Shared Autonomous Reasoning: Interpretations of Habermasian Discourse for The Community of Philosophical Inquiry

Natalie M. Fletcher

INTRODUCTION

Can autonomous reasoning be shared? According to many philosophical perspectives, from Kantian ethics to libertarian theories, this question seems incoherent – the purpose of an individual being able to think rationally for herself, to determine what she finds important and advance her own ends in accordance, seems to lose its appeal if she must engage in this process jointly with others and result in having the same concerns, perspectives, and goals in common. A crucial motivation for many accounts of autonomy is to safeguard the individual's will from the influence or interference of the other, and enable her to authentically express the identity and life she freely chooses for herself. From this vantage point, what could be gained from having the capacity for autonomous reasoning be shareable?

Throughout his writings and notably in his theory of communicative action, Jürgen Habermas examines the potential for shared autonomous reasoning, challenging monological approaches in favour of a discursive understanding that seeks to preserve the emancipatory features of popular notions of self-determination while adding a crucial intersubjective component.[1] From his perspective, it appears that autonomous reasoning not only can be shared but indeed must be shared, given his stance that all human meaning is intersubjectively constituted. But what does this particular construal of autonomy involve? As Gerald Dworkin has noted, autonomy is a term of art—its characterization varies depending on the field and usage, making it a notoriously difficult concept to define without compromising the intricacy of the facets of human reality it denotes.[2] This burden of characterization intensifies when considering the breadth of the Habermasian corpus.

To begin, then, if we define autonomous reasoning simply as the capacity to freely and willingly engage in critical processes of reason generation and justification, the question of whether it can be shared is interpretable in three ways. First, on a descriptive interpretation, we can argue that autonomous reasoning is shared in that it takes place in a context of common understandings and meanings established linguistically that gives us the basis for our reasons and makes interpersonal exchanges possible. Second, on a normative interpretation, we can argue that autonomous reasoning

ought to be shared in that we are responsible for ensuring no one lives under norms they do not themselves endorse, and therefore, we must include them in the collaborative process of generating and justifying reasons, and respect their capacity to do so. Third, on an epistemic interpretation, we can argue that autonomous reasoning benefits from being shared in that the combining of our efforts in terms of both procedure and content can increase knowledge and advance learning in ways that expand the scope and integrity of our collective communicative agency.

Through an exploration of these three interpretations, this essay will contend that autonomous reasoning can and must be shared, as Habermas would maintain, but that to fully benefit from this "sharedness," we must understand it as a capacity comprising a range of faculties driven not only by our commitment to establishing justifiable norms but also by a sense of integrity that recognizes others as epistemic agents whose worth stems from both their discursive aptitudes and concrete particularities. The essay will begin with an overview of the context for Habermas's interest in discursive autonomy, then consider the descriptive, normative and epistemic interpretations in turn, and end with a look at how shared autonomous reasoning might be honed through a Community of Philosophical Inquiry (CPI) practice.

I. Contextualizing Habermas's interest in discursive autonomy

Beyond the influence of Kantian ethics, Habermas's interest in autonomy is largely motivated by what he calls the "unfinished project of modernity."[3] As a philosopher who recognizes the historical situatedness of philosophical concepts and ideologies, Habermas has celebrated the significant positive contributions of modern social life, notably the increases in individual freedom and knowledge as well as the plurality of perspectives and orientations resulting from the decline of dominant religious traditions. At the same time, he has denounced the rise of overly scientistic and instrumentalist worldviews that alienate people from moral engagement and threaten social cohesion, claiming that we have yet to live up to the new expanses of knowledge that modernity has afforded us.[4] And so, modernity remains a project to be completed, in part through the exercise of discursive autonomy that preserves the communication modes necessary for moral norms to be created and followed in a world no longer ruled by divine codes of conduct. As a critical theorist, Habermas has sought to conceive autonomy in ways that protect individuals from exclusion, coercion, groupthink, and political and economic repression, while making them accountable to each other and their social contexts. Politically, he has borrowed ideas from liberal democracy and civic republicanism to express the importance of balancing autonomy's private and public dimensions, which he sees as interdependent: on the one hand autonomy is a kind of self-determination that allows individuals to freely pursue life projects with minimal governmental interference, and on the other hand, it is the collective will of people who self-legislate by recommending their views to their representative governments in the public sphere.[5]

