Whose Constituent Power Is It?

Marco Santambrogio

Abstract. Ferrara maintains that constituent power – i.e., the power to issue a constitution – needs a sovereign actor endowed with singular intentionality, because neither a law nor a constitution can establish itself. At least fifty-three actual constitutions around the world claim authorship on behalf of "the people" for their articles. The question arises: is that actor – the people – an actual subject or, as argued by Juergen Habermas and Hans Kelsen, a merely fictional one? An argument is presented to the effect that it cannot but be fictional. The argument draws on a celebrated result due to Condorcet and generalised by Kenneth Arrow, showing that a plurality of rational subjects, such as a people, is bound to be sometimes irrational, in so far as it harbours cyclical preferences. This is a serious obstacle to holding that an actual people could be endowed with intentionality, which presupposes the possession of, among other things, will, memory, preferences and also rationality.

Keywords: constituent power, rationality, reasonableness, Arrow's theorem, fictional entities

Sovereignty Across Generations is a book of many merits, but the wealth of doctrine and ideas that Alessandro Ferrara offers to clarify, if not solve, some of the fundamental problems of political liberalism is impressive. I would not be able to comment on it in its entirety, so many issues are addressed and so vast is the relevant literature. Fortunately, I have been asked to comment on only one chapter, the fourth, and even of this I will select only one theme: how should we conceive of those

peoples whose deliberations so many constitutions around the world represent themselves as the product?

The chapter P*olitical* L*iberalism and* '*the* P*eople*' opens with a challenging statement due to Carl Schmitt, which Ferrara fully subscribes to: constituent power needs a subject to exercise it, because neither a law nor a constitution can establish itself. The assertion seems to be borne out by numerous existing constitutions that, in their preambles or first articles, refer to the subjects who would be their authors respectively as "the people". The U.S. Constitution, which at the very beginning identifies itself as the deliberation of "We the People", is but one example among many.

It is not immediately obvious that this statement is true, and potential counterexamples immediately come to mind. Are there not customs that are not attributable to any particular subject and yet have legal value – at least as precedents? Are there not financial markets that do not constitute a subject and yet determine the political and even legislative choices of a country – indeed, of many countries?

Ferrara rejects these alleged counterexamples. Undoubtedly, he writes, one cannot impute subjectivity to financial markets. Markets are "mere aggregates of individual preferences". It is indeed true that we speak of the actions and reactions of a market, and an action properly so called (as distinct from, for example, a simple involuntary motion) always presupposes an agent endowed with intentionality. But this is only a figurative way of expressing it: in reality the "actions and reactions" of markets are the simple results of the actions (these in the proper sense) of countless individuals converging while acting independently of one another. When, on the other hand, we attribute constituent power and actions such as that of enacting a constitution to a people, we always assume that there are shared deliberations, exchanges of reasoning among individuals, consultations that ultimately bring about decisions for which the whole collectivity bears responsibility. Thus, it is not a matter of collective will in a merely "statistical" sense – to use a term that Ronald Dworkin contrasts with "communitarian". Financial markets do not act politically, they do not choose one policy in preference to another at the end of a conscious decision procedure binding each of its members: that is why they have no constituent power. Constituent power requires "a sover-

eign agent endowed with individual intentionality", and the "political conception of the people" (Ferrara and Michelman's conception) assumes that peoples instead have it.[1]

I provisionally concede this point to Ferrara. But now a new question arises: is the singular intentionality presupposed in the author of a constitution (because exercising constituent power is undoubtedly an action and not a mere "statistical" regularity) that of a real subject or is it instead that of a fictional, purely imaginary entity that has the same ontological reality as literary characters and other entities that are perfectly respectable but lack intentionality and existence independent of human imagination, such as the metric system?

All of us human beings can assume that we have individual intentionality and subjectivity, which is a rather complex thing and definitely poorly understood by philosophers and psychologists. But the intentionality and subjectivity that we would like to impute to a collectivity such as the people are even more complex and it is not at all clear that they can be attributed to anything other than a flesh-and-blood individual except by pretence. This is the thesis I want to argue. Ferrara reminds us that Juergen Habermas and Hans Kelsen have argued in favour of the fiction hypothesis.[2] I aim to present an independent argument to argue that it could not be otherwise.

