Brandom on Postmodern Ethical Life:

Moral and Political Problems

Franz Knappik

University of Bergen

Robert Brandom's semantic reading of Hegel's *Phenomenology of Spirit* culminates in an account of an ideal form of reciprocal recognition that we are always already committed to whenever we apply a concept, but that still awaits its concrete realization in an adequate recognitive practice replacing modernity. This future practice will be a "postmodern" form of ethical life that is based on a commitment to "forgiveness," a willingness to offer rational reconstructions of past concept applications. On Brandom's reading, Hegel describes the form of reciprocal recognition that is characteristic of postmodern ethical life in the account of forgiveness that concludes the *Spirit* chapter in the *Phenomenology of Spirit*.

I have various doubts about the textual adequacy of this reading, but in the following, my focus will be on the internal philosophical plausibility of this part of Brandom's reading. I am going to explore practical consequences of Brandom's account of postmodern ethical life (PEL): what does it mean more concretely to live in a community of PEL? What constraints does the recognitive structure of PEL impose on the ways in which we act and the ways in which we understand each other's actions? On the one hand, I am going to argue that it is possible to derive moral norms from Brandom's account of PEL. On the other hand, I attempt to show that this account has implications regarding the way we should understand each

other's actions that are morally not acceptable by PEL's own lights. Among other things, PEL leaves no room for moral critique, and it makes us helpless against covert manipulation and oppression by powerful agents. I therefore will conclude that the normative structure of PEL is inconsistent.[1]

I.   Brandom on postmodern ethical life

I begin with a brief summary of those points from Brandom's reading of the *Spirit* chapter in Part Three of *ST* that are essential background to my subsequent discussion.

On Brandom's reading, Hegel holds that the conceptual content of language, thought, and action is constituted by normative attitudes, through which participants of a discursive practice assign each other authority and responsibility. Hegel, on this reading, analyses such normative attitudes as forms of *reciprocal recognition*. In addition, Hegel holds, on this view, that discursive practices involve normative *meta-attitudes*—attitudes that concern the relation between norms, on the one hand, and first-order normative attitudes, on the other hand. Brandom identifies these meta-attitudes with magnanimity (*Edelmütigkeit*) and pusillanimity (*Niederträchtigkeit*), two complementary viewpoints that Hegel introduces in the chapter on *Culture*, and that re-appear in the final section of *Spirit* (explicitly in the case of pusillanimity, implicitly in the case of magnanimity). Brandom understands magnanimity as a view that holds that there are objectively valid norms that transcend our subjective normative attitudes, while pusillanimity, on his reading, denies such objective norms, reducing instead normativity to our subjective normative attitudes.

Pusillanimity is, on Brandom's account, the characteristic normative meta-attitude of modernity, which replaces the naïve form of magnanimity that had characterized ancient ethical life: a form of magnanimity that simply assumed objective norms as given part of reality. Modernity, however, creates alienation, and it has to be overcome eventually by a restored, "postmodern" version of magnanimity that

integrates the modern insight that norms depend on normative attitudes, by acknowledging a reciprocal dependence between norms and normative attitudes. The recognitive structure that corresponds to this form of magnanimity is that of PEL, the culminating stage in the history of discursive practices, which Brandom finds described in Hegel's discussion of forgiveness at the end of the *Spirit* chapter. (In the following, I will reserve the terms "magnanimity" and "magnanimous" to refer to this specifically postmodern form of the meta-attitude, which is realized in PEL.)

More specifically, consciousness learns in the dialectic of confession and forgiveness that whenever we *attempt* to follow norms—whenever we make doxastic claims, aim to act for reasons, etc.—we are always already implicitly committed to magnanimity, and hence also to PEL (*ST*, pp. 596f.). PEL is defined by a *historical* form of recognition: it is defined by *diachronic* relations of recognition that we create by offering *rational reconstructions* of each other's utterances and actions—in general: of our concept applications (*ST*, pp. 600-610). When we apply a concept in thought, language, or action within a practice of PEL, we are like judges in case law in that we treat earlier concept applications as authoritative, and assume the responsibility to integrate them as far as possible into a narrative that presents a progress leading to our present concept application, while having ourselves authority about how exactly we tell this story. At the same time, we authorize and expect future concept-users to relate in exactly the same way to our own present concept application. Only by thus making our concept applications parts of larger traditions of concept-use do we treat them as instances of norm-following at all, rather than as mere effects of causal circumstances. This is why any attempt to apply a concept in thought or action involves, for Brandom, a commitment to PEL—even though we so far have not yet lived up to this commitment, by establishing a practice that is actually articulated by the historical form of recognition that PEL requires.

