Spontaneity without rationality: a Kantian approach to self-consciousness and perceptual content

To appear in *Perceptual Knowledge and Self-Awareness*, edited by Andrea Giananti, Johannes Roessler, and Gianfranco Soldati (Oxford University Press)

Final version, January 25, 2023

Hannah Ginsborg, U.C. Berkeley

A number of philosophers, inspired by Kant's view of self-consciousness as consciousness of spontaneity, have understood self-consciousness as requiring rational agency with respect to one's beliefs and desires. This leads to a demanding conception of self-consciousness, which excludes the ascription of self-consciousness to children who are too young to appreciate reasons. I offer an interpretation of the role of spontaneity in perceptual experience which suggests that there can be spontaneity, in Kant's sense, without rational agency. The consciousness of spontaneity, on my reading of Kant is the consciousness of normative constraint in one's thinking, and, as I argue through consideration of sorting behavior in small children, one can be conscious of normative constraint without recognizing reasons. This leads to a less demanding, but still recognizably Kantian, conception of self-consciousness, one which accommodates the idea that small children can be self-conscious.

I. Introduction

There is a tradition of philosophical thought about the self, largely stemming from Kant, that seeks to explain self-consciousness and self-knowledge in terms of the consciousness of "activity" or "spontaneity" in one's thinking or behavior. On this Kantian way of thinking, self-consciousness —more specifically the distinctive form of self-consciousness reflected in the use of the first person pronoun—is possible for us because we are active in our thinking and in our effects on the world, as opposed to our thoughts and behavior resulting from external influences in accordance with causal laws. Being active or spontaneous is understood, in this tradition, as intrinsically involving the consciousness of spontaneity. To the extent that we are active thinkers and agents, we are not merely passive observers of our own activity, but are conscious of it simply by virtue of engaging in it. This consciousness both answers the question of what self-consciousness consists in, and accounts for the possibility of self-knowledge,

understood roughly as the privileged first-personal knowledge we have of our own mental states and intentional actions in contrast to those of others.

Recent philosophers in this tradition have typically understood the relevant activity as the exercise of rationality, understood as the capacity to be consciously guided by reasons, or, as it is sometimes put, to be responsive to reasons as such.[1] As a result, they have understood self-consciousness and self-knowledge to be intimately associated with the recognition of one's beliefs and other attitudes as rationally constrained. Richard Moran (2001), to take a central example, explains self-consciousness and self-knowledge in terms of our rational agency with respect to our beliefs and desires. I arrive at first-personal knowledge of what I believe not by observing my own beliefs as a spectator, but by addressing the practical, deliberative question of what *to* believe, where that in turn is a matter of considering what I have *reason* to believe. The special authority of self-knowledge derives, not from my occupying an especially privileged viewpoint on my beliefs and other attitudes, but rather from my distinctive role as a rational agent with respect to my beliefs. Relatedly, I am conscious of myself as a subject, not by being presented to myself through introspection in a special quasi-perceptual way, but in virtue of my capacity to recognize my attitudes as answerable to reasons. Moran's approach to self-consciousness is developed, and its Kantian connections made explicit, by Matthew Boyle (2009). Kant's view that the representation "I think" amounts to the consciousness of spontaneity in my thinking corresponds, as Boyle sees it, to Moran's view that self-consciousness depends on my capacity for rational deliberation. Self-consciousness, then, for both Moran and Boyle, is explained by, and depends

[1] The references to "conscious" guidance and to responsiveness to reasons "as such" are intended to distinguish the relevant notion of rationality from a less demanding notion of rationality or responsiveness to reasons which might be ascribed, say, to a gazelle fleeing a lion. The presence of the lion might be regarded as a reason for the gazelle to flee, and the gazelle's fleeing as a response to that reason, but, unlike a typical human being in the presence of a lion, the gazelle does not appreciate that the presence of the lion is a reason for her to flee. For this distinction, see e.g. McDowell 2009, 128-129; Wallace 2015, 261-262; Korsgaard 2018, ch. 3. References to rationality throughout this paper should be understood in the more demanding sense.

on, a person's being able to make up her mind about what to believe and desire through reflection on reasons.[2]

Although Moran and Boyle make a powerful case for understanding self-consciousness in terms of rational agency, we might wonder whether their view sets the bar for self-consciousness too high. Several philosophers have offered accounts of self-consciousness that are considerably less demanding with respect to the required capacities. José Bermúdez (1998), for example, drawing on the fact that external senses like vision provide information not only about outer objects but also about our location and the opportunities for action that the objects make available, takes our external sensory capacities to afford a primitive form of self-consciousness that is available from birth and shared with many non-human animals. For parallel reasons, he takes there to be an equally primitive form of self-consciousness provided by our proprioceptive sensory capacities, which inform us about our posture and the position of our limbs. A third, and somewhat more sophisticated form of self-consciousness is provided, according to Bermúdez, by our capacities to represent an objective environment distinct from our experiences and to navigate a spatial path through that environment: capacities that we acquire in infancy and that we share with animals such as rats and dogs. And a fourth kind of self-consciousness, still designated as "primitive" because of its nonconceptual character, is yielded by capacities for social interaction that emerge in human children at around 9 months of age. These capacities are manifested in behavior such as initiating joint attention, taking turns in games and proto-communicative interactions, and social referencing (looking towards adults for cues as to how to behave in an unfamiliar situation; the form of self-consciousness they make possible goes beyond bodily self-awareness to awareness of oneself as one psychological subject among others. At least some of these forms of awareness have a *prima facie* claim

[2] Rödl (2007) adopts a similar approach, although he rejects Moran's characterization of the stance of rational agency by way of contrast with a stance of causal explanation (52n44). Other philosophers who see a close connection between self-consciousness (or self-knowledge) and rational agency include Bilgrami (2006), Burge (1996), Korsgaard (1996, 92ff; 2009) and O'Brien (2007).

to be labelled self-consciousness,[3] but we clearly do not need not appeal to capacities for rational deliberation in order to account for them.

Defenders of the rational agency view can reply that, whether or not we regard the forms of consciousness identified by Bermúdez as self-consciousness in some the sense, none of them is the kind of self-consciousness they are interested in. Rather, their focusis whatever form of self- consciousness is expressed through the competent use of "I". But their view might still seem too demanding, even as an account of this more sophisticated form of self-consciousness. For there is evidence that children can think about themselves, in the distinctive way reflected in the use of "I", before they are capable of rational deliberation. Although it is not clear at what age children can recognize reasons for their beliefs and desires, it seems unlikely that they can do so before being able to pass the "false belief test" (Wimmer and Perner 1983), which they typically do around age four. But children start using first-person prounouns at around 18 months, and at around two years old they can do so in complex sentences whose use suggests at least some grasp of the first person (for example the 25-month old who says "I'm sad because I popped it [balloon]") (Bretherton et al 1981). The acquisition of first-person pronouns is also correlated with the emergence of a variety of abilities often taken to indicate the presence of a conscious representation of the self. These include the capacities to recognize oneself in a mirror, to engage in pretend play with objects (e.g. "drinking" imaginary milk from a cup), and to feel so-called secondary emotions like embarrassment and empathy (Lewis and Ramsey 1994, Lewis et al 1989; for a useful discussion of some of the relevant data see Musholt 2015). Even in light of recent studies suggesting both that children can already give reasons for their beliefs at three years of age (Köymen and Tomasello 2020), and that the standard false belief test may over-estimate the age at which children can appreciate the possibility of false beliefs (Rubio-Fernandez and Geurts, 2013), it seems far-fetched to suppose that, say, two-year-olds are capable of rational deliberation. Yet it seems implausible to deny them self-

[3] Musholt (2015) argues that the first two are not genuine forms of self-consciousness (ch. 3); she accepts that the fourth constitutes self-consciousness at a low level (ch 6.).

consciousness in something like the sense associated with the use of "I" — or at least in a sense which goes beyond the primitive forms of self-consciousness identified by Bermúdez.

