Reason, Recognition, and Internal Critique

Antti Kauppinen
Published online: 06 Nov 2010.

PLEASE SCROLL DOWN FOR ARTICLE

with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Symposium on Axel Honneth and Recognition

Reason, Recognition, and Internal Critique

Antti Kauppinen
University of Helsinki

Normative political philosophy always refers to a standard against which a society's institutions are judged. In the first, analytical part of the article, the different possible forms of normative criticism are examined according to whether the standards it appeals to are external or internal to the society in question. In the tradition of Socrates and Hegel, it is argued that reconstructing the kind of norms that are implicit in practices enables a critique that does not force the critic's particular views on the addressee and can also be motivationally effective. In the second part of the article, Axel Honneth's theory of recognition is examined as a form of such *reconstructive internal critique*. It is argued that while the implicit norms of recognition made explicit in Honneth's philosophical anthropology help explain progressive social struggles as moral ones, his theory faces two challenges in justifying internal critique. The *Priority Challenge* asks for the reasons why the implicit norms of recognition should be taken as the standard against which other implicit and explicit norms are to be judged. The *Application Challenge* asks why a social group should, by its own lights, extend equal recognition to all its members and even non-members. The kind of functional, prudential, conceptual, and moral considerations that could serve to answer these challenges are sketched.

Normative social criticism by definition appeals to some norm or standard that purports to give a reason to change a social arrangement or policy. It is not just empty talk or an attempt to coerce or manipulate the addressee of the criticism. It aims to change the way things are done by convincing people rationally, by giving them good reasons. It appeals, in part, to facts, such as the fact, if it is such, that cuts in welfare programs have led to increasing numbers of homeless people. If the purported reason is not a fact, it obviously cannot be a good reason. But not all facts give reasons for action, and arguably none does by itself. There has to be something more, something that makes the fact normative for future action, attitude, or emotion. In the above example, it could be something like the following norm: we owe all members of our society the chance to meet the basic physical necessities of living, and that takes priority over tax cuts. This kind of norm makes the mentioned fact a normative reason to change the policy; in other words, it *justifies* a change in the policy. Accepting the norm involved – and its applicability to the current case – makes the fact a *motivational* reason as well (given a suitable

motivational background). After all, as rational agents, we generally act according to our conception of what there is most reason to do. Given that, the facts – or, more conservatively, the attitudes towards them – also serve to *explain* the actions (such as changing the policy). There is therefore a connection between justification and explanation, between what ought to be and what is.

The crucial question for such normative criticism is obviously this: what makes the norms it appeals to genuinely normative? In traditional societies this question was not problematized: the norms embedded in the tradition and expressed in religion and laws had a diffuse *de facto* validity that left open only questions of application to particular cases (cf. Habermas [1996]). But with the advent of modernity this trust was shattered, and grounding moral norms became an acutely felt need. I will not attempt to go through all the answers given to this question from Hobbes on. What I want to do is present a typology of possible sources of normative standards and then apply this to examine the promises and faultlines of Axel Honneth's theory of recognition. For the purpose of exposition, I will present the alternatives in terms of criticizing other societies. This, I believe, brings into focus the problem of relativism. The different forms of critique nevertheless all apply to our own society as well, *mutatis mutandis*.

I. External Critique and Scepticism

So, where do we find the standards for criticizing a society (or policy or arrangement or practice; for brevity, I will speak of a society, although it can be somewhat misleading)? The first answer is simply to use the ones we have ourselves. Here it is irrelevant whether or not the criticized society shares these standards. Therefore this type of criticism deserves to be called *external critique*. It can take two forms depending on the nature the invoked standards are taken to have. The first, weakest kind of external critique is simply based on our existing norms as such, without any pretension that they have some special authority. I term this *ethnocentric external critique*. We have come to believe that it is wrong to force women to wear a veil (whether they are motivated to do so by what we take to be religious indoctrination or by threats), and *therefore* Saudi Arabians should change their institutions and practices. Though this may sound like a caricature, I believe it is close to Richard Rorty's notion of 'frank ethnocentrism' (see, e.g., the essays in Rorty [1991]). This appeal to our contingent norms (or our norms *as* contingent) can give the criticized society neither a normative reason nor a motivational reason to change; a 'so what?' reply is always in order. Our norms are not normative for *us* just because we happen to accept them, so they are hardly likely to be normative for *them* for that reason.

Given that ethnocentric external critique is so weak, the temptation is to seek to provide our norms, or at any rate our ideals, with a special authority deriving from nature, reason, or some such. I call this kind of critique *universalist external critique*. It appeals to standards and norms that are taken to be valid across time and place, regardless of what people actually think. Much of moral philosophy can be seen as an attempt to articulate and justify such universal moral standards, or at least construct universal procedures to arrive at local standards. This is by no means a worthless pursuit. It has given us valuable insights into morality, value, and human nature. Such critique faces a heavy double burden none the less. First, every attempt to find a foundation for moral norms has so far failed to command universal recognition. It seems rather that there are obstacles to finding such normative foundations: to paraphrase G. E. Moore, it seems we can always ask 'But is it moral?' after it has been demonstrated that a course of action is in accordance with the norms of rationality, human nature, or evolution. Second, in a world where value pluralism prevails, universal norms must by their nature be highly abstract and consequently distanced from ordinary motivation. They are therefore likely to fail to give a reason in the motivational sense in concrete situations. Neither of these considerations, of course, is conclusive, but together they suggest that finding alternatives to universalist criticism is at least pragmatically desirable.

The problems of universalist critique have led some to give up on the idea of normative criticism altogether. We find this kind of *normative scepticism* in some explicit pronouncements of Michel Foucault, for example. For Nietzsche, at least as he is normally understood, appeals to moral norms are merely a mask for will to power – what else could they be if they are founded on illusions of a non-existent moral reality? At the very least, finding universally valid norms presupposes reaching a point of view above the fray of contesting claims, and the skeptics argue, plausibly enough, that more often than not this supposedly impartial viewpoint is seriously partial – as male, white, and Western as God himself.