II. The structure of rational reconstructions

Let us now look more closely at the structure of rational reconstructions (*ST*, pp. 603-608). On Brandom's account, when a person A gives a rational reconstruction for a concept application by a target person B, A needs to identify three elements:

- a referent R of the concept application;
- the conception that B has of R, and that is the sense through which R is given to B in the concept application; and
- the conception that A has of R, and that is the sense through which R is given to A in the context of her rational reconstruction of the concept application.

The first and the third items in this list are closely connected for A, since A uses her conception of R in order to single out R. By contrast, B's conception of R is presented in the rational reconstruction as an earlier, imperfect conception of R.

Consider, first, the case where the concept application is (part of) an *utterance* and A aims to rationally reconstruct the semantic content of that (part of an) utterance. In Brandom's example, when a modern reader rationally reconstructs Aristotle's applying the word *cheir*, she will see that word as referring, via a sense that is defined by his imperfect conception of a hand, to the same concept of a hand that she refers to via her own conception of a hand, a conception that she takes to capture that concept more adequately (*ST*, pp. 604f). It is crucial to such a rational reconstruction that it identifies a *shared* concept that (according to this reconstruction) both the reconstructor and the target agent treat as authoritative. In Brandom's words, rational reconstructions arrange "past conceptions" of concepts

> into an expressively progressive tradition of applications of a concept that is seen as having been all along already in play as the norm users of that concept were binding themselves by in making judgments and endorsing purposes […]. (*ST*, p. 606)

In his discussion of Hegel's notion of forgiveness, Brandom makes it clear that he understands rational reconstructions for *actions* in an exactly parallel way. "The

magnanimous forgiving recollector," writes Brandom, "must discern an implicit norm that governs the development of the deed." (*ST*, p. 744) Brandom identifies this implicit norm with the "intention" or "*Absicht*," the agent's plan for the action, which normally includes both a goal and means for realizing that goal and which governs the action as it unfolds over time. So, when A rationally reconstructs an action of B's, she identifies

- the intention that governs the action
- through a conception that is informed by what B actually has done and its consequences, and that improves upon
- B's own conception of her intention.

Moreover, Brandom is explicit that in this case, too, the norm that is identified in the rational reconstruction needs to be shared by the person who offers the rational reconstruction. "The aim" of the rational reconstruction, Brandom writes,

> is to make the deed as rationally reconstructed one that those recollecting it can endorse now, on their own behalf. However unpromising it might have seemed at the outset, the process the agent initiated by acknowledging a practical commitment (*Vorsatz*) is to be seen as turning out to have been a good one, one there is reason [to] have promoted then and to endorse now. (*ST*, p. 737)

And elsewhere:

> So the forgiving agent must endorse the norm being attributed as governing the deed—must acknowledge its authority. […] This is *identifying with* the agent, in the sense of risking and if need be sacrificing one's own attitudes, by subjecting them to normative assessment according to the norm one both attributes and acknowledges, and being subjunctively sensitive to that norm in one's own attitudes. (*ST*, p. 745)

The intention I that is identified in the rational reconstruction is indexed to a particular agent and a particular situation—a situation that, from the vantage-point

of the rational reconstruction, belongs to the past. But the above quotes clearly imply that the person A who offers the rational reconstruction for an action of B adopts an attitude towards I that potentially has consequences also for her own future actions. While Brandom is not explicit on this, we may distinguish the following as four natural options for understanding the content of A's resulting attitude (with "Φ" standing for the type of action specified by the relevant intention I):

> (1) Actions of type Φ are *permissible* for all agents who are in relevantly similar circumstances as B, *and have the same relevant preferences as B*.
>
> (2) Actions of type Φ are *permissible* for all agents who are in relevantly similar circumstances as B, *independently of their preferences*.
>
> (3) Actions of type Φ are *obligatory* for all agents who are in relevantly similar circumstances as B, *and have the same relevant preferences as B*.
>
> (4) Actions of type Φ are *obligatory* for all agents who are in relevantly similar circumstances as B, *independently of their preferences*.