If we reject the rational agency approach as too demanding, it might seem that we have to give up the Kantian idea of self-consciousness as the consciousness of spontaneity. But I propose in this paper that this is not the case. For, as I argue in what follows, the idea of spontaneity is broader than that of rationality. We can be spontaneous in Kant's sense, and thereby conscious of our spontaneity, without the capacity for rational deliberation or for recognizing reasons. My argument depends on driving a wedge between the idea of *normative* constraint, and the idea—on my view, more specific—of *rational* constraint. It is possible to be conscious of what one does as normatively governed, without being conscious of it as something that can be endorsed or rejected on the basis of reasons. And this is a kind of consciousness, I argue, that we can ascribe to children of around the age at which, according to developmental researchers, a conscious representation of the self becomes available. As part of my argument for Kantian spontaneity without rationality, I focus on the role that Kant ascribes to spontaneity in perceptual experience. I articulate a kind of normativity that, as I read Kant, can be recognized in perceptual experience and that cannot be made out in terms of accordance with reasons. I then go on to argue, on the basis of further findings in developmental psychology, that children in roughly the 15-30 month age range—too young to appreciate reasons for beliefs and other attitudes—recognize this normativity, and therefore can be conscious of spontaneity in Kant's sense. The paper is structured as follows. In the next section (section II), I offer a brief sketch of Kant's view of self-consciousness, aiming both to show what is attractive about the Kantian approach to self-consciousness, and to clarify the notion of spontaneity. In section III I describe the account of self-consciousness offered by Moran and Boyle, raising a concern about how their identification of spontaneity with rationality allows us to do justice to the role Kant ascribes to spontaneity in perceptual experience. For this concern to carry weight, we need to make philosophical sense of the idea that perceptual experience involves spontaneity; I aim to do in section IV by invoking the idea of "normative sorting," illustrated with reference to the sorting

behaviour of children in the early stages of language-learning. Section V offers a streamlined version of the overall argument and addresses an objection to my claim that normativity can be divorced from reasons.

II. Kant on spontaneity and self-consciousness

Kant's view of self-consciousness can be motivated in the context of a familiar historical story which begins with Descartes. Reflecting on whether there is anything he can know in the face of evil genius doubt, Descartes concludes that he can know that he is thinking, and therefore that he exists (Descartes 1985 [1637], 127; 1984 [1641], 16-17). In stating the premise "I think," Descartes takes for granted that he is conscious not just of thinking, or of thoughts happening, but of a subject that is doing the thinking, and specifically the very subject who entertains the thought "I think." Descartes's assumption that he is self-conscious, implicit in his use of "I", was challenged by Georg Lichtenberg, in his famous remark that "we know only the existence of our sensations, representations and thoughts...one should say "thinking is going on" [*es denkt*; literally "it thinks"] just as one says "there is lightning" [*es blitzt*; literally "it lightens"] (1971, 412). A related idea was expressed by Hume in another famous passage: "when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe anything but the perception" (Hume 2000 [1739-40], 1.4.6]). Descartes's initially plausible supposition that, in thinking, I am conscious, not just of the occurrence of thinking, but of *myself* thinking, turns out, then, to be problematic. If introspection does not reveal a self to me independent of my thoughts and perceptions, then what gives meaning to my use of "I"? It is no good proposing that I am aware of myself by being aware of my body, for, even assuming that I am my body, my awareness of my body is not awareness of myself *as* myself. It

is always open to me to doubt whether I am my body.[4] The same is true of any object we might suppose to be presented to us in perception, whether inner or outer perception. If introspection did reveal to me some inner item, in addition to my perceptions and thoughts, which I was tempted to regard as the bearer of my perceptions and thoughts, I could always ask "but is that thing *me*?" And that is enough to show that my consciousness of the item could not be self-consciousness in the sense associated with the use of the first person.

Kant offers an indirect answer to the question of self-consciousness in the *Critique of Pure Reason*'s Transcendental Deduction, which argues for the necessary applicability of the pure concepts of understanding (for example *substance* and *causality*) to objects in space and time. The core of the argument is that experience of objects requires an active "synthesis" (a combining or unifying) of the sensory representations that are given to us, and that the pure concepts are, or correspond to, *a priori* rules by which this synthesis is governed. If all spatio-temporal experience of objects is arrived at through the activity of synthesis, then the objects thereby represented must fall under the concepts corresponding to the rules of synthesis, that is to say the pure concepts of understanding. Self-consciousness enters the argument because Kant motivates the need for synthesis by saying that it is required if I am to be able to think of my representations as mine, where that possibility is in turn required if I am to be conscious of them at all (B131-132).[5] In order for me to be conscious of an "I" which remains the same through my various representations—without which consciousness I would have as "as multicolored and diverse a self as the representations of which I am conscious" (B134)—I must, Kant says, both be able to synthesize these representations and to be conscious of this synthesis. The "thoroughgoing identity of the apperception of a manifold given in intuition," that is the consciousness of a single "I" to which the given representations belong, "contains a synthesis of representations, and is possible only through the consciousness of this synthesis" (B133). Although Kant's point is to argue for the applicability of the

[4] For versions of this point, see Martin 1995, Peacocke 2014, 45ff.
[5] Following standard practice, I use A and B page numberings to cite the first and second editions of the *Critique of Pure Reason*. Translations are my own.

categories rather than to give an account of self-consciousness, what he says constitutes an answer to the question of how I can be self-conscious even though I am not presented to myself—that is, not presented to myself *as* myself —in the form of an object of my awareness. My self-consciousness, he suggests, consists in my consciousness, or potential consciousness, of the activity through which I synthesize my representations: a consciousness which he labels "apperception."

What is this activity, and how am I conscious of it? Kant initially introduces synthesis, in the section of the first Critique immediately preceding the Transcendental Deduction, as "the mere effect of imagination, a blind though indispensable function of the soul, without which we would have no cognition at all, but of which we are rarely even conscious" (A78/B103). And in many passages in the Deduction itself, he also ascribes synthesis to the imagination, which he famously describes as a "necessary ingredient of perception" (A120n), suggesting that the role of synthesis is to form unified perceptual images out of the raw sensory data given to us. In the perception of a house, for example, which he describes elsewhere as a complex matter requiring us to unify discrete perceptions of its various spatial parts (A192-3/B237-8), "I make the empirical intuition of [the] house into perception through apprehension of its manifold...and..as it were draw its shape [*Gestalt*]" (B162).[6] This might tempt us to think of the activity as something like the subpersonal visual processing or "binding" through which the perceptual system combines disparate elements of a perceived scene (for example, the colors and shapes of the objects perceived, or the spatial or temporal parts of the scene) to yield the perception of unified, three-dimensional objects (Treisman 1996). In that case, it is hard to see how we could ever come to be conscious of the activity, let alone how our consciousness of it could amount to self-consciousness. But this does not seem to be what Kant has in mind in speaking of synthesis.[7] For he also ascribes synthesis to the faculty of understanding, suggesting that it takes place at the level not only of perception, but also

---

[6] Kant's most detailed explanation of synthesis, so understood, is in his first edition discussion of the "synthesis of apprehension in intuition" and the "synthesis of reproduction in imagination" (see A98-102).

[7] Here I agree with Longuenesse 2017, 41n29.

of explicit judgment. I engage in synthesis, for example, when I combine the representations *body* and *heavy* to make the judgment "bodies are heavy," thus "bring[ing] given cognitions to the objective unity of apperception" (B142).

This might lead us to think that "synthesis" for Kant refers to two distinct kinds of activity: a subpersonal activity, ascribable to imagination, whereby sensory representations are combined to yield perceptual images, and a personal-level activity, ascribable to understanding, of combining concepts to yield judgments. But that would be a mistake, because—and this is crucial to his argument for the applicability of the pure concepts to objects given in experience—Kant takes the activity of imagination in perception and the activity of understanding in judgment to be identical. "It is one and the same spontaneity that, here under the name of imagination, and there under the name of understanding, brings combination into the manifold of intuition" (B162n). We can make sense of this identification if we join Kant in supposing that perceptual experience, for human beings, always involves representing the particulars we perceive as having general features. The role of imagination in perceptual experience is that of creating representations with intentional content: making it possible, for example, that I not only have visual sensations that are caused by the sight of the house, but that I see the house *as* a house. Although the kind of combination is different in the two cases—in perception, the "reproduction" of representations of the same object viewed on previous occasions, or of previously viewed objects of the same kind; in explicit judgment, the combination of one concept with another according to one of the forms of judgment—both count as exercises of judgment in the broad sense of "thinking the particular as contained under the universal" (*Critique of Judgment*, Introduction, IV, 5:179).

We now turn to the question of what it is for me to be conscious of this activity—conscious, that is, in the distinctive way that constitutes self-consciousness. Kant makes clear that it is not a matter of my experiencing it by means of the senses, for the consciousness of activity can be ascribed to the understanding alone. "The representation ['I think'] is an act [*Aktus*] of spontaneity, i.e. it cannot be regarded as belonging to sensibility" (B132; Kant's emphasis omitted); the "synthesis [of understanding],

regarded in itself alone [*für sich allein*], is nothing other than the unity of *the action of which it is conscious as such even without sensibility*" (B153; my emphasis). This excludes the various kinds of empirical knowledge I can have of my own perceiving and judging through observing my own bodily movements or becoming aware of them through proprioception. It likewise excludes the introspective consciousness of my own perceptual and thought processes which we might think is available through inner sense, and which depends on our being "internally affected *by our own selves*" (B156). These various forms of perceptual or quasi-perceptual consciousness of my perceiving and judging present them to me as part of the objective temporal order, that is, as events whose occurrence is governed by natural laws. And it is clear that, for Kant, the consciousness of spontaneity or activity which yields self-consciousness proper—that is, apperception—is supposed to exclude the consciousness of mental goings-on as part of the objective natural order. To be conscious of my activity *as activity* precludes my being conscious of it *as happening*, as a series of events whose occurrence is causally determined in accordance with natural laws. Unlike my experience of external events, or of the temporally extended succession of my own inner states, this distinctive consciousness of the synthesis through which experience is possible does not itself involve synthesis and so does not amount to the representation of my activity as part of a causally determined spatio-temporal order.

