *as something that appears*. If we let $F(p)$ take the propositional value 'p is identical only to p', then $F(p)$ expresses satisfaction of the law of identity for p, or that p is (or consists of) an individual. $F(F(p))$ then expresses that '$F(p)$ is identical only to $F(p)$', or that $F(p)$ is also an individual. This, according to Wittgenstein and others[9], does not establish a simple identity of individuals p and $F(p)$, because the logical structure of $\psi(\phi(p))$ preserves the distinction between the outer argument/individual and the inner one. This pseudo-reflexive structure resembles the subject 'I' being internally aware of the object 'me'. On this account, the subject can only relate to itself as to an object. But subject and object are, by definition, different logical types, therefore the subject does not relate to *itself* by-itself or in-itself, therefore direct reflexive-relating is false.[10]

If this is correct then the only remaining possibility of satisfying the relevant constitutive condition of subjectivity (with one notable exception to be addressed shortly) is via *indirect*, or mediated, relating of individuals.[11] The law of identity entails that every relation other than the relation of self-identity

---

[5] "...subjectivity consists in experiencing oneself as here and now at every moment in every circumstance." (Peters 2010, 419); "...all subjective experience is self-conscious in the weak sense that there is something it is like for the subject to have that experience. This involves a sense that the experience is the subject's experience" (Flanagan 1993, 194); For similar views see (Kriegel 2003, 106), (Nagel 1974, 436), (Zahavi and Parnas 1998, 689-92), (Zahavi 2005, 119). If the attribution of intrinsic self-consciousness to subjectivity is unacceptable to some, my metanormative argument will not be weakened by substituting 'self-conscious subject' for the term 'subject'.

[6] "...no function can have among its values anything which presupposes the function, for if it had, we could not regard the objects ambiguously denoted by the function as definite until the function was definite, while conversely, as we have just seen, the function cannot be definite until its values are definite." (Whitehead and Russell 1927, 39)

[7] "[Self-identity] is certainly a relation formally or logically speaking, but it also holds trivially, it's trivially true of everything..." (Strawson 2013, 21); "...the subject without relation to himself would be condensed into the identity of the in-itself;" (Sartre 1956, 77)

[8] (Wittgenstein 1933, Par 3.333) I have substituted p for the fx used in the original formulation.

[9] A similar line of reasoning is followed by Tarski (1944) in his paper on the semantic conception of truth, where he resolves the liar paradox by showing that any self-referring sentence consists of object- and meta-languages; and by Whitehead and Russell (1927) who resolve Russell's paradox by means of 'logical types'.

[10] "...the 'I' is not 'me' and cannot become a 'me'." (Mead 1934, 174)

[11] This ontological thesis constitutes the foundation of Discourse Ethics developed by Karl-Otto Apel and Jürgen Habermas: "Subjectivity (...) is itself constituted through intersubjective relations to

involves non-identical terms to relate *to* and *from*. Accordingly, if x is not only identical to itself but also reflexively relates, its relating in excess of identity presupposes another, non-identical term y that has the same relation to x as x has to y. This is possible only if x and y are identical in the restricted sense of belonging to the same kind.[12] Formally, if R is a symmetrical binary relation, S is an asymmetrical and transitive binary relation, and there is y which is not identical to x, then the reflexive relation xRx is equivalent to the conjunction of xSy and ySx. Based on these considerations we can formulate the basic *axiom of subjectivity*:

> x is a subject of **f**-kind ($\mathbf{I_f}\{x\}$) iff x is an individual ($\exists!x$) that relates to itself by relating to a different individual y that relates to itself by relating to x, in terms of properties (**f**) common to them both, or
>
> $\exists!x\ \exists!y\ \exists\mathbf{f}\binom{x=y}{\mathbf{f}}\ [x\neq y \wedge xSy \wedge ySx] \leftrightarrow \mathbf{I_f}\{x, y\}$

By the axiom of subjectivity, individuals can be constituted as subjects only by means of reflexive-relating with other individuals of the same kind, which is both an act of identification *with* the kind - in terms of properties (**f**) with respect to which all members of the kind are identical - and of individual differentiation *within* the kind that does not negate **f**. This is relevant to the present metanormative project insofar as whatever is constitutive of the kind (but not necessarily constitutive of subjectivity) is also constitutive of identities of the individuals that belong to that kind, but not everything that is constitutive of individual identity can be constitutive of the kind.