Plausibly, rational reconstructions of actions within PEL may involve each of these attitudes.

III. Moral implications of postmodern ethical life

At one point of his discussion of PEL, Brandom describes magnanimity (and hence, PEL) as a "*moral* necessity" (*ST*, p. 578). In that context, he ascribes to Hegel a Kant-inspired strategy for deriving moral norms from the possibility of self-consciousness:

> Self-recognition, recognizing oneself, treating oneself as a discursive being, as able to undertake determinately contentful commitments, exercise determinately contentful authority and so on, requires recognizing others: attributing that kind of responsibility and authority to them. Any practical or theoretical presupposition of *that* is a structural presupposition of one's own self-consciousness. That is the source of moral requirements on how we treat others. (*ST*, p. 578)

Brandom's point here is that there is a generic moral obligation for us to adopt magnanimity and to establish a practice of PEL. But given the structure of rational reconstructions for actions that we analyzed in the last section, it is possible to show that the form of recognition required by PEL also has more concrete normative implications both for how we may act and for how we ought to interpret each other's actions—and hence, that the moral obligatoriness of PEL entails more specific moral obligations.

We saw in the last section that PEL requires us to interpret other persons' actions in terms of intentions that we can ourselves endorse. This endorsement amounted to acknowledging a norm that deems actions corresponding to that intention permissible or even obligatory (in terms of one of the four attitudes we distinguished at the end of the last section). But as I shall presently attempt to show, there is a principled limit on the types of actions for which a magnanimous agent A can endorse intentions, and which can be deemed permissible or obligatory by the norms A acknowledges. The types of actions that are excluded by this limit are not available as specifications of the intentions that A can ascribe to others in rational reconstructions of their actions. At the same time, given that A cannot acknowledge these types of actions as permissible, they are not available as permissible contents for her own intentions either. So, the limit in question imposes constraints both on the ways in which magnanimous agents ought to interpret the actions of other agents, and on the ways in which they are allowed to act themselves.[2]

Take courses of actions that involve killing someone. Such actions destroy, if successful, the victim's ability to develop rational reconstructions, and hence they diminish the ability of the whole community of PEL to engage in such magnanimous agency. By rationally reconstructing someone else's action in terms of an intention to kill somebody, or by herself killing somebody, a magnanimous agent A would therefore act directly against the interests of PEL. In addition, the action type "killing somebody" has possible instances where A is herself the victim of the deed, and where A's own ability to develop rational reconstructions would be destroyed if the action were carried out. In endorsing an intention of this kind in the context of a

7

rational reconstruction, or in acting herself in this way (in suicide), A would therefore contradict her own status as magnanimous agent, as creator of rational reconstructions.

Or take actions that include the systematic deception of others. Crucial to PEL is the joint enterprise of developing more and more adequate conceptions of concepts, of coming closer and closer to the truth about these concepts. In this sense, PEL aims at truth. This shared aim is undercut by deceptive actions. A magnanimous agent therefore must not ascribe to others intentions to act in such a way in rational reconstructions of their actions, or act herself this way.

Finally, consider actions that involve compulsion, oppression or deprivation of liberty. Such actions create asymmetrical recognitive relations as analyzed by Hegel in the master/slave dialectic, and such relations are incompatible with the symmetrical recognition that PEL calls for. A magnanimous agent must not ascribe an intention to act, and must not act herself, in such ways either, for this would again mean counteracting PEL—supporting or carrying out forms of agency that destroy magnanimous recognitive relations and replace them with defective modes of recognition.