This can seem mysterious. How can I be conscious of my mental goings-on without being conscious of them as a temporally extended process, and so without the aid of inner or outer sense? Although Kant says nothing to address this puzzlement directly, an answer emerges from the contrast between representing something as part of the natural, causally determined order, and representing something as part of a normative order, or as subject to normative constraints. I can be conscious of my activity of synthesis *as* activity—rather than as part of the natural order of events—by being conscious not of how it does happen, but of how it *should* or *ought to* happen. This counts as consciousness of activity because, in representing my mental goings-on as subject to normative rather than natural law, I represent them as to-be-done rather than as happening. In effect I adopt the perspective of an agent, one

who is responsible for the mental goings-on, rather than the perspective of an observer who merely experiences them. To be active or spontaneous, then, is not to be carrying out some distinctive kind of psychological activity, say, forming images, or putting together concepts into judgments; rather, it is to be consciously subject to normative constraints in the performance of such an activity.[8] Correspondingly, I am conscious of my activity in the distinctive way associated with self-consciousness when, in perceiving something as having a certain feature, or making an explicit judgment about it, I am conscious of how I *should be* perceiving or judging, or how the thing in question *is to be* perceived or to be judged. When I perceive the object I am looking at as a house, I am conscious of the various aspects of my perception—that I perceive it as having a determinate shape and colour, as three-dimensional rather than a mere facade, as having side and back walls, windows, a chimney, and so on—as normatively called for. I do not simply feel psychologically impelled to represent the object in those ways; rather, I experience my ways of representing the object as normatively necessitated, as ways that it is *to be* represented, or ways in which I (or anyone else in my situation) ought to represent it. Similarly, when I make the explicit judgment that, say, houses have walls, I take it that the concepts of *house* and of *having walls* should be combined as I am combining them, or that they belong together in the form of a universal judgment.

This interpretation of the consciousness of activity helps explain why, in a move central to the argument of the Deduction, Kant identifies the unity of apperception with objective unity, that is, the unity through which we represent an objective world independent of our thinking. What it is for me to represent the house's color and shape, and its having windows and walls, as features of something existing independently of me, as opposed to those representations' merely figuring together in my subjective consciousness, is just my being conscious of the representations as normatively necessitated. Similarly, what it is for me to make the objective claim that houses have walls, as opposed to the representations *house* and *having walls* merely being associated for me as a matter of psychological law, is for me to take it that I, or anyone else, ought to combine those representations in that way. This

[8] I argue for this interpretation of spontaneity in Ginsborg 2006b; see especially section IV.

normative constraint is the "element of necessity [*etwas von Notwendigkeit*]" associated with the relation of cognition to an object which "prevents our cognitions from being determined haphazardly and arbitrarily" (A104). But also, and more relevantly to the present context, the interpretation of activity in terms of normativity helps us understand why the consciousness of activity or spontaneity should amount to self-consciousness, and so provide an answer to the question raised by Hume and Lichtenberg. Consciousness of a self is not consciousness of some object additional to my representations, nor of objective events of mental processing by which my representations are in fact combined, but of agency with respect to my representations, where the consciousness of agency in turn is the consciousness of the combination of my representations as subject to normative standards, as *to be* combined in this and not some other way.

III. The rational agency interpretation of Kantian self-consciousness

I turn now to Moran's conception of self-consciousness in terms of rational agency. This conception emerges in the context of his account of self-knowledge, understood as the distinctively first-personal knowledge that we have (or are commonly thought to have) of our own mental states in contrast to those of others. Knowledge of this first-personal kind is characterized both by its immediacy, in the sense that we do not need to appeal to behavioral evidence to determine what we (say) believe or desire, and by the special authority which is typically granted to reports of our own beliefs and desires when they are made on the basis of that knowledge. We might be inclined to think that this knowledge arises from the exercise of a special capacity for inner self-observation, as though each of us were the sole spectator at an internal theater in which our mental states are displayed to us. But Moran rejects this spectatorial model of first-personal self-knowledge, arguing instead that our first-personal knowledge of our beliefs and other attitudes is a form of knowledge which is practical or deliberative, as opposed to theoretical or contemplative. It arises directly from our capacity to determine what to believe or what to desire on the

basis of reasons, and this accounts for its immediacy and authority. I come to know what I believe, in the characteristically first-personal way, by coming to know what *to* believe, that is, what I *should* believe, or what I *have most reason to* believe. That is, I come to know my own attitudes through the exercise of rational deliberation or rational agency in the formation of those attitudes, and not through observation of those attitudes or of the process by which I come to arrive at them. Moran develops this point in terms of the "transparency" of one's thinking, a phenomenon identified by Gareth Evans (Moran 2001, 61). According to Evans, "if someone asks me 'Do you think there is going to be a third world war?,' I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question 'Will there be a third world war?'" (Evans 1982, 225). In other words, I must consider whether I should, or have reason to, believe that there will be a third world war. For Moran, that consideration of reasons is the source of my first-personal knowledge of what I believe. I can know immediately that I believe it will rain through recognizing, on the basis of reasons, that I *should* believe it will rain. And a parallel point can be made for desire or intention. I can know what I desire, in a distinctively personal way, by determining what is desirable, and so what I should, or have reason to, desire. Moran is concerned here to explain the possibility of first-personal knowledge of one's mental states, rather than the (arguably more general) capacity for self-consciousness, understood roughly as the capacity to think "I"-thoughts. But there is good reason to think that the account is intended to apply to self-consciousness as well, since "I"-thoughts are paradigmatically thoughts about one's mental states, and Moran's account explains how I come to think about my mental states in a distinctively first-personal way. Moreover, Moran identifies the immediacy of self-knowledge with a kind of immediacy often thought to be central to "I"-thoughts, namely that, in having such a thought, I refer to myself without the mediation of any identifying description of myself (2001: 133-134). So we can understand his view as addressing the same general issue that we discussed in the previous section, that of how I can be conscious of myself as myself. [9] And we can see it as addressing the issue in a way which resembles

[9] This is also how his view is understood by Gertler 2011 (see ch. 7).

Kant's approach in its appeal to agency and to normative constraint. For both Kant and Moran, self-consciousness is a consciousness of agency, made available simply through the exercise of that agency, and amounting to a consciousness of normativity: in Kant's case, that my representations should be combined in this or that way, and in Moran's case, that I should have this or that belief, desire or other attitude.

While Moran briefly acknowledges that his view draws on a Kantian tradition (2001: 89, 127, 138-139), he does not indicate any specific point of connection; nor, in particular, does he connect his approach with Kant's discussion of self-consciousness in the Transcendental Deduction. However, Boyle (2009) develops Moran's approach in a way that relates it explicitly to Kant's. One of Boyle's main aims is to defend Moran's approach to self-knowledge from the objection that it fails to account for our apparently privileged knowledge of non-intentional states like pain. I seem to have immediate and authoritative knowledge that I am in pain, but this cannot be self-knowledge through the exercise of rational deliberation because my being in pain is not the result of deliberation. Boyle's defence appeals to a distinction between what he calls "two kinds of self-knowledge" invoked by Kant, the first being the consciousness of activity or spontaneity which Kant calls apperception, and the second being knowledge of one's own mental states through inner sense.[10] Self-knowledge à la Moran, he argues, corresponds to the first kind of self-knowledge in Kant, whereas the knowledge that, say, I am in pain, corresponds to the second kind. That second kind of knowledge is not arrived at through rational deliberation, so cannot be explained on the same model as knowledge of our beliefs and other intentional states. But, Boyle argues, self-knowledge à la Moran is still essential for the second kind of self-knowledge and, more generally, for self-consciousness. His argument rests on an analysis of the conditions for entertaining a thought like "I am in pain." These include in the first instance, that I am capable of understanding the utterance through which I express that thought, as opposed to merely having a capacity to produce the utterance "I am in

[10] I use scare-quotes because I disagree with Boyle's identification of apperception with a form of knowledge, on the grounds that, for Kant, knowledge requires sensible intuition. This objection is also raised in Khurana 2019. However, it can be disregarded for the purposes of this paper.

pain" in a parrot-like way when I am in pain. Understanding the utterance "I am in pain" requires recognizing that utterance as a complex of meaningful elements capable of figuring in other utterances, which in turn requires a capacity to recognize "relationships of [e.g.] implication, exclusion, and inductive support" among one's utterances" (2009, 150). But that capacity in turn requires the capacity to reflect on my reasons for holding a given utterance true, and to adjust my beliefs based on that reflection; in other words, it requires the capacity to make up my mind on the basis of reasons (2009, 151). Moreover, the understanding of the expression "I" in particular requires not just that I have that capacity, but that I recognize myself to have that capacity: a person's "implicit grasp that he has the power to make up his mind is a condition of his understanding the first person at all" (2009, 155). The upshot is that self-knowledge à la Moran—the knowledge we have of our attitudes in virtue of our capacity for rational deliberation about what to believe and what to do—is a requirement for all self-consciousness and in particular for our self-conscious knowledge of our nonintentional, as well as our intentional, states.