To give up norms, however, need not mean giving up on criticism altogether. In the middle period of Foucault, social criticism becomes a meticulous empirical analysis of particular power relationships, their effects and maintenance, with the aim of bringing to light their contingency and historicality and therefore the possibility of transforming them at local levels. However, as many critics, including Habermas (1985, pp. 330–6), have argued, this pretended 'happy positivism' is hardly free of normative content and intent, which only remains unarticulated and unarticulable within the framework of a theory of power. Normative scepticism is no less an unstable position than normative universalism. As Foucault himself reportedly said, we can no more stop making judgments than stop breathing, and judgments refer to or set up normative standards.

II. The Idea of Internal Critique

But what other alternative is there if we cannot draw on external standards of any kind? The roots of the answer lie at the very beginning of Western philosophy, in the Socratic method and in Aristotle's subsequent notion of dialectical argument. They contain the seeds of internal (or 'immanent') critique, the kind of normative criticism that appeals to the commitments of the addressee of the criticism rather than those of the critic.

Socrates, as we know, always denied he was trying to teach people something they did not already know. Rather, he saw his job as that of a handmaiden: his questioning helped people already pregnant with knowledge to articulate it in explicit form, thus bringing their views and actions under rational control. He frequently showed that what people took for granted turned out to lead to contradictions if probed far enough, thus forcing people to rethink and revise their beliefs without (ideally) presenting any of his own.[1] This idea was formulated in the deep Platonic doctrine of *anamnesis*, literally 'un-forgetting': learning is a matter of retrieving, bringing into light, through dialogical reflection, the buried knowledge we always already have. Aristotle, too, made use of the idea in distinguishing two types of argument, demonstrative and dialectical ones. While demonstrative arguments begin from premises that are self-evident, dialectical arguments are essentially intersubjective, beginning from premises accepted by one's partner to dialogue. Their purpose is, as Robin Smith puts it, to 'make us able to deduce the conclusion we want from premises conceded by the opponent we are faced with' (R. Smith [1995, pp. 60–61]).

In modern philosophy the tradition beginning from Descartes took the argument-type Aristotle labeled demonstrative as fundamental and looked for self-evident or certain foundations from which to derive and deduce philosophical theorems. It was Hegel, of course, who brought back the idea of dialectical argumentation when he argued that history moves ahead through determinate negation of immanent contradictions. His work inspired a minority tradition of immanent or internal critique, which proceeds basically by noting the contradictions inherent in our own forms of life, anticipating the direction of their resolution, and working theoretically and practically to bring that resolution about. The most notable representatives of this tradition are obviously Marx and his heirs, including the Frankfurt school. For example, Horkheimer argued that critical theory is a continuation of pretheoretical critical activity (Horkheimer [1937]; Honneth [1991, p. 15]). Incentive (and justification) for such critical activity comes from subjective experiences of injustice, which are a part of the social reality under investigation. This refusal to separate the 'ought' from the 'is', or, more precisely, to find the 'ought' in the 'is', the reason in the real, is paradigmatic of the Left Hegelian tradition.

Internal critique has two important advantages over external critique. Recall the difficulty of making sure that external standards are not somehow partial, reflecting the history and tradition of a particular society or group. When the addressee of criticism is judged against its own standards, this problem is bypassed. The second advantage is the increased likelihood of practical effectiveness. If the norms are genuinely internal, there will exist some degree of motivation to follow them and to sanction compliance, which can lead to the addressee changing its behaviour without external sanctions once it has become aware that its practices contradict its own standards. For these two reasons, internal critique is at least pragmatically preferable to external.

Simple Internal Critique

To be useful today, the Hegelian concepts of contradiction, negation, and resolution must be formulated and made precise in terms that are free from the taint of suspicious metaphysical notions like self-unfolding of the Concept or *Geist*. I want to make a beginning towards that direction by distinguishing between two forms of internal critique. The first form, which I will call *simple internal critique*, appeals to a contradiction between the explicitly or publicly avowed norms of a society or institution or practice and its actual practice. (More precisely, it could be called 'internal critique based on explicit standards', but that is too clumsy.) This is a rather obvious form of social criticism, found in German *Ideologiekritik* and in its Anglo-Saxon varieties. Honneth characterizes it in *The Critique of Power* as follows:

The classical critique of ideology found in Marxism ... supposed that, in the ideological self-understanding of bourgeois society, its culturally proclaimed goals and legitimations, normative principles are already presented that contain the standards by which the actual social regulation even of this society can be morally criticized. (Honneth [1991, p. 229])

Simple internal critique can further be subdivided into different forms based on whether or not those responsible for running the society or practice are aware of the fact that the actual practice fails to meet its explicitly stated norms or ideals. If they are, they can fairly be termed cynical – they are exploiting the legitimacy brought by the ideals to knowingly advance contrary goals. Criticism of such institutions might be called *exposing* internal *critique*. If they are not aware, there is a failure of self-knowledge, perhaps through self-deception. The purpose of critique is to educate the addressee about the contradiction between the ideal and the real; therefore I will label it *enlightening* internal *critique*. As an example, take the case of the World Bank and IMF, whose explicit purpose is to help bring wealth and

well-being to developing nations, while their actual practices all too often lead to the opposite results, at least for the impoverished majorities. If the people running these institutions are aware of this but continue none the less, they deserve to be subjected to exposing critique. If they continue to believe that their actions further their stated aims when they do not, it could be that they are deceiving themselves, for example to avoid the psychological costs of revising deeply rooted false beliefs about economics. In this case, the most suitable form of critique is enlightening: in so far as we engage in it, we hold out the hope that *if* the addressees of criticism realized the true consequences of their actions, they would change what they do.