There is a general principle behind these examples that we can now formulate. In doing so, we can drop the qualification that such-and-such is not permissible for *magnanimous* agents. For the prohibitions in question are normative implications of PEL which is itself morally obligatory, on Brandom's account: hence, these prohibitions have moral force, too, and are thus valid for all beings who are subject to moral demands. Keeping this in mind, we get the following principle (with "Φ" standing for a type of action, and "φ" for an instance of this type):[3]

(Limit of Forgiveness) If actions of type Φ impair the (ability for) magnanimous agency of (potential or actual)[4] members of a practice of postmodern ethical life, it is not permissible to

- acknowledge a norm that deems permissible or even obligatory actions of type Φ; and hence neither to
- ϕ themselves, nor to
- rationally reconstruct someone else's action in terms of an intention to ϕ.

In the following, I will use the terms "*magnanimity-incompatible*" and "*magnanimity-compatible*" to distinguish those actions, intentions and interpretations which are ruled out by Limit of Forgiveness, from those which are not.

As the above examples show, Limit of Forgiveness puts substantive moral constraints on our actions. This principle can thus be used to complement Brandom's neo-Hegelian theory of normativity with a derivation of moral norms (a possibility that Brandom himself alludes to in the passage that I quoted at the beginning of this section). I will not attempt to discuss here what exact moral norms can be established in that way. Instead, I will simply build on the above examples in my following discussion.

At the same time, Limit of Forgiveness substantively restricts the hermeneutic options that are available to magnanimous interprets. Take, for example, a situation in which A has a significant amount of evidence suggesting that agent B has killed another person C in order to steal C's money. The intention "kill C and steal his money" is magnanimity-incompatible, and hence is not an intention that A can endorse. Given Brandom's account of rational reconstructions of actions, this means that A cannot offer a rational reconstruction of B's action in terms of the intention "kill C and steal his money". Instead, if there is another possible interpretation that is magnanimity-compatible, this interpretation must be preferred.

Hence, interprets in PEL must try as far as possible to make sense of other persons' actions in terms of magnanimity-compatible intentions. Such interprets constantly work on the presumption that her fellow participants in the practice of magnanimity choose magnanimity-compatible actions—in other words, on the presumption that

they respect everyone's role as a participant in that practice (as someone who is able to and supposed to develop rational reconstructions for everyone else).

This presumption is naturally characterized as a form of *trust*. Brandom himself uses the term "trust" both in order to refer to the entire, complex recognitive structure which PEL gives rise to (hence the title *A Spirit of Trust*) (e.g. *ST*, p. 576), and to pick out one particular element of this structure, namely, our petitioning for future judges to rationally reconstruct our present concept applications (*ST*, p. 621). But as turns out now, PEL in Brandom's sense also requires trust in a more ordinary sense, namely that of an attitude of assuming that others will respect other persons' roles as (actual or potential) participants in PEL (and hence not act in ways that impair their ability to offer rational reconstructions). I will in the following call this latter attitude "magnanimous trust."[5]

IV. Moral problems with magnanimous trust

In this section, I am going to argue that magnanimous trust has deeply troublesome consequences. These consequences provide moral reasons that speak against any attempt to establish PEL in real practice.

Consider, first, the following question: can truly magnanimous interprets recognize magnanimity-incompatible actions as such? Imagine a magnanimous interpret A who witnesses a scene that looks as if B killed C and ran away with his money. Initially, the following picture might be tempting: A sees that B kills C and steals his money, and there is no reason to think that B does not do this intentionally. So, A has to ascribe to B the intention "kill C and steal his money". Since this intention is magnanimity-incompatible, A cannot share it, and hence cannot develop a rational reconstruction for B's action.

This picture is mistaken, however. It restricts the role of magnanimous trust to a stage at which A has already decided what B intentionally did and what his intention in doing so was. Magnanimous interpretation adds then merely the agent's own

endorsement of that intention (and possibly helps to decide among rival accounts of the agent's intention in cases that are less clear than our example case). But this restriction is unwarranted. For descriptions under which observers take actions to be intentional, and corresponding ascriptions of intentions to the agent, are not something that is unshakably given. Rather, they are themselves the result of (conscious or unconscious) interpretive decisions under conditions of uncertainty— i.e. decisions that are based on evidence which does not deductively entail a particular account of what has happened. An interpret who lives up to the demands of PEL must bring to bear her magnanimous trust also to this earlier stage of interpretation.