Boyle's discussion brings out clearly two elements that are implicit in Moran's account of self-consciousness: first, the idea that self-consciousness depends on the capacity for rational deliberation, and, second, the connection between that idea and Kant's conception of self-consciousness as the consciousness of spontaneity. Boyle articulates this connection by equating the consciousness of spontaneity, as it figures in the Transcendental Deduction and related passages in the first Critique, with the consciousness of one's capacity for rational deliberation. Implicit in this equation is the assumption that Kantian spontaneity amounts to rational agency. This assumption is widely held. Many philosophers have taken for granted that Kant's talk of the "unity of apperception" refers to a rational unity among our judgments, whereby one judgment is recognized by us as potentially a reason for, or rationally supported by, other judgments.[11] Relatedly, they have assumed that the activity of synthesis essentially involves

[11] Notably Robert Brandom; see especially Brandom 2009, 39, although the point is already implicit in his 2001, chs 1 and 2. Others who endorse this view, or who otherwise identify spontaneity, in Kant's sense, with the capacity to assess reasons, include Rödl 2007 (ch. 4), Kitcher 2011 (see e.g. 121-123), and Longuenesse 2017 (see e.g. 26-28).

consideration of reasons, so that the synthesis whereby I arrive at any given judgment (for example combining the representations *body* and *heavy* to yield the judgment *bodies are heavy*) involves my considering the reasons for making that judgment in the light of other judgments. And these assumptions are especially natural if we are impressed with the thought, articulated in section II, that synthesis involves the recognition of normative constraint in the combination of one's representations. For it might be thought that to recognize normative constraint in the combination of *body* and *heavy* just is to recognize that one has reason to judge that bodies are heavy.

However, I want to call attention to two reasons for questioning the assumption that spontaneity (and by extension normativity) should be understood in terms of rationality. The first, which I presented in section I, and which bears on the primary aim of this paper, has to do with self-consciousness. If we identify spontaneity with rationality, then the Kantian account does not extend to children who are apparently self-conscious, but too young for rational deliberation. The second, which will be my primary focus in the next section, has to do with the fact that synthesis for Kant plays an essential role in perception as well as in explicit judgment. If we discount the idea that synthesis is involved in perception as well as explicit judgment, then it might seem quite plausible that the sense in which I ought to combine *body* and *heavy* in the unity of apperception is that I have good reason for judging that bodies are heavy given my previous experience of bodies or my knowledge of Newtonian physics. But it is difficult to see how the consciousness of normativity in perception could be understood in terms of the recognition of reasons. Unlike our explicit judgments, our perceptions are not responsive to reasons: my perceiving a presented object as a house, or as having walls and windows, is not something I do as a result of rational deliberation, or for which I can offer rational justification. Moreover, there is a difficulty in principle about supposing that perceptual experience requires us to recognize reasons: namely, that this recognition requires grasp of the concepts which figure in the reasons, and—barring thoroughgoing innatism—we need perceptual experience to arrive at the concepts. It might initially be supposed that, if I perceive the house immediately before me as a three-dimensional structure with side and back walls (rather than as a

mere facade), my consciousness that this is how it should be perceived is due to my reasoning inductively along the following lines: this looks like a house, houses I've previously observed have side and back walls, therefore this house has side and back walls, therefore I am perceiving it as I should. But in order to engage in this kind of reasoning I need to have acquired concepts like *house* and *having walls*, and it is unclear how I could do so in advance of being able to perceive objects as being houses or having walls. So even though I can perfectly well carry out that reasoning retrospectively, as I reflect on my perceptual experience, that cannot account for the consciousness of normativity that is involved—on the view that I have ascribed to Kant—in perceptual synthesis itself. If Kantian spontaneity really does figure in perceptual experience, and not just in explicit judgment, then we cannot identify it with the capacity to recognize reasons.

IV Perceptual spontaneity and normative sorting

I have just argued that the identification of spontaneity with rationality prevents us from seeing how spontaneity could figure in perceptual experience as well as in the making of explicit judgments, and I have proposed this as a reason for rejecting that identification. But if this reason is to have force, then we have to understand better what it could be for spontaneity to figure in perceptual experience, that is, for perceptual experience to involve the consciousness of normativity. For otherwise it will be tempting to retain the identification of spontaneity with rationality by dismissing or downplaying Kant's claims for the spontaneity of perceptual synthesis, and instead regarding spontaneity as limited to the making of explicit judgments. In this section I will attempt both to clarify, and to make plausible, the idea that perceptual experience involves spontaneity. I will do so by sketching an account of perceptual experience which I have presented previously in the context of contemporary debates about the content of experience (Ginsborg 2006a, 2011, 2021), and in which the consciousness of normativity—understood independently of rationality—plays a central role.

This account, which I have labelled a "moderate conceptualism" about the content of experience, aims primarily to reconcile two apparently conflicting ideas: first, that perceptual experience (at least for mature humans) involves the perception not just of individuals, but of individuals as having general features; second, that perceptual experience is more primitive than explicit thought and judgment, in the sense that the concepts figuring in explicit thoughts and judgments are acquired on the basis of, rather than antecedently required for, perceptual experience.[12] Philosophers impressed by the first idea have tended to endorse conceptualism about perceptual experience, that is the view that perception has conceptual content; the second has typically motivated the view, either that perception has content which is nonconceptual, or that perceptual experience lacks representational content altogether. The difficulty in reconciling these emerges if—as is natural—we equate the idea of perceiving something as F with the idea of the concept F's figuring in the content of one's perceptual experience of the thing. If perceiving something as a dog is a matter of having a perceptual experience whose content includes the concept *dog*, then it can seem as though one must possess *dog* prior to being able to perceive something as a dog. That is, it can seem as though only someone already capable of making explicit judgments using the concept dog (for example that Lassie is a dog, or that dogs bark) can come to see Lassie as a dog. But that seems to undermine the idea that perceptual experience is prior to judgment, and in particular that we are able to acquire the concepts which figure in our judgments as a result of having the corresponding experiences rather than vice versa.

One way to address the conflict would be to adopt a weaker notion of "perceiving as F" which ties it to the capacity for discriminating Fs from non-Fs, or (roughly equivalently), sorting or classifying Fs together with other Fs. Someone might count as seeing something as a dog, in this less demanding sense, if, on seeing it, she is disposed to respond to it in a way which is distinctive of her responses to dogs in general (for example, by saying the word "dog" or by tokening a mental representation which is

[12] I contrast experience with "explicit" thought and judgment, rather than with thought and judgment *tout court*, because proponents of the first idea often suggest that there is a kind of implicit thinking or judging involved in experience itself (Sellars refers to "thinking-in-presence," 1963, 162).

nomologically correlated with the presence of dogs). A pigeon trained to peck a key when it sees a green light, as opposed to lights of other colours, would count as seeing the light as green if, on seeing the light, it is disposed to peck the key. But philosophers attracted by the first idea typically have something stronger in mind, namely that the general feature figures in the first-personal phenomenology of the experience, as something which is available from the perceiver's point of view. When I see a green light, and, in so doing, see the light as green, the greenness of the light is present to me in a way that it need not be to the pigeon. Even if there is something which it is like from the pigeon's point of view to see a light which is green, so we can say that it sees the green light *in* a distinctive way, this is not yet to say that the pigeon sees the green light *as being* a distinctive way.[13] By the same token, while we might think of the pigeon as sensitive to the presence of greenness, that does not mean that the greenness is represented in its experience.