I will focus on the kind of subjectivity that should be attributed to the people as the subject of a constitution. I will try to argue that, in reality, there is no subjectivity other than individual subjectivity. In other words, there are no subjects other than individuals. Therefore, the people cannot exist as a real collective subject. The many constitutions in the world that refer to "the people" refer to a fictional entity.

Ferrara laments that contemporary political philosophers – including Rawls – simply assume that a people already exists: "We are never

---

[1] "Constituent power, instead, needs a sovereign actor endowed with singular intentionality" (137).

[2] About Kelsen, for example, Ferrara writes: "For Kelsen, the people and its constituent power are postulates – in other words, fictions. The people, putatively exercising constituent power, 'is not, as is often naively imagined, a body or conglomeration, as it were, of actual persons. Rather, it is merely a system of individual human acts regulated by the state legal order'" (94).

told how a people comes into existence as a people". He adds, "It is an unfortunate lacuna of contemporary liberal-democratic theory that this *idolum fori* persists, according to which the formation of the people is understood either as a historical contingency immune from all judgment about its legitimacy or as a retrospective projection, transcendentally 'necessitated' by an accepted constitution" (139). For what I intend to argue – that is, that necessarily the people is a fictional entity – there can be no answer to the problem of how a people constitutes itself as a people. Nor is its existence a historical contingency. "Retrospective projection" is therefore the only viable alternative. For that matter, is there really any need for anything else? Rawls, in A T*heory of* J*ustice*, presents a thought experiment that asks us to imagine our own judgment and that of other individuals – individuals, not collective entities endowed with intentionality – about his two principles of justice: it is this judgment that should convince us that a constitution comprising those two principles is not only just, but will appear just to its citizens – a necessary condition for its stability. If Rawls is right, nothing else is needed to make such a constitution our own.

I now come to the argument. As far as I can tell, philosophers, psychologists and neuroscientists are still far – how far, I do not know – from having clear ideas about what an individual subject is. Each of us is convinced that we are subjects in our own right, but when it comes to attributing the property of being a subject to other individuals (human or otherwise) we are uncertain about the criteria. Some conditions clearly appear necessary. We certainly would not say that a being without a mind can be a subject. And to have a mind it seems necessary to be capable of intentionality: at least beginning with Brentano, intentionality has been taken as the hallmark of the mental. In addition to this, it also seems necessary to be able – at least at a minimal level – to act, to formulate more or less long-term plans of action and projects. So the will is indispensable, and in addition, in order to act, one must also have preferences among the different courses of action available. Perhaps some ability is required to set for oneself conditions to be met in the future and thus to have some notion of the passing of time and to keep track of it. Some conception of oneself seems equally necessary. Is that all? Perhaps not. Perhaps one should add an ability to conceive of the presence (or at least the possibility) of other subjects distinct from, but similar to,

oneself, and an idea, however vague, of the difference between viewing from a particular vantage point and viewing objectively or (if one can say so) from nowhere (Nagel 1986).

I do not intend to go into the details of these conditions – although it seems to me that an entity that is no longer individual, but plural like a people, is very unlikely to be able to satisfy all of them.[3] I am interested in only one condition which I have not yet mentioned and which is perhaps no less arduous than all the others: in order to be able to attribute to someone some form of subjectivity, it is a necessary condition to be able to attribute to her or him at least a minimum of rationality. The difficulty in stating this condition lies in the fact that it is quite hard to draw a clear line of separation between actual irrationality and the simple difference of views and opinions: one must take great care not to mistake extreme and unusual, but consistent, views for irrationality. I know of no effective and safe criteria for doing so. It seems certain, however, that at least in extreme cases, when we realise that we are dealing with beings whose behaviours we are unable to understand and with whom we just cannot communicate – for example, because we have no idea how we could convince them of what seems obvious to us and which we are convinced should be obvious to them as well – we are unwilling

