A popular reading of Kant’s criticism of rationalist doctrines of the soul: Kant had the groundbreaking insight that those doctrines confuse a special kind of awareness (“consciousness of oneself as subject”) with the awareness of a special kind of object: a soul, distinct from the body.

However, a rejoinder against Kant: there is an object one is aware of, in being aware of oneself as subject: oneself as an embodied entity.

BL’s view: Kant is right to maintain that there is a kind of self-consciousness that is only consciousness of oneself as the agent of a certain kind of rational unity among one’s mental contents.

The plan for today’s lecture: explore Kant’s insight into the correlation between consciousness of oneself as subject (grounding the use of ‘I’ in ‘I think’, on the one hand); and consciousness of oneself as a particular spatio-temporal, embodied entity, on the other hand; clarify our notions of person and self, and their respective relation to our use of the first person pronoun.

I-The “Paralogism of Personhood.”

I-1. The general form of a paralogism:

A classic syllogism

\[ \begin{align*}
  MP & \text{ All S’s are P} \\
  mP & \text{ All Rs are S} \\
  C & \text{ So, all Rs are P} \\
  \text{Valid} 
\end{align*} \]

A paralogism

\[ \begin{align*}
  MP & \text{ All S’s are P} \\
  mP & \text{ All R’s are S*} \\
  C & \text{ So, all R’s are P} \\
  \text{Invalid} 
\end{align*} \]

I-2. The paralogism of personhood:

\[ \text{MP} \text{ What is conscious of the numerical identity of itself at different times, is to that extent a person.} \]

\[ \text{mP} \text{ I, as thinking [the soul], am [is] conscious of the identity of myself [itself] at different times.} \]

\[ \text{C So} \text{ I, as thinking [the soul], am [is] a person.*} \]

Removing the equivocation on the middle term:

\[ \text{MP}’ \text{ What is conscious of what is, in fact, its own numerical identity at different times, is, to that extent, a person.} \]

\[ \text{mP}’ \text{ I, as thinking, conceive of myself (have a concept of myself) as numerically identical at different times} \]

\[ \text{C’ So} \text{ I, as thinking, am a person.} \]

* Kant’s actual formulation of the Paralogism: “Something that is conscious of its own identity at different times is, to that extent, a person. Now the soul is etc. So it is a person.”

(“Was sich der numerischen Identität seiner Selbst in verschiedenen Zeiten bewußt ist, ist sofern eine Person. Nun ist die Seele usw. Also ist sie eine Person.”)
The inference is invalid. There is no middle term.

I-3) Thinking the identity at different times of the referent of ‘I’ in ‘I think’

T1 “The identity of the consciousness of myself in different times is therefore only a formal condition of my thoughts and their connection, but it does not prove at all the numerical identity of my subject, in which – despite the logical identity of the I – a change can go on that does not allow it to keep its identity: and this even though all the while the identical-sounding ‘I’ is assigned to it, which in every other state, even in the replacement of the subject, still keeps in view the thought of the previous subject, and thus could also pass it along to the following one.” (A363)

T2 “Assuming substances ... in which representations, together with consciousness of them, flow from one to another, a whole series of such substances may be thought, of which the first would communicate its state together with its consciousness to the second, which would communicate its own state together with that of the previous substance to a third substance, and this in turn would share the states of all the previous ones, together with its own states, because these states would have been carried over to it, together with the consciousness of them. And in spite of this it would not have been the very same person in all these states.” (A363-64n)

T3 “Only because I can comprehend the manifold [of my representations] in a consciousness do I call them all together my representations; for otherwise I would have as multicolored, diverse a self as I have representations of which I am conscious.” (B134)

I-4. My proposal: by “self” (Selbst) Kant means the unifying structure of the contents of my representational states, in virtue of which I have a concept of myself (the thinker) as the agent of that structure. Unless I comprehended the contents of my representational states (what he calls “my representations”) in one whole, I would have “as diverse and multicolored a self,” namely as diverse and multicolored a structure of my representational mental life, as I have representations of which I am conscious – as it were, a particular “self,” structure of mental life, attached to each particular representational content. But I do have a whole of strongly connected representations: concepts, constrained in their meaning by their connection to other concepts; and a whole of perceptual representations of objects related to one another in one spatio-temporal order. So I do have a “self,” a unifying structure of the contents of my representational states, and I thus also have a concept of myself (the referent of ‘I’) as the agent of that unity.

I-5 ‘I’ refers to “the I or he or it (the thing) that thinks” (A346/B404). [Consciousness of] numerical identity of self does not provide any a priori warrant for numerical identity of substance, and thus does not provide any a priori warrant for numerical identity of person.

I-6 Kant on consciousness of the numerical identity at different times of a spatio-temporal entity:

T4 “If I want to cognize through experience the numerical identity of an external object, then I will attend to what is persisting in its appearance, to which, as subject, everything else relates as a determination, and I will notice the identity of the former in the time in which the latter changes.” (A362)

T5 “If I consider myself from the standpoint of another (as an object of his outer intuition), then it is this external observer who originally considers me as in time; [...] Thus from the I that accompanies – and indeed with complete identity – all representations at every time in my
conscio\-usness, although he admits this I, he will still not infer the objective persistence of myself [meiner selbst]. The identity of consciousness of myself at different times is therefore only a formal condition of my thoughts and their connection, but it does not prove at all the numerical identity of my subject.” (A363)

I-7) A possible outcome: a person is an entity that is conscious of what is, in fact, its own numerical identity through time as a spatio-temporal entity: a living being, endowed with psychological states. That consciousness depends on unity of apperception, namely reason-giving unity of consciousness of the contents of mental states, which makes it possible to accompany them with the thought ‘I think.’