What we want, then, is an account of perceiving something as F that demands more than merely perceiving it in a way which manifests sensitivity to F-ness, and that, in particular, does justice to the intuition that the F-ness is somehow present to us in the perception, but that is not so demanding as to require that we already be able to entertain F-thoughts as a prior condition of perceiving something as F. I believe that we can reach the desired middle ground by invoking what I will call "normative sorting": a kind of sorting or discrimination which involves the consciousness of normative constraint. Perceiving something as a dog, on this approach, is perceiving it in such a way as to sort it with previously perceived dogs, where the sorting involves consciousness that one is sorting appropriately, that is, that the thing belongs with, or should be sorted with, the previously perceived dogs. The idea of normative sorting can be illustrated by considering the overt behavior of small children, say between 15 and 30 months old, presented with randomly arranged collections of small objects of two clearly demarcated kinds (say, four

[13] I expand on this distinction in section 1 of Ginsborg 2016c.

yellow cubes and four gray balls, or four dolls and four toy boats).[14] As numerous studies have shown, children in this age range typically engage in "spontaneous sorting": without any prompting, they move the objects around to form two discrete groups, each composed of one kind of object.[15] Moreover they do so in a way which suggests that they take a normative attitude to what they are doing, that is, that they see themselves as putting together the things which belong together, or which should be sorted together. This is clearest in children at the upper end of the age range who have learned to use some normative language. Researchers report that 30-month-olds presented with two mismatched sets of four objects each (in one experiment three circles and a square, and three squares and a circle) respond with comments like "No belongs this way," "Not good in there," "no, they're not on properly."[16] But this normative attittude is also apparent from the behavior of children at the lower end of the age range, who give the impression, in their manipulations of the objects, of aiming to sort them correctly: for example they will inspect objects before putting them together, hesitate as if trying to decide where an object should go, reverse a move they have just made, as if noticing and correcting a mistake, and so on.

We can also see children's early language use as a form of sorting, specifically normative sorting. Alison Gopnik and Andrew Meltzoff have proposed that we understand the early use of names ("names" here including general terms like "cube" or "dog") as a form of sorting or categorization. As they put it, "a name places some of the objects in the world into a particular group."[17] On this proposal, when an 18-month old sees a dog and says "dog," she is behaving like the 18-month-old who adds a cube to the group of cubes she has already assembled. The child's giving the same response to the dog as she has given, or heard given, to dogs encountered earlier, is like her placing a cube in the same location as other cubes that she, or someone else, has gathered together. And, as with the spontaneous sorting of objects into kind-

[14] For a more detailed account of the research on which I am drawing in this section, see section 3 of Ginsborg 2021.
[15] For a review, see Gopnik and Meltzoff 1997, ch. 6.
[16] The first remark is reported in Langer 2001:22, the other two in Sugarman 1983:73
[17] Gopnik and Meltzoff 1992: 1093. I think the point extends beyond names: for example the toddler who says "Up" each time she climbs one stair of a staircase is sorting together each episode of the behavior. But this point is not essential for present purposes.

specific locations, this form of sorting—under kind-specific labels—involves consciousness of normative constraint. This emerges from studies of "false labelling," which investigate children's responses to the incorrect use of names.[18] These studies show that children as young as 16 months respond with apparent protest when they hear a speaker label an object incorrectly by saying, for example, "that's a shoe" while looking at some other object, such as a cat. At 18 months, children respond by saying "no"; among the 16-month-olds, responses included shaking their heads, waving their hands, or, for the example cited, pointing to their own shoe.[19] Although the researchers in the studies I have cited suggest that the behavior should be understood as the negation or rejection of the speaker's false utterance—and even that it might "signal an understanding in infants of what human speakers apparently 'meant' to say"[20] — Gopnik and Meltzoff's proposal makes room for a simpler explanation in terms of normative sorting. The speaker who labels the cat with "shoe" is in effect putting the cat together with the items previously labelled "shoe", and the child who protests, or who points to her own shoe, is expressing her consciousness that this sorting behavior is incorrect, or attempting to correct it by indicating that it is her shoe, and not the cat, that should be sorted with the previously labelled shoes.

We can return now to the question of what it is to perceive something as F, and in particular of how to find a middle ground between mere perceptual sensitivity to the presence of F-ness, and the capacity to make explicit judgments involving the concept F. I have suggested that we can answer the question by identifying perceiving something as F with perceiving it in such a way as to sort it with other Fs, where one's sorting involves the consciousness of normative constraint, or—by way of abbreviation—where the sorting is normative. The idea that normative sorting can be carried out by children in the age range we have been discussing helps us to see why it might help us to the desired middle ground. For on the one hand, the children's sorting behaviour clearly manifests more than the differential responsiveness

---

[18] Pea 1982, Koenig and Echols 2003

[19] Koenig and Echols 2003, 201; this last response was given by five of the sixteen children tested. . I here take the opportunity to correct a mistake in Ginsborg 2021, where I characterized the children as responding to the false labelling of a shoe, rather than to the use of "shoe" to label a non-shoe.

[20] Koenig and Echols 2003, 201

exhibited by trained pigeons. But on the other hand, given the children's level of development, it is implausible to suppose that they are capable of explicitly judging that the objects they are sorting are say, cubes or shoes, let alone of making other judgments whose content includes the concepts *cube* or *shoe*. Even if they can use "shoe" appropriately—for example, pointing to their shoe when asked "Show me your shoe", or producing the word "shoe" themselves under appropriate circumstances—it does not follow that their use of "shoe" amounts to the assertion of something's being a shoe. And even if we do suppose that a child who physically sorts a shoe with other shoes is in some sense judging that the shoe is a shoe, or has some more general property in common with the other shoes, there is no reason to think that her capacity to make this kind of judgment explains her recognition that the different shoes belong together. It is far more plausible to suppose that, if she counts as judging the shoe to be a shoe, this is because of a more primitive capacity to recognize the shoe as belonging with the other shoes, and not the other way around.[21] The existence of normative sorting in children, then, allows us to see that the capacity for normative sorting of objects of a given kind is possible prior to the capacity for making judgments involving the corresponding concepts.

This opens the way for the desired middle-ground account—and here we should note that it applies to adults as well as children—of what it is to perceive something as F. For once we recognize in principle the independence of normative sorting from explicit judging, we can use the idea of normative sorting to make sense of perceiving something as F without having to appeal to the capacity to make explicit judgments involving the concept F. We can do justice to the thought that, when I perceive something as a dog, its being a dog is present in my experience and not just a feature to which I respond, in terms of the thought that I perceive the thing as belonging with, or to be sorted with, previously perceived dogs.[22] As the example of normative sorting in small children shows, that is something I can

---

[21] I discuss this point further in section V.

[22] It might be objected here, especially in the light of issues raised in Kripke 1982, that my perceiving a dog as belonging with a finite set of previously perceived dogs could never be enough to determine that the feature presented to me in my perceptual experience is that of *being a dog* as opposed to, say, *being a*

do without first having identified either what I am perceiving now, or what I perceived previously, as being a dog. It requires only that I be capable of reliably sorting dogs with other dogs and that, in so doing, I recognize my particular sorting response—saying "dog", or putting a toy dog in one pile rather than another—as normatively called for. My having this capacity for normatively sorting dogs is sufficient for its being the case that, whenever my perception of a dog is such as to allow this kind of response (for example if the dog is in clear view, not disguised as a cat, and so forth), I count as seeing it as a dog.[23]

I have presented this account without reference to Kant, but it also helps us understand how, for Kant, spontaneity figures in perceptual experience. To see why, let us return to the idea, discussed in section II, that perception for Kant involves an activity of synthesizing representations—in particular involving the "reproduction" of previous perceptions—that yields a perceptual image with intentional content. I glossed the relevant notion of activity or spontaneity in terms of normativity, claiming that we should understand my consciousness of spontaneity in the synthesis of my representations as

---

dog not on the planet Mars. But this objection misrepresents the role played in the account by the normative element in normative sorting. Although the recognition of normative constraint makes it the case that the feature of being a dog is presented to me in my experience rather than merely being something I am sensitive to, that aspect of the account is not intended to determine *which* feature is presented to me, over and above its being some feature common to the dogs I have previously perceived (as opposed for example, to the feature of being green, or a cat). The work of narrowing down the feature further is done by the part of the account which requires, for perceiving something as F, that I have, or perhaps am disposed to acquire, a reliable disposition to produce a sorting response to Fs. On a robust view of dispositions, that is sufficient to rule out quus- or grue-like alternatives to being a dog. Thanks to Alex Miller for discussion of this point.

[23] As suggested above, I regard this account as amounting to a form of conceptualism about perceptual experience. This reflects my view, argued for in section 2 of Ginsborg 2021, that the debate over conceptualism should be framed in terms of the Kantian notion of concepts as representations which are universal (in contrast to particular or singular), as opposed to the more recent and more commonly cited notion of concepts as inferentially relevant constituents of propositions (in contrast to whole propositions). On my view, it is sufficient for a content's being conceptual, and, more specifically, for its including the concept F, that it involve the representation of something as F in the sense offered here. However, for present purposes, it does not matter whether we think of perceptual content as conceptual or nonconceptual, so long as it is allowed that perceptual experience involves what Kant calls spontaneity, or, equivalently, the exercise of understanding. Thanks to Kristina Musholt for raising this issue.

consciousness that my representations should be, or are to be, synthesized in the way I am synthesizing them. I now want to suggest that this synthesis can be seen as a kind of normative sorting. Consider P. F. Strawson's suggestive characterization of imagination's involvement in perception: "[T]he actual occurrent perception of an enduring object as an object of a certain kind, or as a particular object of that kind, is, as it were, soaked with or animated by, or infused with... the thought of other past or possible perceptions of the same object [or of other objects of the same kind]...non-actual perceptions are in a sense represented in, alive in, the present perception; just as they are represented, by images, in the image-producing activity of the imagination" (1974, 59). Strawson illustrates the idea of past perceptions' being "alive in" the present perception with the example of seeing something as a dog: "To see it as a dog, silent and stationary, is to see it as a possible mover and barker" (1974, 59).