[3] Ferrara cites an eloquent passage in which Schmitt identifies the modern subject of constituent power with the people or the nation: the nation "denotes, specifically, the people as a *unity* capable of political action with the *consciousness* of its political distinctiveness and with the *will* to political existence, while the people not existing as a nation is somehow only something that belongs together ethnically or culturally, but it is not necessarily a bonding of men existing politically" (italics mine). The passage is very clear: Schmitt postulates that the people (or nation) is a subject (a) unitary, (b) endowed with consciousness and (c) will. This is a postulate for which no justification is given. Point (b) I will deal with in a moment. Now I only observe that point (a) takes on a very precise meaning in a context such as Nazi Germany: it turns dissenters into traitors to the state. Ferrara observes in this regard, "Any controversy, in any realm of institutional or social life, could become the vehicle and focus of a 'political' opposition of friends and enemies" (117). It is known that the enemies, according to Schmitt, are not the opposition parties and those who vote for them: they are those who reject those values and "commitments" without which the state would cease to exist. On this point I do not think there is any possible mediation between Schmitt and Rawls' pluralism, especially in *Political Liberalism*.

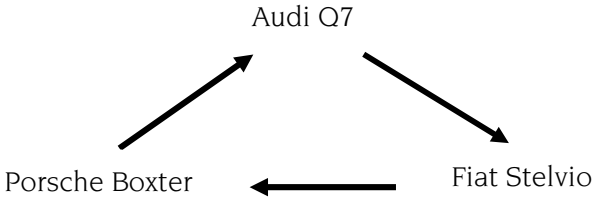
to attribute to them true subjectivity. It naturally happens to all of us to think that someone, even among our closest acquaintances and friends, is occasionally irrational and behaves in a way that is incomprehensible to us, but in all such cases we are convinced that, with time and sufficient goodwill, either we could sooner or later come to convince our interlocutor of her irrationality and get her to admit her error or we would come to understand her behaviour ourselves and realise that her point of view is simply different from ours. But in extreme cases, when we think we are dealing with truly irrational beings who are incomprehensible to us, we cannot attribute subjectivity to them: we can only try to control or cure them.

Is it possible for a real collective entity, such as a people, to satisfy this condition of rationality? Ferrara had assumed, with Schmitt, that the people exercising constituent power possesses "sufficient intentionality to be the author of a constitution" and "the capacity not only to act politically but also to shape its own political conduct". It should therefore be a subject in the full sense of the term and thus also be rational or not obviously irrational. Ferrara does not see substantial differences between an individual and a collective subject. Indeed, in order to explain how a people (in the sense of *ethnos*) can self-constitute itself into a political subject with constituent power (and thus transform itself into a *demos*), he appeals to some theorists of the self-constitution of the subject – from Michel de Montaigne to Christine Koorsgard, to Harry Frankfurt, to Charles Larmore – and adds, "Although their goal is generally *individual* self-constitution, their teachings also apply to *collective* self-constitution" (152). The point I intend to make is all here: in the transition from an individual to a collective entity we necessarily lose not only the guarantee, but the very possibility of rationality. It is at least difficult, therefore, to argue that there is a real subject to be entrusted with constituent power: only by pretence can we refer to the author of a constitution as a subject. In other words, the subject who has constituent power must be a fictional or imaginary subject. (Of course, it cannot be argued without circularity that the constitution itself manifests the intentions and will of the people and, if it is consistent, also demonstrates the consistency and rationality of the people who are its author.)

The argument I want to make is far from new, but I do not know if it has ever been used for the conclusion I am interested in. The starting

point is a simple observation: all rational people, if they prefer A to B and B to C, also prefer A to C and not vice versa, for whatever A, B and C. It is easy to see that this is a necessary condition of rationality.

Indeed, suppose someone you know expresses his preferences in matters automotive and tells you that he prefers the Audi Q7 to the Porsche Boxter and then that he prefers the Porsche Boxter to the Fiat Stelvio. But then between the Audi Q7 and the Fiat Stelvio he prefers the latter. In other words, he has circular preferences that we can graphically represent like this:



Suppose also that the fellow owns a Fiat Stelvio and you happen to have both a Porsche Boxter and an Audi Q7. Hearing his preferences, you offer him your Porsche Boxter in exchange for his Fiat Stelvio and only a thousand euros. Well pleased, he accepts the exchange and gives you the Fiat. Now he owns the Porsche and you own the Fiat, the Audi and an extra thousand euros. He is still not satisfied, however: as we know, he prefers the Audi to the Porsche. And you propose another exchange: your Audi for his Porsche plus another thousand euros. Following his preference, he accepts and you end up with the Porsche, the Fiat and two thousand euros more. He has the Audi and two thousand euros less. Is he satisfied? Not yet: he has revealed to you that he prefers the Fiat Stelvio to the Audi Q7. In a fit of generosity you offer him another exchange, immediately accepted: your Fiat Stelvio for his Audi Q7 plus one thousand euros. Now you find yourselves exactly in the initial situation – he has the Fiat, you have the Porsche and the Audi – except that you now have three thousand euros more and he has three thousand euros less. If you wanted to, and if he had not yet learned his lesson, you could start all over again exchanging cars, each time with his modest outlay of a thousand euros in your favour. You could go on forever – he always dissatisfied, you always richer at his expense. But of course

human subjects endowed with reason may have moments of temporary irrationality but sooner or later they recognise their error (especially if it is explained to them) and correct themselves. That is, they recognise that it is irrational to have circular preferences. We have thus ascertained the starting point of our argument.

Do supposed plural subjects – peoples – behave in the same way? More than two hundred years ago, Marquis Nicolas de Condorcet proved that no, peoples can have circular preferences and there is no way to change their minds. They are therefore irrational in the sense we are interested in. Suppose there are three voters (or three thousand or three million – it makes no difference) and three candidates to choose from: A, B and C. The voters are all rational individuals and have no circular preferences. They also have very marked preferences (perhaps the candidates are enormously different from each other) and are unwilling to change their minds. We represent their preferences graphically in this table:

Voters	*First choice*	*Second choice*	*Third choice*
Voter 1	A	B	C
Voter 2	B	C	A
Voter 3	C	A	B

The first voter prefers A to B and B to C. Since she is rational, by hypothesis, and has no circular preference, she also prefers A to C. Similarly for the other two. Suppose now that each voter expresses her preferences by voting. We can define the preferences of the totality of voters – the people – as those that result from a vote (or a series of votes, depending on the way one votes): we say that the people prefer one candidate to another if a majority of voters express themselves (or would express themselves) in favour of the former. It is easily seen that in the case represented by the table the people have circular preferences: a majority of voters prefer A to B (Voter 1, Voter 3), a majority of voters prefer B to C (Voter 1, Voter 2), a majority of voters prefer C to A (Voter 2, Voter 3).

Condorcet worried about the (decidedly counterintuitive) result whereby, in case of an election of a representative from among A, B, and C, any outcome of the vote would displease a majority of the voters. We are interested in the simple fact of circular preferences that

cannot be remedied, unlike the individual case, simply by bringing it to the attention of the collective subject (the people): we have said that the voters are not willing to change their minds about the candidates, who are strongly characterised, and therefore the people will not be willing to change their minds either. And if we wanted to consider it as a collective entity to which we could attribute authorship of a constitution, we would have to admit that the people would be irredeemably irrational, to the point where it could not be considered a subject similar to individual subjects.

Condorcet's result was generalized by Kenneth Arrow and earned him the Nobel Prize in Economics for the year 1972 (Arrow 1951). His impossibility theorem is a major result concerning the fundamental concepts of political theory, although it would be wrong to draw negative conclusions about the very possibility of democracy. Arrow himself summarised its political significance as follows, "Most systems are not going to work badly all of the time. All I proved is that all can work badly at times." To the contrary, it seems to me that the theory that the author of a constitution should be a real collective subject – the people – is seriously damaged by the argument.

It could perhaps be argued that the people only occasionally are irrational, in much the same way that real individuals are irrational. If the occasional irrationality of individuals is not a sufficient reason to deny them the quality of subjects, why should we deny it to the people? The answer is twofold. First, it can be shown that cases of circular collective preferences are relatively rare in ordinary political elections, when there are many voters and few candidates. (A mathematical theory has been developed that deals with these phenomena and quantifies them exactly.) But if the people, or any other collective entity, were a subject, there would be an indefinite number of occasions when they are called upon to express preferences, and on a far greater percentage of these occasions their preferences would be circular. For example, if we asked the set of guests at a wedding to vote to choose the people with whom to share a dinner table, we would have as many voters as candidates. As we know, leaving the majority of guests unsatisfied is unfortunately a very real possibility.