I-8) Kant’s surprising back-pedaling: “Meanwhile the concept of personhood just like the concepts of substance and of the simple, can remain, insofar as it is merely transcendental, i.e., a unity of the subject that is otherwise unknown to us, but in whose determinations there is a thorough going connection of apperception) and to this extent this concept is also necessary and sufficient for practical use; but we can never boast of it as an extension of our self-knowledge through pure reason, which dazzles us with the uninterrupted continuous duration of the subject drawn from the mere concept of the identical self.” (A365-66)

II- Kant’s Paralogism of Pure Practical Reason

II-1. The I “in sensu stricto” and the soul in Metaphysik L1 (late 1770s)

T-6 “When we speak of the soul a priori, we say of it nothing more than what we can derive from the concept of the I, and insofar as we apply the transcendental concepts to that I.” (AA28-1, p.266)

II-2. Kant’s derivation of the spontaneity of the soul from our use of the concept ‘I.’:

T-7 “Now comes the question ... whether one can attribute to the soul absolute spontaneity. [...] Here again the I must help. [...] When I say: I think, I act, etc., either the word I is falsely appealed to [das Wort Ich wird falsch angebracht] or I am free. If I were not free, I could not say: I do it; but I would have to say: I feel in me a desire to do that someone has elicited in me. But when I say: I do, this means freedom in the transcendental sense.” (AA28-1, p.267-69)

II-3. Kant’s paralogism of spontaneity:

M Insofar as something is conscious of its own self-determination, it is transcendentally free (an uncaused cause).

m I, as thinking the moral “I ought to,” am conscious of my own self-determination.

C So I, as thinking the moral “I ought to,” am transcendentally free.

The ambiguous middle term is: “conscious of (one’s own) self-determination.” There is in fact no middle term. The syllogism is invalid.

II-4 Kant’s possible response to this putative refutation: 1) He does not claim that we know ourselves to be absolutely free. 2) Nevertheless believing that we are free in a metaphysical sense is necessary for any notion of moral accountability to gain traction. This is why a concept of a person that is not the empirical concept of a person as a living entity subject to empirical causal laws could, indeed should remain, as necessary and sufficient for practical use (sufficient, that is, only if one adds to the psychological sense the moral sense of “person”).
II-5. Two responses to Kant’s response: 1) the notion of self-determination as it appears in the minor premise may be dissociated from the ontological notion of uncaused-cause as it appears in the major premise. 2) Kant’s criticism of the paralogism of personhood can serve as a blueprint for criticizing what I call his “paralogism of spontaneity.”

III- Textual support for a concept of a person, in the moral sense: a living being endowed with (moral) personhood.

T-8 “The worth of any object to be acquired by our action is always conditional. Beings the existence of which rests not on our will but on nature, if they are beings without reason, still have only a relative worth, as means, and are therefore called things, whereas rational beings are called persons because their nature already marks them out as an end in itself; that is, as something that may not be used merely as a means, and hence so far limits all choices (and is an object of respect).” (Groundwork: AA4: 429)

T-9 “That a human being can have ‘I’ in its representations elevates it infinitely above all other living beings on earth. It is thereby a person and, thanks to the unity of consciousness in all alterations that may affect it, it is one and the same person, that is to say a being that is different, in rank and value, from things, with which one can do whatever one wants.” (Anthropology from a Pragmatic Standpoint: AA7, 127)

T-10 “Person is that subject who is accountable for her actions (dessen Handlungen einer Zurechnung fähig sind). Moral personality is thus nothing else than the freedom of a rational being under moral laws (whereas psychological personality is only the capacity to become conscious of the identity of one’s own existence in one’s different states); from which it follows that a person is subordinated to no other laws than the laws he gives himself, either alone, or also with others.” (Metaphysics of Morals, AA6, 223).

T-11 “Subjectively the degree of accountability (imputabilitas) of the action is to be evaluated by taking into account the magnitude of the obstacles that had to be overcome. – The bigger the natural hurdles (of sensibility) and the smaller the moral obstacle (of duty), the more praiseworthy is the good action. For instance, if I save from a great misery a human being who is a complete stranger to me, at the cost of a considerable sacrifice for myself. On the contrary, the smaller the natural hurdle and the greater the obstacle on grounds of duty, the bigger the trespass for which one is accountable (open to blame).” (Ibid., AA6, 228).

BL’s suggestion: it remained open to Kant to say that being a person in the moral sense depends on two equally indispensable components: 1) Being an empirically determined, persisting entity, conscious of its own numerical identity through time (person “in the psychological sense,” which has for its necessary condition the unity of apperception and thus the capacity to think ‘I think’). 2) Having the capacity to assign to oneself the moral “I ought to” (person “in the moral sense”).

Concluding remarks
I offered arguments for a naturalized version of Kant’s notion of a person, both in the psychological and in the moral sense.

The notion of person I am advocating is not Peter Strawson’s. It is closer to Harry Frankfurt’s, in “Freedom of the Will and the Concept of a Person.” But it is richer than Frankfurt’s.

More needs to be said to account for the structure of mental life grounding the use of ‘I’ in the moral ‘I ought to.’