An implication of the individual x being a subject of **f**-kind is that both x and some other individual y are subjects of **f**-kind, hence $\mathbf{I_f}\{x, y\}$. This precludes the possibility of indirect reflexive-relating of an individual subject by means of its parts, a thesis advanced by Kriegel (2009, 224-8), since then, by virtue of reflexive symmetry, the individual and at least one of its parts would be simultaneously constituted as *different* subjects, therefore not the same individual subject, therefore contradiction.

It may be objected that, in order to facilitate reflexive-relating of x, the relatum y need not be an individual of the same kind but only a mirror, what is evidently true in a world consisting of strangers, families, friends and mirrors, but x would not be a subject nor could it recognise a reflected image as its own if the outside world consisted *only* of mirrors.[13] When I recognise my image in a mirror, or simply think about myself, I already recognise myself as a member of a particular kind that grounds my identity as a subject. Thomas Nagel, in his famous paper "What is it Like to be a Bat?" (1974, 436), observed that for an organism to have "conscious experience *at all* means, basically, that there is something it is like to *be* that organism." The core phenomenological question - What is it like to be 'me'? - can be objectively answered by a conscious being only by identifying with a kind of beings whose members actually are like that being.[14] Our kind, vaguely enough, may be the human-kind (for Kant) or a communication community (for Habermas), but it is also, unavoidably and precisely, the kind whose **f**-properties include the capacity and reasons for action. More on this later.

The basic function of reflexive-relating, as expressed by the axiom of subjectivity, is evidently still too 'flat' to evade circularity (subjectivity of x presupposes subjectivity of y, which in turn presupposes subjectivity of x). This deficiency can be overcome by postulating subjectivity not as a fixed, reciprocally entailed property for all individuals (this would prove too much), but as a matter of

---

others. The individual self will only emerge through the course of social externalization, and can only be *stabilized* within the network of undamaged relations of mutual recognition." (Habermas 2003, 34); "...the self of the practical relation-to-self cannot reassure itself about itself through direct reflection but only via the perspective of others." (Habermas 1992, 186)

[12] Two individuals (a, b) belong to the same kind iff "a has to b the relation of identity as restricted to things that **f**; or, more formally, $(\exists\mathbf{f})\binom{a=b}{\mathbf{f}}$" (Wiggins 2001, 17). Another way, (the same) **f** is a property or part of both a and b, but a is not b.

[13] (Prinz 2012, 54)

[14] A similar thesis is central to George Mead's theory of subjectivity: "...the general conversation which constitutes the process of thinking - is carried on by the individual from the standpoint of the 'generalised other'." (Mead 1934, 155) See also (Vandenberghe 2010)

degree determined asymmetrically by the *kind* of individual relating to other individuals, where perfect reciprocity is only the logical limit of reflexive-relating.

Between the limit scenario of a non-reflexive world consisting only of objects and a maximally reflexive world populated only by 'soulmates', we can plausibly infer degrees of existence of an individual as a subject corresponding to the degrees of reflexive-relating with other individuals of the **f**-kind.[15] It is possible to reformulate the axiom of subjectivity to define the minimum (μ) and the maximum (M) degree of existence of an individual x as a subject $\mathbf{I_f}\{x\}$, expressed as the degree of identity of x to $\mathbf{I_f}\{x\}$, in a world populated by other individuals (n) of the same kind:

$$\exists! x\ [\forall n\ \exists \mathbf{f} \left(\tbinom{x=n}{\mathbf{f}}\right)(x \neq n) \wedge \exists! n(xSn \wedge nSx)] \leftrightarrow \mathbf{Id}(x, \mathbf{I_f}\{x\}) = \mu$$

(the degree of existence of the subject $\mathbf{I_f}\{x\}$ is minimal iff x is an individual that relates to itself by relating to only one other individual n that relates to itself by relating to x, in terms of properties (**f**) common to them both)

$$\exists! x\ [\forall n\ \exists \mathbf{f} \left(\tbinom{x=n}{\mathbf{f}}\right)(x \neq n \wedge xSn \wedge nSx)] \leftrightarrow \mathbf{Id}(x, \mathbf{I_f}\{x\}) = \mathrm{M}$$

(the degree of existence of the subject $\mathbf{I_f}\{x\}$ is maximal iff x is an individual that relates to itself by relating to all other individuals n that relate to themselves by relating to x, in terms of properties (**f**) common to them all)

On this account, the individual x exists as a subject (identifies as 'myself') only as much as x reflexively-relates with other individuals of the kind that grounds x's identity as a subject. Consequently, were x to performatively contradict any constitutive property of its kind in another individual it would negate a unique pathway of reflexivity and the associated means of self-constitution, and thus its degree of existence as a subject would be diminished. What matters here is not the fact of existence or inexistence of another individual, but the act of negation or affirmation[16] of an identity that is already performatively constituted via reflexive-relating of an individual with *other* individuals, already actualised by others 'for me' as someone who shares capacities, commitments or other properties of my identity-grounding kind.[17]