Second, it is true that we are all occasionally irrational, but, unfortunate as it is, no one has ever made a big deal out of it: why should we get

over-worried in the case of collective subjects? There is a difference, which I have already alluded to, between our individual irrational behaviours and those of the supposed subject, the people. We know how to correct our mistakes – at least, when someone points them out to us. That is why we cannot hope to get rich at the expense of our car-loving and occasionally confused friends. But in a case of circular preferences the people cannot correct themselves if their members do not change their minds. The percentage of cases of irrationality is not in question here: it only takes a single case to declare that someone is hopelessly irrational. Suppose someone is able to add any two numbers without making mistakes but for some mysterious reason cannot calculate 2 + 3 at any cost and despite all our explanations goes on endlessly to say that 2 + 3 = 1729. Or even worse, suppose we know that in only one case does he make such an error, but we don't know what that case is. Would you take him as an accountant? ("Come on, it's a very small error in only one case!").

The conclusion seems forced to me that the people as a collective subject must be a fiction, like the present king of France and the phlogiston.[4] After some individuals formulated the text of a constitution, after it was put to a vote and approved (presumably by a majority, because unanimity seems unattainable), we can retrospectively pretend that the people as a whole were its concordant author and by that themselves, exercising intelligence and will, made commitments for themselves in the future and constituted themselves as a people (*demos*). But this is precisely only a pretence.

Even in our case when, instead of electoral systems, we are dealing with the rationality of the subject who is the author of a constitution, the conclusion is not troubling (except for the theory that it is the actual people who are the author of modern constitutions). Indeed, what benefit could we have expected if it had been established that the people is a real subject and not merely imaginary? Ferrara cares to distinguish the people exercis-

[4] Of course, nothing prevents us from imagining the present king of France, but Russell's theory of descriptions, which has much authority among contemporary philosophers, allows us to dispense with even this fictitious individual (the real one is obviously nonexistent) while still recognising that utterances such as 'The present king of France is bald' and 'The present king of France does not exist' are not truth-valueless (Russell 1905).

ing constituent power from the set of individuals interacting in a market. In a market the interests of individuals are conflicting or at any rate not all jointly satisfiable. Each person thinks only of himself. This would not be the case in a constitutionally regulated society that is – to use a characterisation by John Rawls – a cooperative venture for mutual advantage, even though it is typically marked by a conflict as much as by an identity of interests.[5] To account for this (partial) convergence of purpose and action among the citizens of a state it seems to Ferrara necessary to assume that some form of unity exists among them. He claims that this unity can be realised only if the citizens form a single subject, a people that acts politically, in the sense that it chooses certain policies in preference to others, in a decision-making process that binds each citizen and is capable of sufficient shared intentionality to be the author of a constitution.[6]

I, however, fail to see how this plural subject – the people – can serve to satisfy that need. In fact, I think it is an impediment. If it were a subject endowed with intentionality, in what relationship would it stand with the other subjects, the individual citizens who are part of it? Necessarily they should be different subjects and external to each other. Of course, citizens are part of the people but we cannot say – it would make no sense – that individual intentionalities are part of the collective intentionality. One intentionality (one mind) cannot be part of another. And two distinct intentionalities can get along just as well as they can conflict. If we wanted to insist that the collective one somehow realises the concord of citizens and does not conflict with them, we would have to postulate another intentionality that 'includes' both the collective and individual ones. Once again the Third Man proves to be a powerful, and lethal, argument.

---

[5] "Then, although a society is a cooperative venture for mutual advantage, it is typically marked by a conflict as well as by an identity of interests" (Rawls 1971, 4).

[6] "[A] people can *act politically*, in the sense of publicly choosing, in law-regulated ways, one policy over another, thus prioritising one collective end aver another in a decision-making process that binds, or at least significantly affects, every member" (143) and "our 'political conception of a people' must assume that each of the peoples whose constitution was cited above is capable of enough shared intentionality to author it" (143).

Moreover, the opposition between the people and the market can also be questioned, and here again we have much to learn from a debate at the beginning of political philosophy. "It is best that the whole state should be as much of a unity as possible" – Socrates says at one point, in Plato's R*epublic* (R*epublic* 422e ff, 462a ff.) Aristotle's main objection to this view of the state is that it overvalues unity and uniformity.