Reconstructive Internal Critique

Simple internal critique can be an effective weapon, particularly in our liberal societies, whose avowed ideals of freedom, equality, and tolerance are indeed quite admirable. But not all societies and institutions have such ideals in the first place, or if they do their interpretations of what they amount to in practice may differ radically. We may assume that the actions of the Taliban regime or Nazi Germany fell in line with their stated aims, which means that we cannot criticize them for failing to meet their explicit ideals. Does this force a retreat to external critique or scepticism about the possibility of normative criticism? No, because not all norms or ideals are explicitly stated or even voluntary. This makes possible what I will call *reconstructive internal critique*, critique based on making implicit standards explicit. The idea is that there are norms and normative pressures in our social practices that are not (necessarily) explicitly stated. They are manifest, for example, in the form of unarticulated emotions and informal sanctions, or perhaps partially articulated in the form of proverbs, religious prohibitions, or laws. They can be theoretically reconstructed and thus brought into the game of giving and asking for reasons. For example, a Taliban recruit may have a vague feeling of guilt for beating a woman who is outside her home at the wrong time, even if he is thoroughly convinced that this action is prescribed by the religious norms he believes in. The norm in question may nevertheless cohere with all of his explicit beliefs and ideals. We can begin to give him a reason to stop and question his normative beliefs by making explicit the normative expectations that his actions violate, that is, the fundamentally gender non-specific reciprocal norms of respect for persons that are, to anticipate the discussion below, found in some (however rudimentary and implicit) form in all human societies by anthropological necessity. In this kind of case, there is *ex hypothesi* no contradiction between theory and practice, but a contradiction *within* a practice, a contradiction rooted in normative expectations that are constitutive for human social practices.

Reconstructive internal critique can take different forms depending on whether the implicit norms appealed to are particular or universal. *Weak*

Type of critique	Appeals to	Representatives	Problems
Normative skepticism	–	Nietzsche, Foucault	Cryptonormative
Ethnocentric external critique	Our norms	Rorty, later Rawls (?)	Relativist
Universalist external critique	Metaphysically grounded norms	Plato, Kant, Utilitarians	Suits God, not us
Simple internal critique	The contradiction between expressed norms and actual practice	*Ideologiekritik*	Cannot deal with fanatics; also, explicit norms are easy to change
Reconstructive internal critique	The contradiction between norms embedded in practices and actual practices	Habermas, Honneth	??

reconstructive critique relies on norms that simply happen to structure the practices of a particular society. Its *strong* form appeals to universal implicit norms. Here it is not a matter of what ideals a society *happens* to have, but what it *has* to have. Human social life is shot through with normativity, with 'existing reason' in the Hegelian sense (cf. Honneth [2001b, pp. 8–9]). Habermas's theory of communicative action is clearly of this type. It reconstructs the 'unavoidable idealizing presuppositions' of speech acts, such as the fact that by making claims at all we submit ourselves to an authority that transcends all factual authority, and uses the degree to which these idealizations are realized in practices and institutions as a yardstick of their rationality. He famously argues that modernity has to find its normativity within itself, and presents his model of communicative rationality as the only credible non-metaphysical alternative to radical scepticism (e.g. Habermas [1985]). If I am not wrong, Honneth's formal theory of ethical life aims to provide normative tools for reconstructive internal critique in the same strong sense as Habermas's. The difference is that, as we will see, Honneth aims to give a better account of the motivational and explanatory aspects of existing reason with his anthropological approach. The table summarizes the main options presented above.[2]

III. Honneth: The Implicit Norms of Recognition

Honneth's theory of struggles for recognition aims to provide a framework for

internal critique that draws from normative sources implicitly embedded in all human practices. He emphasizes that 'our critical standards are bound together with reality', and that 'the normative demands of critical social theory have their genuine social place in this dimension of recognition' (1999). Because these normative demands arise within the practices themselves rather than result from external evaluation, it is a conceptual and empirical possibility that they become motivationally efficacious for the participants themselves. Thus, they can serve to explain social struggles and changes as specifically *moral* struggles, rather than struggles for self-preservation or for satisfying material interests (e.g. Honneth [1995, p. 2; cf. p. 144]).

This dual aim of explaining and justifying social struggles is familiar from such theories as Habermas's theory of communicative action. In contrast to Habermas, however, Honneth seeks to provide an *anthropological* rather than linguistic (transcendental-pragmatic) foundation for critical theory. His theory takes its point of departure in the intersubjective processes of identity formation that form an important part of the cultural (as opposed to material) reproduction of human societies. At its center is the claim that human beings can develop certain types of identity-constituting self-relation only if they are in certain kinds of relations with others, and the nature of those relations to self depends on the nature of the relations to others.[3] In other words, the claim is that individuals can take certain attitudes towards themselves and their beliefs, desires, values, and the like, if and only if others manifest certain attitudes towards them. For some attitudes, there is an additional requirement that the individual herself, in turn, take certain (perhaps the same) attitudes towards those others; this is the demand of *mutuality* that Hegel famously analyzed in the dialectic of the Master and the Slave. In a word, *no identity without (mutual) recognition*.

'Recognition' is here a name for a complex of attitudes towards others, not for a cognitive state. These attitudes have in common that their *expression* forms a part of their content. Recognitive attitudes cannot be merely internal: if you 'recognize' someone in thought, you have not yet really recognized her.[4] Depending on the context and type of recognition, this expression can take many forms from gestures like a simple nod (Honneth [2001a, p. 116]) to declaring a cease-fire so that the Red Cross can collect the wounded. Because recognition is in principle public, it is possible for one to experience herself as recognized, to experience the other's attitude towards oneself. This experience, in turn, makes possible and influences one's own attitudes towards oneself.