To spell out in more detail what this means, it will be useful to draw on Bayesian theories of hypothesis testing, which are generally seen as providing normatively valid models of data-interpretation under conditions of uncertainty (e.g. Jaynes 2003). Bayes's rule holds that

(Bayes's rule) $p(H|E) = p(E|H)*p(H)/p(E)$

where "p" stands for a (subjective) probability, "H" for a given hypothesis—in the case of interpretative tasks, a candidate interpretation—and "E" for a given piece of evidence. Let us quickly go through the various factors in Bayes's rule:

- $p(H|E)$, the probability of the hypothesis in question given the evidence, is known as "posterior probability." On Bayesian accounts of interpretation, agents compare the posterior probabilities of all relevant candidate interpretations (hypotheses) for the given evidence and select the hypothesis with highest posterior probability as their preferred interpretation.
- $p(E|H)$, the probability that the given evidence occurs if the hypothesis in question if true, is called "likelihood." In Bayesian accounts of interpretive tasks, this factor is used to model the fit between an interpretive hypothesis and the given evidence: if a given interpretive hypothesis has high likelihood, this means that it matches the evidence well. (For example, if the task is to interpret the fact that there is smoke on the horizon, the hypothesis that

there is a forest fire in that direction fits the evidence of there being smoke well: if that hypothesis were true, it would be very likely that there would be smoke in that direction.)

- p(H), the unconditional probability of the hypothesis, is called the "prior probability" of the hypothesis. This factor is used to model the internal plausibility of a given interpretation. (For example, a hypothesis on which the smoke on the horizon is caused by a UFO that has just landed, has much lower prior probability than the forest fire hypothesis.)
- p(E), the unconditional probability of the evidence, is needed as normalizing factor, to yield posterior probabilities between 0 and 1.

In these Bayesian terms, magnanimous trust—the presumption that others will always act in magnanimity-compatible ways—is equivalent to a particular assessment of prior probabilities for magnanimity-incompatible interpretations, namely, an assessment on which such interpretations have a prior probability of 0. For setting the prior probability for magnanimity-incompatible interpretations to 0 means to set them to the lowest possible value, and we need to do this in order to live up to PEL with its demand that we do what we can in order to provide rational reconstructions.[6]

Given Bayes's rule, such priors have the result that the posterior probability of magnanimity-incompatible interpretive hypotheses will automatically be 0, too. So, it is never rational for magnanimous agents to select magnanimity-incompatible interpretations as their winning hypothesis.

In situations where an action would have clearly seemed to be governed by a magnanimity-incompatible intention—and hence would have seemed to be "unforgivable"— this probabilistic effect of magnanimous trust can have either of two consequences. First, there may be another, magnanimity-compatible interpretation available that otherwise would have lost to the magnanimity-incompatible interpretation because of lower likelihood, prior probability or both, but that now—due to the effect of magnanimous trust—has the highest posterior probability and is adopted as winning hypothesis. Take again the above example,

where the magnanimous interpret A has witnessed a scene that looked as if B killed C and ran away with his money. Due to A's magnanimous trust, the interpretation that would have otherwise won—B as acting on the intention of killing C and stealing his money—will have posterior probability 0. By contrast, an interpretation on which the scene that A witnessed was, say, a skillfully staged artistic performance would otherwise have lost because of its relatively low prior probability, but it now may be adopted as best interpretation of B's action.

Second, there can be cases in which no winning hypothesis emerges. It is always possible to think of magnanimity-compatible "revisionist" interpretations,[7] but there may be a threshold of posterior probability that hypotheses need to reach in order to be selected as winning interpretation. In a given case, all available revisionist accounts may fail to reach this threshold, such that no interpretation is adopted at all. The result in this case is that the interpret suspends judgment about how to interpret the evidence.

As a consequence, wherever magnanimous interprets are confronted with actions that really are magnanimity-incompatible, magnanimous trust will prevent them from understanding these actions, and produce instead more or less bizarre false interpretations, and/or persisting ignorance.