If our existence as subjects is indeed determined by how we relate to other individuals of the same kind, and if our relations with respect to other members of the kind are inconsistent, then our relationship with the kind is also inconsistent. For example, in case we related to all individuals of our kind as subjects we would be unambiguously constituted as subjects of *that* kind, but in case we related to some subjects as if they were not subjects, or as if they were subjects of a different kind, then our subjectivity would be realised to a lower degree (or corrupted to a degree) on account of inconsistent reflexivity with our ontological kind. In other words, my 'reflection' would not be true to the kind of an individual that I objectively am, and thus I would lack what Korsgaard (1996, 102) calls *integrity*. The essential point is that the degree of existence as a subject of a particular kind is equal to the degree of performatively affirmed identity (or *reflexive kinship*) with the kind as collectively constituted by all instances of reflexive-relating among co-existent individuals of that kind. If this is correct then the axiom of subjectivity amounts to a *universal law*. It does not of itself tell us explicitly

---

[15] If my interpretation is correct, this line of reasoning is compatible with Badiou's axioms of existence: "the degree of existence of *a* is equal to its degree of identity to *b*" (2009, 257), and "the *common* of degrees of existence is equal to the degree of identity of the terms in question" (2009, 356).

[16] Performative affirmation of a fact, hereafter *affirmation*, consists of actions or attitudes that require, or could not possibly be intended without, the agent's acceptance of that fact: the relevant fact is presupposed by the affirmative act or attitude.

[17] According to Korsgaard (2009, 25), "an action that is less successful at constituting its agent is to that extent less of an action. So on this conception, 'action' is an idea that admits of degrees." The lowest degree of agency is that of a "tyrannized soul" who is no longer an agent but "a mere force of nature, an object, a thing" (2009, 173). Korsgaard deals with the question of action as "an idea that admits of degrees" in a broadly qualitative way and does not develop a general theory of how the degrees of existence could be quantified.

how we ought to act but formalises a causal connection between how we relate to others and the degree of our own existence.

For analytical clarity of the ontological part of my argument I have so far avoided making claims about what is constitutive of agency or intentional action, even though the kind of relating that I am concerned with here is applicable only to subjects who act.[18] In the following section I will examine metanormative consequences of the ontological thesis in light of value-commitments presupposed by reasons for action.

An Ontological-Transcendental Account of Normative Truth

I have argued that existence of a subject entails co-existence with other individuals of the same ontological kind. In other words, 'my' existence implies that there is a kind to which 'I' belong and that belonging to 'my' kind is constitutive of 'my' subjectivity, but these constitutive conditions do not yet account for practical normativity.

According to Hume's Law, it is not possible to obtain a prescriptive 'ought' from a descriptive 'is'.[19] Something more must be given to consistently and uncontroversially make the transition from ontological truth to normative truth. It is not even sufficient to provide evidence of intrinsic goodness of something, because the notion of 'good' does not necessarily entail an 'ought' either. One can plausibly know what is objectively good and yet have no overriding reason or motivation to pursue it.[20] Consequently, metaethical or moral realism does not necessarily entail metanormative realism, or vice versa, and I do not wish to takes sides on this question here. Henceforth, I employ the term 'good' exclusively for private or subjective good (something valued by 'me') without committing to *reasons internalism* or *motivational internalism* about the good in the objective sense: that the *objective* good is intrinsically reasons-giving or motivating.[21]

In order to demonstrate normative truth that is independent of social conventions and individual convictions about value, at least in the practical sense that interests me here, it is necessary to show that: a) some principle, state, relation or thing (*p*) is good or valuable to all possible agents in a given world, b) there are reasons to pursue *p* that would have to be overridingly motivating for any rational[22] and sufficiently informed agent, c) for all possible agents, certain actions or attitudes with respect to any other agent amount to either enhancing or diminishing *p*.