Although Strawson does not explicitly relate this idea to Kant's talk of "reproduction," it is natural to read the passage as interpreting Kant's idea that aspects of past perceptions are reproduced in ("alive in") present perceptions. I see the dog as a possible barker—where this is an aspect of how I actually see the dog, not a judgment, external to my perception, that the dog I am now seeing might bark—by reproducing, in my present perception of the now-silent dog, past perceptions of dogs that were barking.[24] The crucial point for our purposes is that this reproduction of past representations is a kind of sorting. My perception of the now-silent dog as a possible barker, that is, my reproduction, in my current perceptual situation, of previous perceptions of barking, amounts to my sorting the presently perceived dog with previously perceived dogs, just as a child's responding to a newly perceived dog with "dog" amounts to her sorting the dog with the other objects to which she has heard the word "dog" applied. When we combine this with the idea that my imaginative activity involves consciousness of normative constraint, so that I recognize the representation of barking as normatively called for in this perceptual context, we reach the thought that perceptual synthesis just is a kind of normative sorting. So my account

[24] Strawson himself denies that, for Kant, I literally reproduce past perceptions of dogs, as if my present perception included superimposed images drawn from my past perceptions. The question how to make psychological sense of "reproduction" is difficult; for present purposes we can leave it aside.

of perceiving something as F, on which the intentional content of perception is explained in terms of the recognition of normative constraint in our sorting of objects, can be seen as a generalization of Kant's account of how spontaneity figures in perception, on which the intentional content of perception is explained in terms of the recognition of normative constraint in our perceptual synthesis. The core idea of both accounts is that we perceive things as having general features by virtue of our capacity to sort them normatively with other things which share those features: whether the sorting takes the form of physical manipulation, the use of words, or the recalling of previous perceptions.[25] As the example of children's sorting shows, we can exercise this capacity without having to be able to judge explicitly that the things sorted have the corresponding feature. But this exercise requires the consciousness of normative constraint—the consciousness of things as to be placed in a given location, or pointed to on hearing a word, or recalled as part of the formation of a perceptual image—and that is why it involves what Kant calls the consciousness of spontaneity.

V. Normative sorting without reasons

I have argued that Kantian spontaneity is broader than rationality, and I have done so by appealing to Kant's view that spontaneity is involved in perception, and not just in explicit judgment. I have also attempted to defend Kant's view that spontaneity is involved in perception by invoking the idea of what I have called "normative sorting," as manifested in the behavior of small children. The upshot is an interpretation of Kantian spontaneity on which the consciousness of spontaneity—and hence self-consciousness—does not require the recognition of reasons for one's beliefs and actions. It is enough

[25] Another way to connect children's normative sorting behavior with imaginative synthesis for Kant would be to appeal to the "extended mind" hypothesis (Clark and Chalmers 1998) to claim that children's physical sorting of objects or early language use *just is* the mental activity of imaginative synthesis. But I will not pursue that possibility further here.

that one be able to engage in the synthesis of imagination which makes possible the perception of objects as having general features. The account of self-consciousness I am ascribing to Kant preserves the idea, emphasized by recent proponents of the Kantian tradition like Moran and Boyle, that self-consciousness is consciousness of a kind of agency, construed in turn as consciousness of normative constraint, of things (taken in a very general sense) *as to be done* rather than *as happening*. But it abandons the requirement that the agency and normativity be understood as rational agency and rational constraint. Perceptual experience for Kant, I have argued, involves recognition of a more primitive form of normative constraint governing the imaginative synthesis which is responsible for perception's having intentional content. Perceiving the object as a dog involves recognizing various aspects of my perception—my perceiving it, say, as a possible mover and barker—as normatively called for. But this is not the recognition that I have reason to represent it as a barker, say on the grounds that it is a dog and that dogs bark. Rather, it is a more primitive recognition of this dog as belonging with the dogs I have perceived in the past, one which leads me, without reasoning and yet with a consciousness of normative constraint, to reproduce in my present perception elements, like barking, which figured in my previous perceptions of dogs. This consciousness of normativity in my perceptual synthesis is sufficient for the consciousness of spontaneity or agency which, for Kant, constitutes self-consciousness. We have thus arrived at an account of self-consciousness of the kind I was aiming to offer in this paper: an account which shares the Kantian orientation of the rational agency approach, but which is accommodating enough to allow self-consciousness to children not yet capable of rational deliberation.

In developing this account, I have taken a somewhat circuitous route, drawing both on an interpretation of Kant's own views on spontaneity and self-consciousness, and on considerations about the content of perceptual experience. But it is worth noting that the primary desideratum of the account—to makes sense of a broadly Kantian notion of self-consciousness which does not require the capacity to recognize reasons—can be satisfied simply by appeal to the notion of normative sorting, without any reference to Kant or to perception. This yields a more direct argument for the possibility of agential self-

consciousness without rationality. If it is granted that the children in the spontaneous sorting studies not only put the same-kind objects together, but, in each instance of putting an object with others, take it that this is where the object *should* go, then we can ascribe to them an agential attitude to what they are doing. The 18-month old who hesitates over whether to put a doll with the other dolls or with the boats is experiencing her behavior not just as something happening, but as something with respect to which she is active, something for which she is responsible.[26] Moran's insight about the importance of the agential attitude to self-consciousness can thus be preserved, but without requiring that this attitude involves appreciation of reasons for what one is doing.

---

[26] It might be objected here that normative sorting behavior is not distinctive in this respect, and that the same could be said of earlier forms of purposive behavior, for example a 3-month-old's pulling a chain to shake a rattle (vividly described in Piaget 1952: 162-164) or a 9-month-old's pointing at an object to draw her caregiver's attention to it. Aren't the children in these activities aware of their agency, as suggested by Bermúdez (1998: 261-262)? And isn't there also a form of normativity here, in that the child rccognizes that the string is to be, or should be, pulled, in order to make the rattle sound? The answer is that these earlier forms of behavior are different from the normative sorting behavior because they are directly motivated by desire for immediate reward, e.g. the entertaining rattle sound or the socially pleasurable joint attention. While the infants may feel themselves drawn towards the behavior by the anticipation of the result, this does not imply that they take a normative attitude towards their behavior: they may be responding to an affordance, but that does not amount to recognition that this is how they *should* respond. (Here I dissent from Korsgaard's characterization of animals' responses to the affordances they perceive as "primitively normative," involving a sense of the response as "appropriate" to, or "called for by" what occasions it (2009, 110-113)). So, even if the infants have a distinctive experience of the connection between their internally felt efforts and the externally observed results, this is not a consciousness of agency in the sense associated with normativity. The normative sorting which appears in the older age group is different because it is not directly reward-driven. The child's consciousness of one doll as "to be put with" the others cannot be reduced to her being drawn to put them together by the anticipation of a pleasant or interesting effect. If she does feel pleasure in putting the dolls together, it is because she takes it that the dolls should be put together, and enjoys doing what she takes to be appropriate. The point is related to Gopnik and Meltzoff's observation that, in contrast to other forms of discriminative behavior, the child's sorting responses of "placing objects in one particular location rather than another" are "arbitrary" and "removed from any immediate reward," and to their conclusions that "categorizing...is its own reward" and that sorting behaviors have a "purely cognitive motivation" (1997, 172).

At this point it might be objected that I have been putting too much weight on potentially debatable empirical considerations, in particular regarding the age at which children can offer or recognize reasons for belief or action. If I am right that children in the 15-30 month age range do indeed take a normative attitude to their sorting responses, according to the objection, perhaps that just goes to show that children are capable of recognizing reasons earlier than developmental psychologists have typically been willing to allow. Perhaps, when the child who is given the mixed-up boats and dolls picks up one doll and puts it beside another, it is because she has already recognized, of each of the dolls individually, that they are dolls (or that they are people, or look like Mommy, or have a head and legs, or display some other feature which distinguish them from the boats). And this thought might be encouraged by the fact that the children in spontaneous sorting studies often name the things they are sorting, or features of those things, as they sort them. Susan Sugarman gives examples of a two-year-old pointing to two dolls in succession and saying "She's lady, that is a lady"; another two-year-old, grouping four rectangles, says "*One* brick, *two* brick... brick, brick, *four* brick"; another, collecting two blue blocks together, repeats "Blue, blue blue blue" (1983: 172). It might seem, then, that we can account for children's normative attitudes in their sorting by supposing that, say, the child who says of the two dolls "She's lady, that is a lady" is reasoning, albeit inchoately, along the following lines: *this is a lady*, *and that is a lady, so this should go with that*.