This implication is troublesome for several reasons. First, false beliefs and ignorance about other persons' actions are damaging for magnanimous agents. Not only do they conflict with the epistemic aims of such agents, they also make it impossible for them to choose appropriate actions in order to protect themselves and the whole practice of PEL against the magnanimity-incompatible actions of others—by trying to intervene in on-going magnanimity-incompatible actions, or to prevent them in the future.

Second, if nobody is able to detect a magnanimity-incompatible action, it also follows that nobody can be made responsible and criticized for such an action. Without room for moral critique, moral improvement is hard to achieve. Moreover,

authors of magnanimity-incompatible actions like murder or enslavement would never have to face any consequences such as reproach or punishment for their actions, which would be unjust. Finally, one can argue that moral norms whose infringements are impossible to detect—at least for agents who comply with the demands of PEL that are the very source of these norms—are "private" norms, like those described by Wittgenstein in the private language argument. One should therefore doubt whether they can be considered norms at all, even more so in the context of a Hegelian view of normativity that puts so much weight on the social character of normativity.

So, there are several deeply worrisome consequences of magnanimous trust, and hence several reasons why we ought not to adopt such trust. Moreover, these reasons are moral reasons. Magnanimous trust makes us unable to protect PEL against magnanimity-incompatible actions that endanger this practice. So, it fulfils itself the criteria for magnanimity-incompatibility, and resistance to such trust is morally bidden as a consequence of the principle Limit of Forgiveness. And since magnanimous trust leaves no room for moral responsibility and critique, the moral norms that follow from Limit of Forgiveness require themselves that we do not adopt magnanimous trust—for absence of magnanimous trust is a condition of possibility of practically implementing these norms.

PEL thus turns out to create a normative contradiction: on the one hand, it implies that magnanimous trust is morally bidden; on the other hand, it implies that magnanimous trust is morally wrong. Notice, moreover, that this is not a case of a moral dilemma, where various moral demands happen to conflict with each other in particular cases. Rather, the contradiction in question is a systematic one, which follows from the very nature of PEL.

To all this, the following two objections can be made. First, the assaults against which, according to the foregoing argument, it must be possible to protect PEL consist in magnanimity-incompatible behavior. Should we not expect that such

behavior is overcome in the transition to PEL, so that there is nothing left against which the practice and its participants have to be protected?

It may be true that the transition to PEL requires all magnanimity-incompatible behavior to stop (although Brandom nowhere gets explicit about these aspects of the transition to PEL). But even if this is so, one cannot simply assume that PEL, once it is established, will forever be immune to such phenomena. In order to be able to persist in its existence, a practice of PEL needs to be able to protect itself against any regression into them, and my point is that magnanimous trust undermines this ability.

Second, it might be argued that PEL need not necessarily take a form where all agents set the prior probability for magnanimity-incompatible actions to 0. Brandom himself explicitly allows for magnanimity-incompatible actions that are recognized as such, and therefore are not amenable to rational reconstructions:

> It seems that the metanormative criteria of adequacy for *successful* forgiveness, both reparative and recollective, are in many cases *impossible* to satisfy. Some things people have done strike us, even upon due reflection, as simply *unforgivable*. In these cases, though we might try to mitigate the consequences of evil doings, we have no idea at all how to go about discerning the emergence of a governing norm we could ourselves endorse. (*ST*, pp. 746f.)

Does this not mean that PEL need not have the dramatic consequences I described above?

To this, I respond that there are two senses in which an action can be "unforgivable": a normative sense, according to which a rational agent must not forgive (rationally reconstruct) the doing; and a psychological sense, according to which a given interpret lacks sufficient psychological resources, and does not manage to forge a rational reconstruction (e.g. because she lacks sufficient imagination, or because she

15

does not manage to actually treat the prior probability of magnanimity-incompatible actions always as 0).[8] I do not deny that actions may be unforgivable within PEL in the second sense, and in such cases, magnanimity-incompatible actions can actually be detected as such by magnanimous interprets. But this also means that these interprets fail to live up to the demands of PEL. By contrast, an interpret who does comply with these demands will never recognize a magnanimity-incompatible action as such, and so there are no unforgivable actions in the first sense.

This seems to be Brandom's view, too. For after the passage I quoted above, he goes on to explain (*ST*, pp. 747f.) that in cases in which magnanimous interprets find themselves unable to rationally reconstruct a given action, they have to admit that this is partly their own fault, and that they fail to live up to the ideal of forgiveness they are committed to.