Honneth argues that there are three basic forms of recognition: love (expressed centrally in parental care), respect (expressed by granting rights and holding responsible), and social esteem. In being loved, we are recognized as needy yet independent individuals. When others respect us

(taken here in a Kantian sense), we are recognized as responsible, autonomous agents, and when they esteem us for our particular traits or abilities, we are recognized as particular kind of agents. To these forms of recognition correspond three forms of practical self-relation: basic self-confidence, self-respect, and self-esteem. Honneth backs the correlation between recognition and self-relation by referring to empirical research that shows how children develop their sense of independent selfhood and of the worth of their desires in emotional interaction with their parents, how our conception of ourselves as responsible, autonomous agents depends on having rights to enter into legal relationships and public deliberation on equal footing with others, and how our valuing of our particular traits and abilities depends on social esteem given in communities of value (Honneth [1995, ch. 5]).[5] Both within these categories and between them there is a kind of teleological pressure towards more encompassing forms of recognition: for example, while a formal right to vote expresses partial recognition of each citizen's ability to enter into democratic decision-making, it can in practice soon be seen as being in need of supplementation by a general right to education, and so on.[6]

For reasons that I will return to in the final section, this anthropological necessity gives rise to an implicit norm of recognition that is present in some form in all human societies. Following Robert Brandom's suggestion, we can speak of implicit norms when a group (or, indeed, an individual) treats a type of performance as correct or incorrect in practice without applying an explicit rule to it (e.g. Brandom [1994, p. 32]). Norms of action, in general, are (roughly) generalized behavioral expectations with some kind of sanctions attached to their violation. For example, take the norm delimiting the range of appropriate topics for dinner-table conversation. Few of us could ever articulate such a complex and context-dependent rule. But we are generally able to discriminate between success and failure in complying with it. To use a well-worn distinction, it is a matter of knowing how rather than knowing that. We can tell when someone – a child, a foreigner, a drunk, an insensitive person – violates the norm, and, depending on the context, there is a reaction; a hidden irritation, a meaningful silence, a raising of the eyebrows, a rebuke, in brief, some form of external sanction. The person who has violated the norm may feel a pang of guilt or shame, in general, some form of internal sanction.[7] We generally become conscious of the norm's existence only when it is violated. This can lead to a rough, partial verbal formulation, such as 'Don't bring up the Pacific War with the Japanese', which can be used as an aid in teaching this particular normative skill. What matters for our purposes is that the norm was already present in the practice, fallibly governing the actions and reactions (including at least the emotions and motivations, as well as possibly judgments that do not explicitly refer to the norm) of the participants, before it was explicitly (and partially) formulated. After the

norm has been formulated explicitly, one can ask for grounds for it or use it as a premise in deliberation and practical inferences – in other words, bring it to the 'game of giving and asking for reasons' (cf. Brandom [1994, pp. 105–7]). Working out precise identity criteria and epistemology for implicit norms is a task yet to be fulfilled, but the basic idea is solid.

IV. Recognition and Moral Explanation

So much for an overview of Honneth's theory. From an explanatory point of view, it represents a significant advance over crudely materialistic interpretations of Marxism as well as interest-based rational choice or evolutionary models of moral and societal development. The point of departure is experiences of misrecognition, which could be divided into *non-recognition* (failure of recognition, passive disrespect – such as not noticing one or failing to acknowledge one's achievements) and *malrecognition* ('negative' recognition, active disrespect – such as treating one like an animal or making fun of one's aspirations). They are not felt merely as personal disappointments but as specifically moral injuries, violations of (possibly unarticulated) legitimate normative expectations. The ensuing negative moral emotions – indignation, shame, humiliation, and so on – play a double role in the theory: *epistemologically*, they serve as defeasible evidence of a failure of recognition and thus the presence of a normative reason for change, and *motivationally* they move one to seek it. If they are read as typical for a group of people, they can lead to a social struggle for recognition (Honneth [1995, p. 162]). It should be noted that Honneth wisely does not advance the reductive thesis that *all* social struggles are struggles for recognition (ibid., p. 165); that would amount to an implausible reduction of all politics to identity politics. Some situations can lead to old-fashioned redistributive politics.[8]

The double role of moral emotions is essential in the theory. They make the difference between interest-based explanation and genuinely *moral* explanation, because they serve as the psychological link between wrong-making features and action.[9] It seems, however, that Honneth has not adequately dealt with the implications of their less than perfect reliability. There are two kinds of failure: either moral emotions (and consequently motivation) are present, although there is in fact no failure of recognition, or there is a failure of recognition, but one lacks the emotions and motivation. Actual moral feelings are therefore neither necessary nor sufficient for there to be a normative reason for a social struggle. I will next examine the two cases where the link between real failure of recognition and moral emotions is broken.

Emotions Without Normative Reason

It can surely be the case that a person or even a group of persons has moral feelings that explain its actions without there being a justification for them. For example, a refugee may attribute her failure to be admitted to a university to misrecognition in the form of racism even when the decision has in fact been made in ignorance of her ethnic background, the questions have not been biased, and so on. The felt disrespect does not justify her anger with the university, because it is itself not justified; in fact, no disrespect was involved. Moreover, not all particular identities deserve to be recognized. It is no cause for normative criticism if the institutions of a society make it hard for a neo-Nazi to feel proud of himself. If there is disrespect, it is merited.

To get a grip on the problem, we can ask what makes the emotions unwarranted in these two cases. In the first, the refugee has false *factual beliefs* about the criteria the university has used. In the second, the hurt neo-Nazi has false *evaluative beliefs* about the value (and thus recognition-worthiness) of his particular identity. (This, naturally, is a judgment that must be made from a substantive ethical point of view; to be able to tell which identities deserve recognition, we need other, more substantial moral norms.) What these cases suggest, therefore, is that the existence of negative emotions is evidence of genuine misrecognition only if they are based on true factual and evaluative beliefs. This is because true beliefs are needed for the psychological link between wrong-making features and emotional motivation to be valid. It may seem that emotions drop out of the picture here as redundant, but they do have defeasible epistemological significance, because they serve to direct one's attention to *possible* failures of recognition. An emotionless being might well fail to catch an insulting remark or a racist slur. Moreover, from an agent's own point of view, it may make sense to trust one's emotions at the expense of one's (evaluative) beliefs; after a proper moral education, they form a central part of one's total moral sensibility and in favorable conditions serve as generally reliable indicators of value and disvalue (cf. Helm [2001]).