> The state consists not merely of a plurality of men, but of different kinds of men; you cannot make a state out of men who are all alike. Consider in this connection the difference between a state and an alliance: the purpose of an alliance is military assistance, and its usefulness depends on the amount of that assistance, not on any differentiation in kind; the greater the weight, the greater the pull. [...] On the other hand, constituents which must form a single unity differ in kind. Hence, as I have already stated in my E*thics*, 3 it is *reciprocal equivalence* that keeps a state in being. (T*he* P*olitics*, 1261a22) ['Reciprocal equivalence' – *to ison to antipeponthos* (N*icomachean* E*thics*, V) – is the principle of mutually supporting diversity of function, whereby (to take a simple example) a shoemaker provides shoes for a bakery who provides bread in return] (Aristotle 1962, 103).

And little beyond that:

> Undoubtedly there must be some unity in the state, as in a family, but not total unity. On the road to gradual unification, at some point the state, if it does not fail altogether as a state, is endangered and becomes worse. It is as if one wants to reduce harmony to unison and rhythm to a single beat. As I have already said, a state is a plurality that must let unity be produced by education (Aristotle 1962, 116).

In a Greek city-state, individuals, families, tribes, freemen and slaves, citizens and foreigners, who had different and competing skills, functions, and economic interests, met and clashed and supported each other. But doesn't the same hold true in a marketplace? Ferrara contrasts the "independent but convergent" actions of individuals acting in a marketplace with the exchange of reasons and consultations that precedes

deliberations among the members of a people.[7] But it does not seem to me that the actions of individuals in a market are independent: supply and demand evidently take into account each other and coordinate.

Observe then that the assumption that the people would be a real subject endowed with intentionality, and not a mere fiction, directly conflicts with Rawls' own critique of all forms of utilitarianism in one of the opening paragraphs (§5) of A T*heory of* J*ustice*:

> It is a conspicuous feature of the utilitarian position on justice that the way in which this sum of satisfactions [e.g., of rational desires] is distributed among individuals matters no more (except indirectly) than the way in which an individual person distributes his satisfactions over time. [...] The most natural way to arrive at utilitarianism (though of course it is not the only way to do so) is to adopt for society as a whole the principle of rational choice for a single person. Once this point is recognized, one immediately understands the place of the impartial spectator and the insistence on sympathy in the history of utilitarian thought (Rawls 1971, 26-27).

We know what Rawls' critique of this position is:

> This position on social cooperation is the consequence of the extension to society of the principle of choice for a single person and then, to make this extension work, of the merging of all persons into one through the acts of imagination of the sympathetic impartial spectator. Utilitarianism does not take the distinction of persons seriously (Rawls 1971, 27).

What else does the hypothesis of the people as real individuals endowed with intentionality amount to, if not precisely the fusion of all people into one and the refusal to take seriously the distinction of per-

[7] "When we attribute to a market, or to a social system, positive or negative reactions to circumstances, we are really using those terms as a shorthand for what millions of individuals, independently but convergingly, do. When instead we attribute to a *people* positive or negative reactions to possible options, we imagine that some sort of inter-individual exchange of reasons – however minimal, anonymous, or impersonal – does take place, a minimal consultation according to some mechanism that in the end, if only via simple majority rule or acclamation, selects one or the other option as *imputable to the whole collectivity*" (143).

sons? Note that Rawls' critique primarily strikes at the assumption that the unbiased onlooker can even imagine the people as one person, but *a fortiori* it also strikes at the stronger assumption that the people is not just a fictitious entity like the current king of France and the phlogiston, but has real existence of its own, either as *ethnos* or *demos*.