But in any case, the fact that there can be practices of PEL in which magnanimity-incompatible actions are recognized as such because interprets fail to fully realize magnanimous trust, does not assuage the above worries: the inconsistency that we had identified was located at the level of the normative implications of PEL, and considerations about psychological failures to live up to these demands cannot resolve that inconsistency.

V. Political problems with magnanimous trust

Yet could there not be a normative compromise that does in some sense resolve the inconsistency? Given the problems we pointed out in the last section, one might propose that all that PEL can reasonably demand from us is to realize magnanimous trust to *some* (imperfect) degree, such that there remain possibilities for moral critique, protection etc. In such a compromise solution, magnanimous trust might have only a very weak impact on prior probabilities, or even merely consist in a policy of preferring magnanimity-compatible over magnanimity-incompatible interpretations when both have equal posterior probability. This would enable magnanimous interprets to detect and criticize at least *most* of the magnanimous-

incompatible actions that are visible to a non-magnanimous interpret, and this would probably be enough to protect PEL against existential threats.

Insofar as this solution would require a systematic and far-reaching failure of compliance with what PEL really demands, it is questionable whether such a compromise would be in the spirit of PEL. In addition, even if it is granted that this compromise would resolve the problems that we rehearsed above, it still would have very problematic consequences. In particular, there are forms of political agency that create a problem for this proposal. I want to suggest that there is an important category of harmful political actions that share the following features:

(a) they pursue manipulative, oppressive, or other magnanimity-incompatible intentions;

(b) their agents have particular power;

(c) they are covert, in that they involve strategies that make them difficult to detect and/or criticize; and

(d) they are particularly efficient and harmful because of (b) and (c).

A first example is *propaganda*, which deliberately manipulates people's beliefs and emotions through lies, distortions, and misleading rhetoric. Propaganda is magnanimity-incompatible as it runs counter to the truth-aimedness of PEL. At the same time, propaganda is covert since it presents itself as truthful discourse, and its authors often explain the fact that their claims contradict those of other voices (e.g. of serious journalistic media) by providing conspiracy theories aimed at denigrating those voices. Finally, propaganda is typically efficient in manipulating people *because* its manipulative agenda is hidden behind the "official" facade of truthfulness; and it is efficient in manipulating at a large scale when its authors have power positions that enable them to reach a large audience, impose their agenda on state media, etc.

Further examples include:

- colonialism with its legitimizing narrative of a "civilizing mission" (Rist 2008, pp. 47-58);

- warfare for pretended humanitarian reasons;

- political communication that is deliberately ambiguous between a highly harmful (e.g. racist, sexist, fascist ...) and an innocuous message.

(In this last case—a favorite ploy in recent right-wing populist and extremist politics[9]—the offensive message does its harm, while the agent can point to the harmless message in order to defend his behavior against critique, and to denigrate his critics as persons who are exaggerating or making up things.)

Why do phenomena like these present a problem for the restricted version of PEL that we are considering? Given the covert nature of these phenomena, "friendly" interpretations (such as interpretations of propaganda as truthful, or of ambiguous political communication as harmless) will normally have, at a first assessment by average interprets, at least the same posterior probability as critical interpretations (which uncover the true character of the actions in question). But this means that a *magnanimous* interpret will believe the friendly version, even when magnanimous trust is restricted in the proposed way. This, in its turn, allows the action to fully unfold its deleterious effects, and makes it very hard if not impossible for the interpret to defend herself and/or others against these effects. Given the particularly efficient nature of the actions in question, these effects can cause harm to others (and to their magnanimous agency) at a large scale, and therefore pose a substantive threat to PEL. Any resistance against such forms of political agency requires in the first place insight into their existence and working. But it takes a more critical attitude than magnanimous trust (even in its restricted form) to gain such insight: an attitude that takes any sign for manipulation, oppression and the like very seriously, and treats it at least as sufficient reason for further investigation.[10]

I therefore conclude that the proposed compromise with its restricted version of magnanimous trust is not a satisfactory solution either. Together with the argument in the last section, this shows that PEL is a practice with contradictory normative implications; and it is not plausible that there is a moral requirement or necessary implicit commitment to create a practice of this kind.