Misrecognition Without Motivation

Historical and contemporary experience tells us that failure of recognition does not necessarily lead to any kind of struggle. It is possible that people do not in fact have moral feelings and motivation when they would have a reason to have them. To take an extreme example, many refugee women in the West voluntarily undergo genital mutilation without experiencing it as a failure to respect their physical inviolability. This may result from mutual ignorance within the community or conscious manipulation. Ideology and indoctrination are not simply means to lead people to misidentify their interests, they

can also amount to a miseducation of the emotions: a woman who learns the 'woman's place' will not be offended when she is belittled or bossed around by her man. In part this is because emotions are responsive to judgments about value – if the woman comes to believe that she is of lesser value, she will feel that she is treated properly (although there may still be contradictory feelings) – but lifelong training can surely dull emotional responses even at the most basic, non-intellectual level.

Even when emotions exist, agents may be led to misidentify their causes or fail to identify them altogether, or fail to be motivated for any action. Perhaps they lack the conceptual resources either to articulate the misrecognition as such in the first place or to argue for their intuitive take on the situation. This can lead to a conflict between their explicit judgments, which attribute the frustration and bitterness to personal problems, for example, and emotional reaction, which is (rightly) directed at the unjust circumstances.[10] As Honneth acknowledges, the agents need a 'suitable interpretative language' (Honneth [1995, p. 162]) before their emotions can translate into (collective) action.

V. Recognition and Internal Critique

The normative core of Honneth's argument is that violating the implicit norm of (mutual) recognition gives a reason for any group G to change its practices. From the participant perspective, it justifies a struggle for recognition to change the practice in question. From the perspective of a social critic – who may or may not be a member of G herself – it justifies internal critique. This simple sketch leaves open several questions about justification. To begin with, G will have several other implicit and explicit norms. What is it that makes norms of recognition a proper yardstick for the G's actions and possibly its other norms? Why should they not rather give up or modify the norms of recognition? How does the empirical importance of recognition for the development of identity translate into a normative demand? These are questions that serious normative social criticism must answer lest it collapse into mere assertion. As a form of internal critique, it must show that there is a *reason* from the group's own perspective to hold on to the norm of recognition.

We can begin examining the status of norms of recognition by looking more closely into the kind of cases where they are violated. It is important to notice that it makes no sense to attribute an implicit norm to G in the first place if its members never act according to it or sanction each other for failures. In the case of norms of recognition, the typical situation is therefore that some persons are recognized while others are recognized either inadequately or not at all. For example, eighteenth-century Americans, both at the individual and institutional level, recognized white men as full persons

with human rights, while Blacks, and to a certain extent also white women, were treated as non-persons, fit to be bought and sold or excluded from political participation. These cases raise two analytically distinct problems for criticism. First, it may be that within *G*, other norms, either explicit or implicit, override or silence the norms of recognition. Perhaps there are norms within *G* prescribing, for example, distinct public roles for men and women. Here the internal critic will have to show that *G*'s commitment to recognition norms is somehow more basic or fundamental. I call this the *Priority Challenge*. The second problem concerns the extension of recognition. Even if norms of recognition are admitted to be basic to human interaction, there is a further question about why they should govern behavior with regard, for example, to outgroups. The challenge for the would-be internal critic is to show why Americans would have had a good reason in the light of *their own* commitments and interests to extend full recognition to the disadvantaged groups. I will call this the *Application Challenge*.

How can one reply to these challenges? It is perhaps useful to begin by considering what kind of reasons one can appeal to. As far as I know, there is no generally accepted list of possible sources of reasons, but we can distinguish at least instrumental, prudential, functional, conceptual, and moral ones. Looking at the two challenges, it must first be admitted that recognition is not always supported by *instrumental* or *prudential* considerations. It is true that we all need to be recognized, which presupposes that we recognize others in the cases that require mutuality, but we do not need to be recognized by everyone or in every situation to maintain our identity, self-respect, and so on. The pre-Civil War Dixie farmer may well have no instrumental reason to treat his slaves as full persons.

Functional considerations fare somewhat better. Given the (empirical) importance of recognition for individual development, any society will need norms sanctioning it. In addition, mutual recognition is a key precondition of *collective* agency (possibly even conceptually). Genuine we-attitudes such as we-intentions (see Tuomela [1995, ch. 3]) will not be possible unless group members recognize each other as persons capable of forming rational plans that take the plans of others into account. In collective action, a degree of mutual recognition is not optional. Functionally speaking, a group whose members recognize each other will be able to act collectively in ways that give it a competitive advantage. From the point of view of evolutionary psychology, this helps explain our innate predispositions to recognize other persons. Most other social norms will not have similar non-contingent functional significance, so these considerations might serve to answer the Priority Challenge from any society's own point of view. They do not, however, help with the Application Challenge, because it might well not be functional for a group to recognize outgroup individuals.

Broadly *conceptual* considerations provide the best non-moral answer for

the Application Challenge. It is plausible that recognition-worthiness supervenes on other properties (call them *S*-properties): we recognize agents as autonomous because they are capable of acting on reasons they reflectively accept or hold a football player in esteem in virtue of his ability to give precise long-distance passes. Now, when fully spelled out, the subvenient properties will not in most cases include such things as being Finnish or being a man – it would be very much *ad hoc* to suggest that only men should have the vote, because they are capable of responsible rational thought *and* because they are men. The only possible argument is that the non-recognized lack the *S*-properties, and indeed misrecognition is often rationalized with false factual beliefs. This presents a practical problem for the critic, who often has to struggle against prejudice and dogma, but from a normative point of view this makes no difference. The members of *G* have themselves committed to the norms prescribing inference from *S*-properties (of which, moreover, they may have false beliefs) to recognition, and as usual these commitments can outrun their actual performance – after all, it is precisely this feature of norms that makes them genuinely binding and makes possible making mistakes (Brandom [1994, pp. 50–55]). Whether or not they acknowledge it, they have a normative (logical-conceptual) reason to apply the norms of recognition to all who have *S*-properties, whether they are privileged members of *G* or not.