An immediate implication of the axiom of subjectivity, given that agents are subjects of a kind whose members have the capacity to act, is that my existence as an agent depends on reflexive-relating with other agents. It follows that if I value my own existence as an agent I then ought, on pain of diminished self-existence, but in consideration of any conflicting factors that are also critical to preservation of my existence, aim to preserve the agency of others and whatever is constitutive of the

---

[18] I take *action* to mean possessing causal features that are not fully determined prior to the relevant act or intention being realised.
[19] (Hume 1965, 469-70)
[20] "Whether the recognition of moral facts provides reasons for action depends upon whether the agent has reasons to do what morality requires." (Brink 1984, 114); "An account that takes agents to be oriented toward the good will only be adequate if the account of the good is such that it makes sense as a constitutive aim of agency (or better, rational agency). And it will only make sense as a constitutive aim of agency if it is clear how being good (or good for someone), so conceived, could answer the question of what to do, thereby providing the agent with reasons for action." (Rosati 2003)
[21] (Brink 1984, 113)
[22] Why be rational? What is so special about rationality to motivate us to act rationally? I understand (minimal) *rationality* as adherence to the law of identity and the law of non-contradiction - the necessary conditions of *sense*. We have to act rationally in order to reliably get what we want out of action, and aiming to get what one wants out of action is just what 'being an agent' means. A different argument in support of the thesis that *one ought to be rational* was recently presented by Lord (2017).

agent-kind.[23] But what if I were to assert and believe that I do not (sufficiently) value my own existence? In order to properly address this doubt will require some preparation.

According to the classical view, intentional (purposive) action is paradigmatically directed towards preferences about acting in a particular way and therefore towards the contemporaneous good of the agent: the *sense* of preference is a positive value-commitment.[24] It is irrelevant whether intentions and the associated preferences are realised exactly as intended (or not at all): having first-person awareness that one is trying to do something is already acting intentionally, with a purpose, and an affirmation of preference for what is intended.[25] Without preferring an action on reasons taken to be 'my' reasons it would be a mystery in what sense intention differs from compulsion, and 'my' action from something merely happening with me. Even if it made sense to intend an action that were believed to be bad or of no value for the agent, as the opponents of the classical view typically argue, in *every* case of intending (be it good or bad) there still is a second-order preference for having the first-order intention to act in a particular way satisfied rather than not satisfied. In intending to ϕ we affirm that it is preferable (or more valuable) to ϕ than not to ϕ, where ϕ stands for 'exercise the capacity <to act in a particular way>'. Since ϕ amounts to '*now* being an agent', the commitment to value ϕ is a value-commitment to '*now* being an agent'. I will call this *'value internalism about reasons' of the second-order*.

Value internalism of the first-order is arguably too 'thin' to generate any normative content. If all agent-reasons are good to the agent at the time of intending and *only* these reasons count towards underwriting the value of actions, then there is no logical space remaining for objectively normative reasons.[26] Any normative principle based on a flatly positive valuation of actions but declaring some actions as being of negative or no value would be contradicting a condition of its own authority. I will therefore proceed without committing to value internalism about reasons of the first-order but only to value internalism of the second-order, which does not depend on the value of specific actions or reasons. I will later consider potential consequences for the present metanormative project in case value internalism of the second-order were shown to be false or normatively trivial.

If value internalism about reasons of the second-order is true and normatively non-trivial, then every action, insofar as action is intentional, already affirms that exercising my capacity to act in a particular way is unequivocally valuable to me. In addition, if the *sense* (or the constitutive aim) of intentional action is to satisfy my reasons to act, and certain actions are more valuable than others for satisfying my reasons to act, then having the freedom to discriminate between more or less valuable actions is a constitutive condition of what I am already committed to value and may be valuable also.[27]

---

[23] Brink seems to approve of the kind of normative logic I am pursuing here: "Everyone has reason to promote his own well-being, and everyone has reason to promote the well-being of others at least to the extent that his own well-being is tied up with theirs. Presumably, any plausible theory of human needs, wants, and capacities will show that the satisfaction of these desiderata for any given individual will depend to a large extent on the well-being of others." (Brink 1984)

[24] This account of intentional action, which may be characterised as *value internalism about reasons*, is defended by Joseph Raz: "...the classical approach, it may be called, can be characterized as holding that the central type of human action is intentional action; that intentional action is action for a reason; and that reasons are facts in virtue of which those actions are good in some respect and to some degree." (Raz 1999, 22); "...having intentions involves belief in the value of what they intend" (Raz 2015, 1); see also (Raz 2016) (Schueler 2003, 104) (Gewirth 1978, 48-53) (Nagel 1970, 35) (Anscombe 1957, 76) (Kant 2015, 5:58-5:60) (Aristotle 1984a, 431). For arguments opposing the classical view see (Stocker 1979) (Velleman 1992) (Setiya 2003) (Setiya 2007). For an overview of the debate see (Orsi 2015).

[25] Much of Setiya's (2007, 26) argument against the classical view hinges on the premise that 'trying to do something' does not constitute action. I am not clear what Setiya means by 'trying to do' that isn't already *acting*. Aren't we always only 'trying to do' something, with no guarantee of success?