References

Baier, Annette (1986) "Trust and Antitrust", *Ethics* 96/2, pp. 231–260.

Brandom, Robert (2019) *A Spirit of Trust: A Reading of Hegel's Phenomenology* (Cambridge, MA: Harvard University Press).

Jaynes, Edwin Thomas (2003) *Probability Theory: The Logic of Science*, edited by G. L. Bretthorst (Cambridge: Cambridge University Press).

Jones, Karen (1996) "Trust as an Affective Attitude", *Ethics* 107/1, pp. 4–25.

Knappik, Franz (2013) *Im Reich der Freiheit: Hegels Theorie autonomer Vernunft* (Berlin/Boston: de Gruyter).

Korsgaard, Christin (1985) "Kant's Formula of Universal Law", *Pacific Philosophical Quarterly* 66, pp. 24–47.

Rist, Gilbert (2008) *The History of Development: From Western Origins to Global Faith*, translated by P. Camiller, third edition (London/New York: Zed Books).

Stefanoni, Franco (2017) "Salvini con gestore della 'spiaggia fascista': 'Non processiamo le idee'", *Corriere della Sera*, July 18 (available online).

Warren, Mark (2018) "Trust and Democracy", in E. Uslane (ed.) *The Oxford Handbook of Social and Political Trust* (Oxford: Oxford University Press), pp. 75–94.

---

[1] I discuss further problems with Brandom's reading of the *Spirit* chapter in Knappik 2013, esp. pp. 125–134 and pp. 452–461.

[2] The following discussion is indebted to Korsgaard's discussion of Kant's Categorical Imperative and her "practical contradiction" interpretation of the Formula of Universal Law in Korsgaard 1985. Drawing on Kantian moral theory in this Hegelian context is justified by the fact that Brandom himself characterizes the strategy by which Hegel (on his reading) establishes the moral obligatoriness of PEL, as Kantian (*ST*, pp. 559f.).

[3] While my account of the various attitudes through which an agent can endorse an intention in section II was indexed to circumstances and—in cases (1) and (3)—preferences of the agent, I drop these qualifications here: as the above examples

suggest, types of actions that impair magnanimous agency normally do so independently of particular circumstances and preferences.

[4] An action like murder destroys the victim's ability for magnanimous action even when the victim does not live in a practice of PEL.

[5] One problem with Brandom's account that I shall bracket in the following is that there is an influential view (see, for instance, Baier 1986) on which trust cannot be adopted at will for practical (as opposed to epistemic) reasons, where the "moral necessity" of PEL would be such a practical reason.

[6] Assuming that we can adopt a particular probabilistic assessment for practical reasons. If this is not possible, it is implausible that there is a moral requirement to adopt a particular interpretive stance in the first place (see also footnote 5).

[7] Even when she has collected a lot of relevant evidence—e.g., in the above example, verified that C is dead—it is always possible that the interpret is only dreaming, or that some similar skeptical hypothesis is true.

[8] Brandom makes a similar distinction at *ST*, p. 747.

[9] To give just one example: in July 2017, the present Italian Minister of the Interior, Matteo Salvini, made a public visit to a beach resort whose tenant had put on display signposts with quotes by Mussolini and other references to fascism and Nazism, and was under legal investigation for this reason. Salvini had pictures taken of himself embracing the tenant and pointing smilingly to the signposts and declared on this occasion: "This is not a political visit, but one of support for a professional activity that employs dozens and dozens of persons." (Stefanoni 2017, my translation).

[10] This attitude can still be one of trust (and in many cases, it should be: see Warren 2018 on trust in institutions as precondition for democracy). But it needs to be a *critical* form of trust where the trustor is open to evidence suggesting that the trust is, or has become, unwarranted, and ready to eventually revoke his trust—as opposed to a form of trust where the trustor is overly biased in favor of the trustee, and hence *blind* to signs that the trustee may not be trustworthy (see Jones 1996, pp. 11–13, who gives the example of trust in a "particularly charming and particularly irresponsible" friend).