But it is important to see here that the possibility of normative sorting without reasons is motivated by philosophical as well as empirical considerations. Specifically, the capacity for normative sorting is more primitive than the capacity to recognize reasons, not just in the sense that normative sorting precedes the recognition of reasons in the course of a child's development, but in the sense that the recognition of reasons depends on the capacity for normative sorting. It is easiest to see this dependence if we adopt the commonly held and *prima facie* plausible assumption that the recognition of reasons as such requires some linguistic competence, at least to the extent of being able to understand sentences expressing the reasons that are recognized. For the child to be able to recognize that *this is a*

lady, and to regard that consideration as part of the reason why she should sort it with another doll which she has also recognized as a lady, she must have mastered, at least to some extent, the use of "lady". And if, as I have suggested, the early use of names is a form of normative sorting behavior, that means she must already be able to recognize, of anything she calls "lady," that it belongs with the other ladies — her mother, the daycare helper, the policewoman on TV, and so on.[27] So even if we were to suppose that, in the spontaneous sorting study, she takes one doll to belong with another on the grounds that they are both ladies, her being able to entertain that reason for sorting them together would depend on the more primitive normative sorting manifested in her use of "lady" in contexts beyond that of her behavior with the dolls and the boats. And that normative sorting, at least on the assumption that the recognition of reasons requires linguistic competence, could not in turn be based on the recognition of reasons. Since it seems that we have to ascribe normative sorting without reasons to the child in any case, it makes sense to interpret the child's repeated use of "lady" in the physical sorting task, not as an inchoate form of reason-giving but, rather, as a reinforcement of her physical sorting behavior: she is giving the dolls the same label as well as putting them in the same location.

What if we deny that the child needs to have grasped the use of an expression in a public language in order to think that *this is a lady*? What if we instead suppose, with Fodor, that she needs only a mental symbol-type in a language of thought, the tokening of which has come to be nomologically linked to the presence of things of the kind she will later come to call "ladies"? The difficulty here, briefly put, is that, on this kind of account, the thought *this is a lady* is not the kind of thing which could figure in someone's own personal-level reasoning about how she should sort the objects presented to her. Thinking *this is a lady* is being in a mental or neural state which is distinctively associated with the presence of the property of being a lady, without that mental state having intrinsic intentional content

---

[27] This might be disputed by Boyle, given his suggestion that children's early uses of language might be "parrot-like" (2009, 144). But it is more plausible (given, for example, the behavior of the 16-month-olds in the false labelling studies mentioned in section IV), and also makes better sense of the child's transition to a fully comprehending use of language, to suppose that children at the one-word stage already take a normative attitude to their uses of words.

beyond that which (according to the account) it has in virtue of instantiating a type associated with that property. So, on this view, when the child thinks *this is a lady* on seeing one of the dolls, the general feature of being a lady does not figure in her thought in such a way that she can represent the thing's being a lady as a reason to sort it with other things about which she has had the same thought (that is, which have activated the same kind of mental state), rather than sorting it with the things which have activated the thought *this is a boat*. We can certainly understand how thinking *this is a lady* so understood, could *cause* a child to sort a thing with the dolls rather than with the boats, but it would not amount to grasp of a *reason* for sorting the thing that way—at least not grasp of a reason *as such*.[28]

This might still leave the worry that, in the absence of the recognition of a reason, we cannot take the child to be genuinely conscious that she is sorting as she ought. Perhaps the considerations which might lead us to interpret a child in the 15- to 30-month age range as taking the dolls to belong together—

---

[28] See note ***1 for more on the idea of grasping reasons as such. We might ask here how the capacity to grasp reasons (as such) relates to the capacity for making explicit judgments and the capacity for rational deliberation. I understand these capacities to be of a piece: even though, on any particular occasion, someone can recognize a reason without explicitly judging that the reason obtains or engaging in rational deliberation, her capacity to recognize the reason depends on her being able to articulate it and to cite it as justification for her behavior. It might be argued, however, that the recognition of reasons need not require the other two capacities, given that small children can apparently be responsive to evidence—and thus, seemingly, recognize reasons—even if they are not in a position to articulate their reasons. Consider a case in which a child apparently corrects her sorting behavior in response to evidence that an object is not as it seems to be. For example, having just sorted a sponge that looks like a rock with other rocks rather than sponges, she is now shown that the sponge can be squeezed, and, as a result, she moves the sponge over to the other sponges, perhaps saying "No, it belongs here." Isn't this a case of the child's recognizing the squeezability of the sponge as a reason for sorting it with the sponges rather than the rocks? It could be, but it doesn't have have to be. When the child moves the sponge, with the attitude that this is where it belongs, she might simply be doing so as a causal response to the sponge's perceived squeezability. She recognizes the squeezability as a reason only if she is capable of thinking, not just *this should go here* — that is, the consciousness of normative constraint —*this belongs here because you can squeeze it*. If she cannot entertain that second thought, perhaps because she hasn't yet learned "squeeze" or "squishy," then the perception of the sponge's squeezability can still bring about, in a causal way, the normative consciousness of the sponge as belonging together with the other sponges. But that normative consciousness does not yet amount to the consciousness of a reason for sorting the sponge in that way. Thanks to an anonymous reviewer for pressing this line of questioning; for an overlapping objection, see Lauer 2021, 218.

her attentiveness to them, the seeming deliberateness of her behavior, her apparent corrections of her own and others' sorting behavior, and (at the higher end of the age range) the use of normative language accompanying her behavior—reflect only a desire to put the dolls together, and not a consciousness that they *should* be put together. The same, it might be said, holds for the apparently normative attitude which children take to language use, as indicated by their reactions in the false labelling studies: the 16 month-old who objects to a speaker's use of "shoe" for a cat prefers that the cat not be given the same linguistic treatment as shoes, but does not see the speaker as doing anything incorrect or inappropriate.

However—and this is intended as a *reductio*—if this non-normative understanding of small children's sorting behavior is motivated by the thought that sorting cannot be normative unless the supposedly normative attitude is grounded in the recognition of a reason, then we should be equally motivated to deny the possibility of normative sorting in adults. For the possibility of recognizing reasons for one's sorting behavior, in adults as well as children, depends on being able to sort things normatively in a way which does not depend on the possibility of reasons. Consider an adult who is given the dolls and boats and invited, without further specification, to sort them into groups. We can imagine that she might justify her putting one doll with another doll by saying (for example, to a small child looking on): "See, these go together because they are both dolls, and the others are boats." But the fact that they are both dolls is a plausible candidate for being a reason in this particular context only on the assumption that, in general, any one doll should be sorted with other dolls rather than, say, with boats. Without this assumption, the adult could equally well justify putting the doll with a boat on the grounds that they are both "bolls" (boats examined before time *t* or dolls examined after time *t*, for some suitably specified *t*) rather than "doats" (dolls examined before time *t* or boats examined after time *t*). Now of course, outside a philosophical context, and more specifically without the influence of Goodman's *grue* or Kripke's *quus*,[29] this kind of justification would never occur to us. Although we might recognize, on philosophical reflection, that there are indeed properties of the kind picked out by "boll" and "doat," as defined here,

[29] Goodman 1955, Kripke 1982

such properties strike us as obviously irrelevant to the question of what should be sorted with, or what "goes with," what. . But the fact that these properties are not salient to us, or that we find them irrelevant to the justification of our overt decisions about how to sort things, reflects a more primitive normative sorting which is manifest inthe meanings of our words and in the concepts we grasp. The English word "doll" means what it does for us—that is, it means *doll* rather than, say, *boll* or *doat* or any number of grue-like alternatives—only because, given any newly presented doll, we take "doll" rather than "boat" to be the *mot juste* for it, thus manifesting our attitude that the doll should be sorted with previously presented dolls rather than with previously presented boats. If we did not already recognize dolls in general as preferentially belonging together, then we could not grasp the concept *doll*, and we would be unable to invoke considerations like *that is a doll* as part of our reason, in the more limited context of the physical sorting exercise, for putting a doll in one place rather than another.[30]

I conclude with a brief recapitulation. My primary aim has been to show that we can adopt a Kantian approach to self-consciousness—one which identifies self-consciousness with consciousness of activity or spontaneity—without supposing that self-consciousness requires the capacity for rational agency. I argued this primarily by appealing to the role Kant ascribes to spontaneity in perceptual experience, in which we see things as having general features, that is to say as to-be-sorted-together in determinate ways, but not on the basis of reasons. This allows us, I claimed, to distinguish a general notion of spontaneity from the more specific notion of rationality, opening up the possibility of a consciousness of spontaneity which is not a consciousness of rational constraint. In developing the argument, I attempted to clarify the notion of spontaneity without rationality in terms of a certain kind of sorting behaviour manifested by chidren under three, which, I argued, manifests the recognition of

---

[30] What about the consideration that *this is the same as that*? Recognizing that consideration as a reason for sorting the dolls together requires grasping the concept *same* or the meaning of "same," and that in turn requires being able to recognize different instances of same-kind pairs as preferentially belonging together. So normative sorting (and indeed of a kind more difficult than that associated with grasping "doll") is still presupposed. The same point applies in the case of *this is similar to that* or *this is more similar to that than to the other things*.

normative constraint without the recognition of reasons. This pointed the way to an abbreviated form of the argument which does not rely on assumptions either about Kant's own views or about the nature of perceptual experience. Appeal to the normative sorting behavior of small children, both in the physical sorting of objects in spatial locations and in the very early use of language, shows that there can be consciousness of agency without rational deliberation. This can be invoked to yield an account of self-consciousness that is less demanding than the rational agency view, but retains the spirit of the broadly Kantian tradition from which that view emerges.[31]

.