[26] On this point I agree with Setiya: "When an agent acts on a reason, he takes it *as* a reason, but that means he takes it as *his* reason, not that he takes it to be a *good* reason on which to act." (Setiya 2003, 380)

[27] Cf. "...if one cares about one's capacity for agency, then one should (by virtue of consistency) care about how one exercises it." (Arruda 2016).

It seems logical that if *b* is a constitutive condition of *a* and I am committed to *a*, then I am also committed to *b*, otherwise my commitment to *a* would be incomplete and possibly inconsistent.[28] For example, if I am committed to paying taxes but not to income declaration and tax assessment procedures, the tax authorities may contend that I am not in fact committed to paying taxes (on account of my non-compliance with the necessary conditions of paying taxes). Nonetheless, this evaluative principle does not always hold true for complex conditionals.[29] The relevant exception is the case of being committed to goodness of *a* and badness of *b*, where *c* is a constitutive condition of both *a* and *b*. To put it differently, attribution of value to a constitutive condition on the basis of what is constituted is frustrated if there are some instances of what the relevant condition is constitutive-of that are mutually inconsistent in value. For example, firearm possession is a constitutive condition (and arguably also a *sufficient* condition for at least some agents) of firearm violence, which may be regarded as good when it is defensive, proportionate and necessary to protect lives or property, *and* regarded as bad when it is offensive, excessive, unnecessary or criminal in intent. In this case it is not possible to generalise about the value of firearm possession solely on account of its positively valued uses or aims.[30]

The two conflicting examples demonstrate that it is not logically necessary for a necessary condition of value to be valuable: the value of a condition of value depends on properties of the entire constitutive structure with respect to which evaluation is made. In the case of the constitutive conditions of agency or action, inconsistency does not frustrate evaluation since it is not logically possible for an agent to intentionally act against and thus be negatively committed to one's own contemporaneous preference for acting intentionally, or to exercise the freedom to discriminate between more or less valuable actions without already unequivocally affirming the preference for acting in a particular way - what amounts to having a preference for '*now* being an agent'. Commitment to the value of '*now* being an agent' is a necessary condition of *all* contemporaneous value-commitments.[31]

A committed sceptic could still argue that even though she indeed values her own existence, her capacity to act in a particular way, and her freedom to discriminate between more or less valuable actions, these conditions are only of instrumental value and she prefers non-existence as the ultimate or final good. I argue that this objection is inconsistent on at least two levels: being *intentionally* presented from the position of existence it is a performative affirmation of the value of agential existence only, even if this value is expressly negated,[32] and her preference to pursue non-existence is an affirmation of (subjectively construed) self-interest in satisfying *her* preference, and therefore a second-order value-commitment to the intending self. Self-interest is expressed in every act of the self, including the act of self-nihilation or self-sacrifice, because it asserts authority of the self over the conditions of its own existence, even if the motivating reasons are self-negating.[33] Another way, if I would value some aim (X) as much as my agency I would be committed to potentially renouncing my

---

[28] See Gewirth (1978, 48) and Arruda (2016).

[29] "...infection makes penicillin valuable, but infection isn't therefore valuable, much less intrinsically valuable." (Markovits 2014, 105); Cf. "...it is not logically necessary for the condition of goodness to be good itself, either conditionally or unconditionally." (Rabinowicz and Rønnow-Rasmussen 2000, 48)

[30] For a similar argument see Kerstein (2001, 37).

[31] Cf. "I must see myself as having unconditional value - as being an end in myself and the condition of the value of my chosen ends - in virtue of my capacity to bestow worth on my ends by rationally choosing them." (Markovits 2014, 103)

[32] Enoch (2011) argues that an attack on the sceptic on the basis of an unavoidable performative contradiction does not alone succeed in refuting the sceptical challenge. The essence of his argument is that we should not mistake finding flaws with the sceptic for vindicating our own position, and so we still need a substantive answer to the sceptical challenge. This may be true in some cases but does not amount to a general principle. If the sceptical challenge involves a performative contradiction then the fault is not just with the sceptic but with the sceptical challenge as well, potentially rendering it false. For example, if the sceptic argues by ϕ-ing that not-ϕ-ing then the argument is self-negating.

[33] "If you are already an agent, then any authoritative recommendation you receive - even a recommendation to cease being an agent - owes its authority to agency's constitutive norm. And so the force of such a recommendation could never undermine the authority of that norm." (Silverstein 2015); Cf. (Korsgaard 1996, 160-164)

agency in case of conflict between the two values. Since things have value for me only in virtue of my capacity to value, therefore agency, I would then be also committed to potentially negating the value that X has for me, therefore contradiction. It seems I can value things and act in favour of them only insofar as I value 'my capacity to value things' more than I value the things themselves.