Ferrara says that "Rawls unreservedly sides with Schmitt in affirming the non-fictional, not merely retroactively or 'constructed', existence of constituent power as the power of a subject — the people, in democratic theory — 'to establish a new regime', to bring into being 'a framework to regulate ordinary power, and to articulate in a constitution its political ideal 'to govern itself in a certain way.'" (123) I do not read the passages from Rawls quoted by Ferrara (on pages 231 and following of *Political Liberalism*) in the same way. Indeed, I find this other passage little further: "The idea of right and just constitutions and basic laws is always ascertained by the most reasonable political conception of justice and *not by the result of an actual political process*." (italics mine, Rawls 1993, 233)

But above all, I do not see why Rawls should resort to a real subject, or even an imaginary subject, to make it the author of a constitution. Consistent with the last quoted passage that states, in essence, that the judgment on the justice of a constitution is of the same kind as the judgment on the reasonableness of a philosophical conception or theory, and not the verdict of a vote, Rawls resorts to a mental experiment. In *A Theory of Justice*, the thought experiment is crystal clear: we are asked to imagine ourselves in the hypothetical situation of the original position and to verify that, in that situation, the principles of justice formulated by the theory would appear acceptable to us. The original position in a certain way forces us to be impartial and not to privilege the social position in which in fact each of us finds ourselves. This impartiality is the same thing as reciprocity and reasonableness.[8] If Rawls is right, if each of us is convinced that in the original position she herself would accept those principles of justice, what else is required for a constitution that respects them to be embraced by all citizens (more realistically, by almost all) and recognised as just and stable?

[8] See the characterisation of reasonable inclusive doctrines on pp. 58ff. and especially p. 62 of Rawls 1993.

But perhaps Ferrara, who distinctly prefers P*olitical* L*iberalism* to A T*heory of* J*ustice*, thinks Rawls must abandon that thought experiment once he embraces the kind of normativity implied in the second work: "this normativity [of P*olitical* L*iberalism*] cannot be that of 'justice as fairness' which is the result of the original position discussed in A T*heory of* J*ustice*. That interpretation is precluded by footnote 7 of the second lecture in P*olitical* L*iberalism*." (126) Footnote 7 immediately follows this period in the main text, "To see justice as fairness as trying to derive the reasonable from the rational misinterprets the original position". Here is the footnote: "Here I correct a remark in T*heory* [*of* J*ustice*], p. 16, where it is said that the theory of justice is a part of the theory of rational decision. From what we have just said, this is simply incorrect. What should have been said is that the account of the parties, and of their reasoning, uses the theory of rational decision, though only in an intuitive way. This theory is itself part of a political conception of justice, one that tries to give an account of reasonable principles of justice. There is no thought of deriving those principles from the concept of rationality as the sole normative concept. I believe that the text of T*heory* as a whole supports this interpretation" (Rawls 1993, 53).

It seems to me that Ferrara takes footnote 7 as a rejection by Rawls of his earlier (T*heory's*) position, as if the thought experiment of the original position asked us to imagine purely rational subjects who lack a sense of justice and do not recognise the independent validity of other subjects' claims. Instead, I think Rawls in that note meant only that already in A T*heory of* J*ustice* (despite the unfortunate remark on page 16) the reasonable and the rational are complementary ideas and neither can stand without the other. As I have said, the principle of reciprocity (i.e., reasonableness) is imposed on subjects in the original position by the veil of ignorance, which makes it inevitable to acknowledge the validity of other subjects' demands as being on the same footing as ours – simply because their demands might be ours. It is thus not a discovery of P*olitical* L*iberalism* that merely rational, and not also reasonable, agents can be psychopaths if their only interest is in advancing their own welfare.[9] My

[9] "Rational agents approach being psychopathic when their interests are solely in benefits to themselves" (Rawls 1993, 51).

conclusion, then, is that the thought experiment of A *Theory of Justice*, or at any rate a reasoning of the same kind that philosophers use to judge whether a philosophical theory is acceptable – and not the result of an actual vote or other manifestation of the will of a supposed plural subject such as the people – continues to be for Rawls the way in which each of us, and therefore all citizens belonging to a people, could accept a just constitution and feel it as their own.

References

Aristotle (1962), *Politics*, London, Penguin.

Arrow K. (1951), *Social Choice and Individual Values*, New York, Wiley.

Ferrara A. (2023), *Sovereignty Across Generations. Constituent Power and Political Liberalism*, Oxford, Oxford University Press.

Nagel T. (1986), A *View from Nowhere*, Oxford, Oxford University Press.

Rawls J. (1993), *Political Liberalism*, New York, Columbia University Press.

– 1971, A *Theory of Justice*, Cambridge (MA), Harvard University Press.

Russell B. (1905), "On Denoting", *Mind*, vol. 14, n. 56, pp. 479-493.