[31] I presented earlier versions of this paper at conferences in Fribourg and Warwick, and at the University of Otago and the University of Auckland. For comments and discussion, on those occasions and others, I am grateful to Matt Boyle, Lucy Campbell, Naomi Eilan, Andrea Giananti, Naomi Ginsborg-Warren, Mariel Goddu, Alison Gopnik, Maya Kronfeld, Guy Longworth, Mike Martin, Alex Miller, Richard Moran, Kristina Musholt, Charles Pigden, Johannes Roessler, Daniel Warren, and an anonymous reviewer for OUP.

References

Bermúdez, José (1998), *The Paradox of Self-Consciousness* (Cambridge, MA: MIT Press).

Bilgrami, Akeel (2006), *Self-Knowledge and Resentment* (Cambridge, MA: Harvard University Press).

Boyle, Matthew (2009), 'Two Kinds of Self-Knowledge', *Philosophy and Phenomenological Research,* 78(1), 133-64.

Brandom, Robert (1994), *Making It Explicit* (Cambridge, Mass.: Harvard University Press).

Brandom, Robert (2009), *Reason in Philosophy: Animating Ideas* (Cambridge, Mass.: Harvard University Press).

Bretherton, Inge, McNew, Sandra, and Beeghly-Smith, Marjorie (1981), 'Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a "theory of mind"', *Infant social cognition,* 333, 73.

Burge, Tyler (1996), 'Our Entitlement to Self-Knowledge I', *Proceedings of the Aristotelian Society,* 96, 91-116.

Clark, Andy and Chalmers, David (1998), 'The extended mind', *Analysis,* 58(1), 7-19.

Descartes, René (1984 [1641]), *Meditations on First Philosophy*, in John Cottingham (ed.), *The Philosophical Writings of Descartes, II* (Cambridge: Cambridge University Press), 3-62.

--- (1985 [1637]), *Discourse on the Method*, in John Cottingham (ed.), *The Philosophical Writings of Descartes, I* (Cambridge: Cambridge University Press), 111-51.

Evans, Gareth (1982), *The Varieties of Reference* (Oxford: Oxford University Press).

Ginsborg, Hannah (2006a), 'Empirical concepts and the content of experience', *European Journal of Philosophy,* 14(3), 349-72.

--- (2006b), 'Kant and the Problem of Experience', *Philosophical Topics,* 34(1-2), 59-106.

--- (2006c), 'Aesthetic Judgment and Perceptual Normativity', *Inquiry*, 49(5), 403-437.

--- (2011), 'Perception, Generality and Reasons', in Andrew Reisner and Asbjørn Steglich-Petersen (eds.), *Reasons for Belief* (Cambridge: Cambridge University Press), 131-57.

--- (2021), 'Conceptualism and the Notion of a Concept', in Christoph Demmerling and Dirk Schröder (eds.), *Concepts in Thought, Action and Emotion* (New York and Oxford: Routledge), 42-59.

Goodman, Nelson (1955), 'The New Riddle of Induction', in *Fact, Fiction and Forecast* (Cambridge, Mass.: Harvard University Press.

Gopnik, Alison and Meltzoff, Andrew N (1992), 'Categorization and naming: Basic-level sorting in eighteen-month-olds and its relation to language', *Child Development,* 63(5), 1091-103.

--- (1997), *Words, Thoughts, and Theories* (Cambridge, MA: MIT Press).

Hume, David (2007 [1739-40]), *A Treatise of Human Nature*, eds David Fate Norton and Mary J. Norton (Oxford: Oxford University Press).

Khurana, Thomas (2019), ''I do not cognize myself through being conscious of myself as thinking': Self-knowledge and the irreducibility of self-objectification in Kant', *Canadian Journal of Philosophy,* 49(7), 956-79.

Kitcher, Patricia (2011), *Kant's Thinker* (Oxford: Oxford University Press).

Koenig, Melissa A and Echols, Catharine H (2003), 'Infants' understanding of false labeling events: The referential roles of words and the speakers who use them', *Cognition,* 87(3), 179-208.

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

--- (2018), *Fellow Creatures: Our Obligations to the Other Animals* (Oxford: Oxford University Press).

Köymen, Bahar and Tomasello, Michael (2020), 'The Early Ontogeny of Reason Giving', *Child Development Perspectives,* 14 (4), 215-20.

Kripke, Saul (1982), *Wittgenstein on Rules and Private Language* (Cambridge, Mass.: Harvard University Press).

Langer, Jonas (2001), 'The mosaic evolution of cognitive and linguistic ontogeny', in Melissa Bowerman and Stephen C. Levinson (eds.), *Language acquisition and conceptual development* (Cambridge: Cambridge University Press), 19-44.

Lauer, David (2021), 'Concepts, Normativity, and Self-Knowledge: On Ginsborg's Notion of Primitive Normativity', in Christoph Demmerling and Dirk Schröder (eds.), *Concepts in Thought, Action, and Perception* (Oxford and New York: Routledge), 117-38.

Lewis, Michael and Ramsay, Douglas (2004), 'Development of self-recognition, personal pronoun use, and pretend play during the 2nd year', *Child development,* 75(6), 1821-31.

Lewis, Michael, et al. (1989), 'Self development and self-conscious emotions', 60(1), *Child development*, 146-56.

Lichtenberg, Georg (1971), *Schriften und Briefe, II*, ed. W. Promies (Munich: Carl Hanser).

Longuenesse, Beatrice (2017), *I, Me, Mine: Back to Kant, and Back Again* (Oxford: Oxford University Press).

Martin, M.G.F. (1995), 'Bodily Awareness: A Sense of Ownership', in José Luis Bermúdez, Anthony Marcel, and Naomi Eilan (eds.), *The Body and the Self* (Cambridge, MA: MIT Press).

McDowell, John (2009), *Having the World in View* (Cambridge, Mass.: Harvard University Press).

Moran, Richard (2001), *Authority and Estrangement* (Cambridge, MA: Harvard University Press).

Musholt, Kristina (2015), 'Thinking About Oneself', (Cambridge, MA: MIT Press).

O'Brien, Lucy (2007), *Self-Knowing Agents* (Oxford: Oxford University Press).

Pea, Roy D (1982), 'Origins of verbal logic: Spontaneous denials by two-and three-year olds', *Journal of Child Language,* 9(3), 597-626.

Peacocke, Christopher (2014), *The Mirror of the World: Subjects, Consciousness and Self-Consciousness* (Oxford: Oxford University Press).

Piaget, Jean (1952), *The Origins of Intelligence in Children*, trans. Margaret Cook (New York: International Universities Press).

Rödl, Sebastian (2007), Self-Consciousness (Cambridge, MA: Harvard University PressRubio-Fernández, P. and Geurts, B. (2013), 'How to pass the false-belief task before your fourth birthday', *Psychol Sci,* 24(1), 27-33.

Sellars, Wilfrid (1963), 'Empiricism and the Philosophy of Mind', in *Science, Perception and Reality* (London: Routledge and Kegan Paul).

Strawson, P.F. (1974), 'Imagination and Perception', in *Freedom and Resentment* (Oxford and New York: Routledge), 50-72.

Sugarman, Susan (1983), *Children's Early Thought: Developments in classification* (Cambridge: Cambridge University Press).

Treisman, Anne (1996), 'The binding problem', *Current opinion in neurobiology,* 6(2), 171-78.

Wallace, R. Jay (2015), 'The fugitive thought: Blackburn on Reasons', in Robert N. Johnson and Michael Smith (eds.), *Passions and Projections: Themes from the Philosophy of Simon Blackburn* (Oxford: Oxford University Press), 246-66.

Wimmer, Heinz and Perner, Josef (1983), 'Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception', *Cognition,* 13 (1), 103-28.