If my argument in this section is correct, then 'my' existence, 'my' capacity to act and 'my' freedom to discriminate between more or less valuable actions are not only constitutive conditions of agency but also universally affirmed, non-contingent value-commitments intrinsic to agency.[34] In case my argument were shown to be defective, possibly because value internalism about reasons of the second-order were shown to be defective, I would then fail to demonstrate *realism* of the resulting metanormative claims, but (*ceteris paribus*) these would still apply contingently: only to those agents who happen to value their own existence as agents. Conversely, if value internalism about reasons of the second-order is true but agents are not *necessarily* committed to value being agents *above all else* and may rationally hold something other than agency as the highest value, the commitment to value being an agent would still be normative in a weaker sense, insofar as we are rationally committed to take it into account when serving other aims.

This brings me to Enoch's (2006) claim that in order to substantiate normative truth "it is not sufficient to show that some aims or motives or capacities are constitutive of agency" or that agency is unavoidable, but also that "we have reasons to be agents". I agree, but this is a moot point since every reason to act is already a reason to be an agent if the *sense* of 'to act' is to 'be an agent'. The relevant question here is not, as Enoch (2011) puts it, whether playing chess suffices for having a reason to checkmate, but whether having a reason to play chess amounts to having a reason to be a chess-player of sorts. Similarly, having a reason to terminate one's own agency amounts to having a reason to be an agent who 'terminates one's own agency', rather than not to be an agent at all.[35] I have argued that this kind of reason is self-negating, but even if it were rational 'to act in order not to act' (or 'to exist in order to not-exist') and thus cease being an agent, a non-agent is not a problem for metaethical or metanormative theory because moral value and normativity extend exclusively to intentional actions and therefore to agents; not to events in general.

I do agree with Enoch (2011) insofar as he criticises inconsistency of "the move from someone caring about something, immediately to it being the case that she should care about it, or at least that she has a reason to so care"; I disagree that it is necessary to show that one *should* care about being an agent but only that some value-commitments are intrinsic to agency. I claim that universal, practical normative consequences follow from this premise as long as every agent is ontologically dependent on reflexive-relating with other agents.

By showing that certain value-commitments are intrinsic to agency I have satisfied my first condition of practical metanormative realism: that *p* is good or valuable to every possible agent. Since *p* corresponds to a fundamental value-commitment of every agent, without which no other value, interest or contingent reason for action could be pursued (or rejected), the second condition is satisfied by implication: there are non-contingent reasons to pursue *p* that would have to be overridingly motivating for any rational and sufficiently informed agent. It still remains to be shown that for all possible agents, certain actions or attitudes with respect to *any* other agent amount to either enhancing or diminishing *p*. In order to accomplish this final step I appeal to the ontological part of my argument.

I have argued that the degree of existence of an individual as an agent is determined by the degree of consistency of reflexive-relating with other individuals of the agent-kind: the **f**-kind that ground the identity of every agent. This I called the ontological thesis. If every agent is rationally committed to value *above all else* being an agent, what consists in exercising the capacity to act and having the freedom to discriminate between more or less valuable actions, "then we have reason not only to preserve it, but also to manifest it whenever the opportunity presents itself" (Kolodny 2005, 545). We are thus rationally committed not just to sufficiency but to maximisation of the degree of existence as

[34] These 'properties' of agency are consistent with Korsgaard's (1996, 101) definition of practical identity.

[35] Cf. "Choosing not to act makes not acting a kind of action, makes it something that you do." (Korsgaard 2009, 1)

an agent whenever we intend to act. This I called the transcendental thesis. The logical consequence of the two theses working together is that every rational and sufficiently informed agent *ought* to act (on account of non-contingent reasons) in the best way to preserve the constitutive conditions of agency for all agents, including oneself. Preservation of my good - the capacity to act and the freedom to discriminate between more or less valuable actions - is conditional on preservation of the same kind of good for all agents, even if there are contingent factors that are immediately more critical to preservation of my existence. There is logical space here to accommodate the case of defensive violence necessary for self-preservation as an individual, for preservation of 'my' capacity to act and of 'my' freedom to discriminate between actions. While *direct* self-preservation takes absolute precedence in the logical order of value-commitments, it is nonetheless practically on par with preservation of agential existence of others. Another way, we may have good contingent reasons to kill in order to preserve our own existence in the absolute sense, but we simultaneously have a non-contingent reason to avoid killing in order to preserve the degree of our existence as agents. It is unclear to me whether or how the scope of agency could be intentionally expanded beyond the present **f**-properties of my agent-kind.