Groups may also have *moral* reasons to uphold and extend the norm of recognition. In this respect, Honneth appears to argue that any society should promote recognition, because recognition is a necessary condition for the kind of self-relation that makes possible full self-realization, and thus living a good life (Honneth [1995, pp. 172–3]). These conditions are *formal* in the sense that they are held to be necessary for *any* kind of good life regardless of content: a good life as a monk or as a thief, or a good life in Sweden or in the jungles of Indonesia, for example. Because these conditions are, on the one hand, formal and universal (although empirically grounded and thus contingent), and, on the other hand, oriented towards self-realization and good life, the theory combines elements from both Kantian and communitarian ethics (ibid.). Occasionally, Honneth substitutes other highly valued abstract goods like well-being, freedom, personal identity, or integrity for self-realization and good life. Each of these, it seems, is made possible by proper recognition. Misrecognition, in turn, causes problems in one's self-relation, and Honneth claims that moral injuries are precisely injuries to self-relation (Honneth [1997]).

It looks as if we have a straightforward hypothetical imperative here: if you value self-realization (freedom, personal identity, etc.), you should value relations of recognition, which make it possible. But if this is so, recognition as such plays no normative role in the theory: struggles for recognition turn out to be ultimately struggles for self-realization. A just society is one in which people are able to 'realize themselves', and recognition is only a means

for that. This suspicion has been raised before, for example by Michael O. Hardimon (1997) and Christopher Zurn (2000). Here is how Hardimon puts it:

What is doing the normative work in Honneth's critical theory? At the outset of the discussion, it appeared that *recognition* was the basic normative notion of the theory. But, then, when we get to the final chapter of the book, it seems that *self-realization* is the basic normative notion, recognition being important as only as a precondition of self-realization. (Hardimon [1997, p. 54])

He argues that it follows from this that Honneth's theory 'embodies a particular vision of the good life', which would seem to defeat its universalistic ambitions. It could only serve as a vehicle for weak internal critique. The challenge is, in other words, that the norms of recognition are *derivative* from more fundamental moral norms of self-realization, and these fundamental norms are relativistic, being specific to Western liberal societies that prize individual self-realization over other values.

How can Honneth solve this problem of moral justification? It seems he has two possibilities here, a foundationalist and a non-foundationalist one. First, he could bite the bullet and admit that recognition is secondary and self-realization is the real source of teleological pressure, and then argue that this does not in fact mean relying on a particular vision of the good life. Or he could give up the notion of looking for a foundation for normative criticism and instead try to find a coherentist justification for recognition. I will briefly sketch both of these alternatives; Honneth seems to favor a version of the first, while I myself believe the second is more promising. Both will turn out to blur the distinction between internal and external critique.

The first response goes something like this: perhaps recognition does not have a direct normative role. It is a good thing, because it, as a matter of fact, enables self-realization, freedom, well-being, and other such good things. However, according to this line of response, these values are not particular to historical traditions, but universal. They might even be necessary. Contemporary Kantians like Christine Korsgaard (1996) argue for this in the following manner: If you value anything at all, you must value freedom to pursue the things you value and which constitute a good life for you. And as human beings we must value things, because by virtue of the reflective structure of our consciousness we are not caused to act but must make choices according to our values, in the light of which we reflectively endorse certain desires. What the anthropological version of critical theory adds, on this line of response, is that as valuing beings we are always already at least implicitly committed to according value (that is, recognizing) to other valuing beings, because valuing implies valuing freedom and freedom requires mutual recognition.

There is, however, a dilemma for this approach. Arguably, the necessary

conditions for the pursuit of good life are either too vague and general to have any critical bite or they are so specific and concrete that they already embody a particular, Western (and liberal democratic) conception of the good life (one of the many conceptions available in the Western societies). Quasi-transcendental arguments like the Kantian one sketched above seem to lead to the first horn, while gripping the second one would require some kind of Minerva's Owl-style teleological argument showing that this particular form of the good life is somehow on a higher developmental level than others – in other words, that it is present as an implicit *telos* of other forms of life – so that its conception of the good can provide criteria for criticizing others. This may be what Honneth has in mind, but arguments are as of yet wanting. In any case, such solutions blur the distinction between internal and external critique, because they refer to some kind of a priori standards – albeit ones that are assumed to be internal.

The second, non-foundationalist response begins from the notion that perhaps there is no need to look for a single source of moral normativity, such as self-realization. As John McDowell puts it, moral reflection is 'Neurathian' (1996, p. 197): we begin from a situation where there always already are moral commitments and conceptions that can be weighed against each other. We need not begin from nowhere: the criticized group will have *some* moral norms and ideals, and they will necessarily overlap with ours, or we will not be able to recognize them as *moral* in the first place. Morally justifying the norms of recognition means then showing that they cohere positively with the rest of the group's moral values, intuitions, and norms – in short, showing that they feature in a *reflective equilibrium*. They need not, therefore, rely on the possibly questionable value of self-realization (although, given that *we* value self-realization, its relation to recognition is one reason for *us* to recognize each other).

Of course, this does not yet excise the specter of relativism, because there are many possible coherent normative systems. Some kind of external check is needed if we want to ensure priority and universal application for the norms of recognition. As long as we are in the business of internal critique this may seem impossible. But further reflection reveals that the norms and ideals of any group or society will have an intimate relation to their factual beliefs, which are constitutively governed by norms of responsiveness to independent reality (see Velleman [1996, pp. 707–89). This opens up an indirect strategy: given that the members of *G* are committed to believing truths and forming their moral conceptions in response to their beliefs, they are also committed to changing their moral conceptions if they turn out to be based on falsehoods. In other words, they are committed to reaching a *wide reflective equilibrium*, which balances moral intuitions and principles not only with each other but also scientific evidence (Rawls [1996, pp. 8, 384n16]). A possible reply to challenges, therefore, is that postulating an implicit norm of recognition

makes best sense of the group's intuitive judgments, emotions, and norms *after* they are, where necessary, corrected in the light of our best (social) scientific understanding, which includes the empirical theory of the importance of recognition. The hope is that moving from narrow to wide reflective equilibrium – which is itself internally justified by the norm of truth that is constitutive of any cognitive practices – will force the members of the criticized group to acknowledge that it follows from their own commitments that they ought to recognize each other and non-members. Since the indirect strategy may require the criticized group to change its normative conceptions, it, too, stands on the borderline between internal and external critique.