According to the ontological thesis, every action or attitude that would intentionally diminish the constitutive conditions of agency for any other agent would also diminish the degree of existence and the associated capacity to act of the offending agent. Such actions obviously do not result in the offending agent's imminent disappearance from the world, nor is it evident that abusive and selfish agents are materially less real for me than agents who are caring and supportive. Returning to the formula defining the minimum and the maximum degree of existence in terms of identity of the individual x to the subject $\mathbf{I}_{\mathbf{f}}\{x\}$, or $\mathbf{Id}(x, \mathbf{I}_{\mathbf{f}}\{x\})$, an individual who would diminish any constitutive condition of agency for any other agent is, or is in the process of becoming, necessarily less real *for himself*. It follows that any individual with a diminished degree of existence as an agent does not partly disappear from the world as a material being (an object) but may experience detachment, demoralisation, depersonalisation or identity fragmentation (as a subject).[36] This may in turn correlate with a low degree of reflexive-relating and, consequently, with a diminished capacity to act intentionally or to discriminate between more or less valuable actions. This kind of consequence for acting against the private good of others is psychologically plausible, but evaluation of the present project from the point of view of psychology is beyond the scope of this paper.

Conclusion

The present account of practical metanormative realism consists of two complementary theses, ontological and transcendental. First, I have formulated the axiom of subjectivity which provides that a measure of reflexive kinship of individuals with their ontological kind is also a measure of their degrees of existence as agents: the ontological thesis. An agent has the maximum degree of existence - is maximally real - if the agent respects the constitutive conditions of agency of *all* individuals of the same ontological kind. There is no possibility of exceeding this logical limit for a given scope of common properties that characterise the kind. Conversely, if an agent relates in a way that respects the constitutive conditions of agency of only *some* individuals of the agent-kind then the agent is (or is bound to progressively become) to a degree unreal.

I have subsequently argued that in acting or in intending to act on a reason, every possible agent is rationally committed to value *above all else* being an agent, what consists in exercising the capacity to act and having the freedom to discriminate between more or less valuable actions, and to maximise the degree of existence as an agent: the transcendental thesis. According to the ontological thesis, realisation of this non-contingent self-interest depends on how we relate to *all* other agents, irrespective of contingent and variable desires, conventions or beliefs, therefore certain actions and attitudes may be objectively wrong (or right) for all agents.

Whether this account of practical metanormative realism also satisfies the desiderata of metaethical realism will depend on how moral value is understood. If moral value is defined as a special kind of

[36] I may be vindicating here Korsgaard's intuitions about consequences of being a 'tyrant'. (2009, 173)

good that is perhaps not intrinsically reasons-giving or motivating then the present argument does not substantiate metaethical realism but only metanormative realism.

On a more speculative note, the present argument may have implications for the concept of *justice*. If some objective wrongs have inescapable metaphysical consequences, expressed as changes in the degree of existence as an agent, then, in a sense, our world embodies justice; metaphysical justice. The inhuman things we do dehumanise us, turn us into animals and ultimately to stone.

References

Anscombe, G.E.M. *Intention.* Cambridge: Harvard University Press, 1957.

Aristotle. *The Complete Works of Aristotle (Volume I).* Princeton: Princeton University Press, 1984a.

—. *The Complete Works of Aristotle (Volume II).* Princeton: Princeton University Press, 1984b.

Arruda, Caroline T. "Why Care about Being an Agent?" *Australasian Journal of Philosophy*, 2016.

Badiou, Alain. *Logics of Worlds: Being and Event II.* London: Continuum, 2009.

Brink, David O. *Moral Realism and the Foundations of Ethics.* Cambridge: Cambridge University Press, 1989.

Brink, David O. "Moral Realism and the Sceptical Arguments from Disagreement and Queerness." *Australasian Journal of Philosophy*, 1984: 111-125.

Enoch, David. "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action." *The Philosophical Review*, 2006: 169-198.

Enoch, David. "Shmagency Revisited." In *New Waves in Metaethics*, by Michael Brady, 208-233. Palgrave Macmilan UK, 2011.

Enoch, David. "The Epistemological Challenge to Metanormative Realism." *Philosophical Studies*, 2010: 413-438.

Flanagan, Owen. *Consciousness Reconsidered.* Cambridge: MIT Press, 1993.

Gewirth, Alan. *Reason and Morality.* Chicago: The University of Chicago Press, 1978.

Habermas, Jürgen. *Postmetaphysical Thinking.* Cambridge: MIT Press, 1992.

—. *The Future of Human Nature.* Cambridge: Polity Press, 2003.

Hume, David. *A Treatise of Human Nature.* Oxford: Oxford University Press, 1965.

Kant, Immanuel. *Critique of Practical Reason.* Cambridge: Cambridge University Press, 2015.

Kerstein, Samuel J. "Korsgaard's Kantian Arguments for the Value of Humanity." *Canadian Journal of Philosophy*, 2001: 23-52.

Kolodny, Niko. "Why Be Rational?" *Mind*, 2005: 509-563.

Korsgaard, Christine M. *Self-Constitution: Agency, Identity, and Integrity.* Oxford: Oxford University Press, 2009.

—. *The Sources of Normativity.* Cambridge: Cambridge University Press, 1996.

Kriegel, Uriah. "Consciousness as Intransitive Self-Consciousness: Two Views and an Argument." *Canadian Journal of Philosophy*, 2003: 103-132.

—. Subjective Consciousness: A Self-Representational Theory. Oxford: Oxford University Press, 2009.

Lord, Errol. "What You're Rationally Required to Do and What You Ought to Do (Are the Same Thing!)." *Mind*, 2017: (forthcoming).

Markovits, Julia. *Moral Reason.* Oxford: Oxford University Press, 2014.

Mead, George H. *Mind, Self, and Society.* Chicago and London: The University of Chicago Press, 1934.

Nagel, Thomas. *The Possibility of Altruism.* Oxford: Clarendon Press, 1970.

Nagel, Thomas. “What Is It Like to Be a Bat?” *The Philosophical Review*, 1974: 435-450.

Nozick, Robert. *Philosophical Explanations.* Cambridge: Harvard University Press, 1981.

Orsi, Francesco. “The Guise of the Good.” *Philosophy Compass*, 2015: 714-724.

Peters, Frederic. “Consciousness as Recursive, Spatiotemporal Self-location.” *Psychological Research*, 2010: 407-421.

Prinz, Wolfgang. Open Minds: The Social Making of Agency and Intentionality. Cambridge: The MIT Press, 2012.

Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen. “A Distinction in Value: Intrinsic and For Its Own Sake.” *Proceedings of the Aristotelian Society*, 2000: 33-51.

Raz, Joseph. *Engaging Rerason: On the Theory of Value and Action.* Oxford: Oxford University Press, 1999.

Raz, Joseph. *Intention and Value.* Public Law & Legal Theory Working Paper, New York: Columbia Law School, 2015.

Raz, Joseph. “The Guise of the Bad.” *Journal of Ethics & Social Philosophy*, 2016.

Rosati, Connie S. “Agency and the Open Question Argument.” *Ethics*, 2003: 490-527.

Sartre, Jean-Paul. *Being and Nothingness.* New York: Philosophical Library, 1956.

Schueler, G. F. Reasons and Purposes: Human Rationality and the Teleological Explanation of Action. Oxford: Oxford University Press, 2003.

Setiya, Kieran. “Explaining Action.” *The Philosophical Review*, 2003: 339-393.

—. *Reasons without Rationalism.* Princeton and Oxford: Princeton University Press, 2007.

Silverstein, Matthew. “The Shmagency Question.” *Philosophical Studies*, 2015: 1127-1142.

Stocker, Michael. “Desiring the Bad: An Essay in Moral Psychology.” *The Journal of Philosophy*, 1979: 738-53.

Strawson, Galen. “'Self-intimation'.” *Phenomenology and the Cognitive Sciences*, 2013: 1-31.

Tarski, Alfred. “The Semantic Conception of Truth.” *Philosophy and Phenomenological Research*, 1944: 341-376.

Vandenberghe, Frederic. “Pragmatist and Hermeneutic Reflections on the Internal Conversations that We Are.” In *Conversations About Reflexivity*, by Margaret S. Archer, 55-74. New York: Routledge, 2010.

Velleman, J.D. “The Guise of the Good.” *Nous*, 1992: 3-26.

Whitehead, Alfred North, and Bertrand Russell. *Principia Mathematica (Volume I).* Cambridge: Cambridge University Press, 1927.

Wiggins, David. *Sameness and Substance Renewed.* Cambridge: Cambridge University Press, 2001.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus.* London: Routledge & Kegan Paul, 1933.

Zahavi, Dan. Subjectivity and Selfhood: Investigating the First-Person Perspective. Cambridge: The MIT Press, 2005.

Zahavi, Dan, and Josef Parnas. “Phenomenal Consciousness and Self-Awareness: a Phenomenological Critique of Representational Theory.” *Journal of Consciousness Studies*, 1998: 687-705.