How Internal Critique Might Work

We are now in a position to see how the theory of recognition might (ideally) work as a tool for reconstructive internal critique. The critic has at her disposal well-supported empirical hypotheses about the significance of recognition for individual psychological development, collective action, and social reproduction. She also knows that because of this significance the criticized group will need to have at least implicit norms that prescribe a degree of recognition for its members. In addition, she has a rough epistemology for the way these implicit norms are indirectly manifest in moral emotions, informal sanctions, and judgments. With the help of this theoretical information, she can try to make the norms and ideals that inform them explicit in dialogue with the members of criticized groups and show where, from her point of view, their actual practice contradicts them. Of course, in the end it is always possible that a society's practices simply contain no standards or ideals that would allow one to appeal to what it itself takes to be reasons. Such cases can move the critic to examine the rationale behind her own conceptions. If she finds no fault with them, she has no choice but to use them as the basis of criticism, externally (cf. Blackburn [1998, p. 265]). However, given the functional, conceptual, and moral support norms of recognition enjoy, this kind of situation is less likely than it is sometimes thought.

If internal critique succeeds, it leads to a new normative self-understanding in the criticized group (cf. Anderson [1993, p. 104]) and thus, at least in so far as the members are practically rational, to changes in their motivation (see, e.g., M. Smith [1995]). Of course, we all know how easy it is for people to discard criticism, but there is a real chance that this kind of critique will motivate at least some members of the criticized society to demand more adequate practices of recognition. After all, in the absence of massive manipulation – which can itself be countered with disclosing critique – there already exists, in the form of feelings like guilt and shame, a motivational potential within the society that fails to recognize adequately. Internal

critique, in so far as it remains true to its Socratic ideals, merely makes explicit the normative expectations that inform these reactions, and so makes them available for further reflection and action.[11]

NOTES

1 In the actual dialogues of Plato, the Socrates character no doubt feeds and forces his own ideas on his interlocutors. But the Socratic ideal remains what it is.

2 There are several possible forms of social criticism that are missing from the table. One is what could be termed *visionary* or *utopian* critique, which would be openly based on invented ideals and standards not yet realized in any society. (It may be, though, that this collapses back to other forms, since it is hard to see how we would be able to judge a proposed ideal as desirable if not in relation to what we already at some level accept as such.) Another important form is *genealogical* critique, which Honneth characterizes as 'the general attempt to criticize a social order by showing the means and processes whereby bounded ideas and norms become historically legitimated disciplining or repressive practices' (2001b, p. 7). For reasons of space, I must leave these forms out of consideration here.

3 It is not obvious at the outset just what the status of this claim is: is it empirical or perhaps conceptual or transcendental? It sometimes seems as if Honneth would like to answer 'all of the above', but that will hardly do. Hegel's theory should probably be understood as making *conceptual* or *transcendental* claims. On this interpretation, recognition is not a *means* for us to become autonomous, for example; rather, being recognized is *constitutive* of autonomy and other forms of self-relation. A free agent in the Hegelian sense does not face other agents as external obstacles to her; rather, she is free in so far as she recognizes herself in the others and the others themselves in her. Standing in relations of mutual recognition is what constitutes freedom. This construal might solve some of the problems I introduce below, but Honneth does not, to my knowledge, advance arguments for it.

4 Charles Taylor has often argued that attitudes and emotions that are not expressed exist only in an implicit or potential form, possibly inchoately and mixed with other attitudes. Elizabeth Anderson has a nice example of what this means for a recognitive attitude: 'The geek who claims to love his girlfriend, but who gives her clumsy or insulting gifts, has thereby failed to love her adequately. His love is not realized, and indeed there is room to question whether it really is romantic love as opposed to mere fondness, or perhaps a goofy infatuation, or an immature fear of intimacy' (Anderson [1996, p. 547]). Anderson's point is that there are intersubjective norms for expression that form a part of the truth conditions for attitude attributions.

5 One major question is whether Honneth has the means to distinguish between recognition by concrete others and recognition by institutions. He seems to waver between the two views. The problem is that it is empirically plausible that people are able to acquire the kinds of self-relation Honneth speaks about in person-to-person contexts rather than in institutional ones – it is, for example, extremely implausible that no Finn was autonomous a thousand years ago, which would be the case if Honneth's argument that equal legal recognition is a precondition of personal autonomy were correct. However, if and when this is so, then Honneth cannot argue that the struggle for legal recognition (for example) is motivated by the quest for autonomous self-relation. Unless Honneth manages to show us the relevance of institutional recognition, he is pushed away from critical social theory to moral philosophy and thus loses an important part of the explanatory force of the theory.

6 While there may be pressure toward more and more extensive recognition, there can also be tension and conflict between the different forms: people might lose self-esteem while gaining self-respect, for example. Avishai Margalit, whose independently developed theory shares important similarities with Honneth's, uses the example of former socialist countries, where people 'could easily find themselves in a position of losing their self-esteem because they lack a useful role in the new economic and social order, yet they would no longer be

forced to compromise their integrity and self-respect, as was the case under the old regime' (Margalit [1996, p. 46]).

7 The sanction may serve as partial motivation for complying with the norm, but, as Elizabeth Anderson has argued, that cannot be the whole story, because how do we then explain the motivation for sanctioning? (Anderson [2000, pp. 181–4]). It cannot be that sanctioning itself is sanctioned – although that it often is in the case of moral norms, which are located a step or two higher on what Simon Blackburn calls the 'staircase of emotional ascent' (1998, p. 9) – because that would lead to infinite regress. If, then, we are motivated to sanction because of our belief in the norm's validity, then we can by the same token be directly motivated to act by that belief (Anderson, op. cit.).

8 Honneth does suggest, however, that even redistributive struggles have a recognitional dimension. He refers to E. P. Thompson's studies, which show that economic hardship as such are not enough for birth of resistance movements: 'what counts as an unbearable level of economic provision is to be measured in terms of the moral expectations that people consensually bring to the organization of the community' (Honneth [1995, p. 166]). One could begin to investigate the relatively easy acceptance of low-wage and dangerous child and adult labor conditions in Third World *maquiladoras* from this angle.

9 Note that Honneth's theory, as I am construing it here, allows indirect social scientific explanatory relevance to wrong-making properties themselves: the fact that a state of affairs is morally wrong (i.e. there is no proper recognition) (*ceteris paribus*) explains emotional reactions and judgments, which in turn (*ceteris paribus*) explain action. It is thus structurally similar to naturalist moral realism (e.g. Sturgeon [1988]; Brink [1989]; for recent criticism, Leiter [2001]). For reasons of space, I cannot further examine this possible convergence here.

10 From the point of view of normative criticism, this situation calls for a 'disclosing critique of society', such as the one Honneth finds in his recent re-reading of *The Dialectic of the Enlightenment* (Honneth [2000]). He characterizes it as a 'calculated attempt to change the preconditions under which evaluative discourses on the ends of common action are conducted in a society' by means of provoking 'a changed perception of parts of our apparently familiar lifeworld so that we will become attentive to their pathological character' (2000, p. 124). Honneth suggests that Adorno and Horkheimer achieve this by using disturbing rhetorical devices like narrative metaphor (Ulysses), chiasmus ('culture industry'), and exaggeration (de Sade as the exemplary bourgeois) (op. cit., pp. 124–5).

11 I thank Joel Anderson, Dara Connolly, Stephen Dodd, Heikki Ikäheimo, Arto Laitinen, Jarno Rautio, and Teemu Toppinen for stimulating comments and discussion.

REFERENCES

Anderson, Elizabeth 1993. *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.
Anderson, Elizabeth 1996. 'Reasons, Attitudes, and Values: Replies to Sturgeon and Piper.' *Ethics* 106, 538–54.
Anderson, Elizabeth 2000. 'Beyond Homo Economicus: Recent Developments in the Theory of Norms.' *Philosophy and Public Affairs* 29, 170–200.
Blackburn, Simon 1998. *Ruling Passions. A Theory of Practical Reasoning*. Oxford: Clarendon Press.
Brandom, Robert 1994. *Making It Explicit*. Cambridge, MA: Harvard University Press.
Brink, David 1988. *Moral Realism and the Foundations of Ethics*. Oxford: Oxford University Press.
Habermas, Jürgen 1985. *Der philosophische Diskurs der Moderne*. Frankfurt am Main: Suhrkamp.
Habermas, Jürgen 1996. 'Eine genealogische Betrachtung zum kognitiven Gehalt der Moral', in *Die Einbeziehung des Anderen*. Frankfurt am Main: Suhrkamp.
Hardimon, Michael O. 1997. Review of *The Struggle for Recognition*. *The Journal of Philosophy* 94 (1), 46–54.

498   *Antti Kauppinen*

Helm, Bennett 2001. *Emotional Reason. Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press.

Honneth, Axel 1991. *The Critique of Power. Reflective Stages in Critical Social Theory*. Trans. Kenneth Baynes. Cambridge, MA: MIT Press.

Honneth, Axel 1995. *The Struggle for Recognition. The Moral Grammar of Social Conflicts*. Trans. Joel Anderson. Cambridge: Polity Press.

Honneth, Axel 1997. 'Anerkennung und moralische Verpflichtung.' *Zeitschrift für philosophische Forschung* 51 (1), 25–41.

Honneth, Axel 1999. 'Mutual Recognition as a Key for a Universal Ethics.' Trans. John M. M. Farrell. On the Internet at http://www.unesco.or.kr/kor/science_s/project/universal_ethics/asianvalues/honneth.htm

Honneth, Axel 2000. 'The Possibility of a Disclosing Critique of Society: The *Dialectic of the Enlightenment* in Light of Current Debates in Social Criticism.' *Constellations* 7 (1).

Honneth, Axel 2001a. 'Invisibility: On the Epistemology of Recognition.' Trans. Maeve Cooke and Jeff Seitzer. *Proceedings of the Aristotelian Society Supplementary Volume* 75, 111–26.

Honneth, Axel 2001b. 'Reconstructive Social Critique with a Genealogical Reservation: On the Idea of Critique in the Frankfurt School.' *Graduate Faculty Philosophy Journal* 22 (2), 3–12.

Horkheimer, Max 1937. 'Traditionelle und kritische Theorie', in *Kritische Theorie. Band II*. Ed. Alfred Schmidt. Frankfurt am Main: S. Fischer Verlag, pp. 137–91.

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

Leiter, Brian 2001. 'Moral Facts and Best Explanations.' *Social Philosophy and Policy* 18 (Summer 2001), 79–101.

Margalit, Avishai 1996. *The Decent Society*. Trans Naomi Goldblum. Cambridge, MA: Harvard University Press.

McDowell, John 1996. 'Two Sorts of Naturalism', reprinted in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press, pp. 167–97.

Rawls, John 1996. *Political Liberalism*. Paperback Edition. New York: Columbia University Press.

Rorty, Richard 1991. *Objectivity, Realism, and Truth. Philosophical Papers Volume 1*. Cambridge: Cambridge University Press.

Smith, Michael 1995. 'Internal Reasons.' *Philosophy and Phenomenological Research* LV (1), 109–30.

Smith, Robin 1995. 'Logic', in *The Cambridge Companion to Aristotle*, ed. Jonathan Barnes. Cambridge: Cambridge University Press, pp. 27–65.

Sturgeon, Nicholas 1988. 'Moral Explanations,' in *Essays on Moral Realism*, ed. Geoffrey Sayre-McCord. Ithaca, NY: Cornell University Press, pp. 229–55.

Tuomela, Raimo 1995. *The Importance of Us. A Philosophical Study of Basic Social Notions*. Stanford: Stanford University Press.

Velleman, David 1996. 'The Possibility of Practical Reason.' *Ethics* 106, 694–726.

Zurn, Christopher 2000. 'Anthropology and Normativity: a Critique of Axel Honneth's "Formal Conception of Ethical Life".' *Philosophy and Social Criticism* 26 (1), 115–24.

Antti Kauppinen, Department of Moral and Social Philosophy, PL 9 00014 University of Helsinki, Finland. E-mail: antti.kauppinen@helsinki.fi