Through an interdisciplinary rational "reconstructive" method combining pragmatist theory with empirical sciences,[6] Habermas proposes a dualistic view of society as comprising two realms of activity in which autonomy can be exercised, albeit in very distinct ways: the "lifeworld" and the "system." The lifeworld consists of "the culturally transmitted and linguistically organized stock of interpretive patterns"[7] that shape our everyday unregimented interactions with others, from our families and communities to our mass media and grassroots political projects: it contains the diversity of shared meanings and understandings—the "vast and incalculable web of presuppositions"[8]—that make communicative action possible. As a self-sustaining force that enables cultural reproduction, the lifeworld can only undergo gradual change given its deep-rooted complexity. In contrast, the system designates the sphere of society's material reproduction of goods and services governed by money and power, which in turn are wielded by instrumental and strategic action. When people act instrumentally in the system, they use their reasoning to calculate how best to reach their desired ends through targeted means, including acting strategically to influence others in ways that will support the realization of their chosen ends. Unlike Marxists and fellow Frankfurt School theorists, Habermas recognizes the significance of action in both realms; however, he argues that since the system lacks transparency in its aims and imposes external restrictions on agency, if its reach extends too far, it can lead to "systematically distorted communication" that creates social pathologies like alienation, demoralization and instability that upset the lifeworld and result in its colonization. He writes: "Such communication pathologies can be conceived of as the result of a confusion between actions oriented to reaching understanding and actions oriented to success."[9]

In response to these increasingly complex circumstances, we must protect the lifeworld by exercising our autonomous reasoning to come to mutual understandings and agreements about how best to live and how we ought to treat one another—questions of the good life (ethics) and questions of the right and the just (morality). As subsequent sections will strive to elucidate, this process of communicative action occurs through our everyday exchanges, whereby we coordinate our actions by using validity claims and, when disagreement occurs, by engaging in specific types of discourse that enable us to determine the ethical values and moral norms best suited to guide our actions. For Habermas, through the process of modernization, ethics and morality have grown apart: without the overarching grip of religious codes and beliefs, people have had to collaboratively decide for themselves what moral norms ought to universally apply to maintain social order and resolve conflicts, while acknowledging that the broadening array of worldviews and orientations has resulted in vastly different ideas about what makes a life worthwhile for individuals and their respective communities. As Habermas explains,

> At first glance, moral theory and ethics appear to be oriented to the same question: What ought I, or what out we, to do? But the 'ought' has a different sense once we are no longer asking about rights and duties that everyone ascribes to one another from an inclusive 'we' perspective and ask what is best 'for me' or 'for us' in the

> long run and all things considered. Such ethical questions regarding our own weal and woe arise in the context of a particular life history or a unique form of life. They are wedded to questions of identity: how we should understand ourselves, who we are and want to be.[10]

In both the moral and ethical domains, shared autonomous reasoning plays a key role, but questions of the right and the just require that it be exercised in very particular ways toward the crucial objective of establishing reasonable norms to which we can all agree by virtue of our common humanity.[11]

And so, it would appear that on Habermas's account, autonomy functions differently—albeit always in some sense intersubjectively—depending on the context in which it is employed. Here, Joe Anderson's analysis of the five senses of autonomy that emerge from Habermas's theory of communicative action is helpful: (i) within the theoretical context of deliberative democracy, political autonomy involves the freedom from "illegitimate domination by others" and appropriate integration into "processes of collective self-determination;" (ii) within the theoretical context of moral philosophy, moral autonomy involves the capacity of allowing "intersubjectively shared reasons to determine one's will;" (iii) within the theoretical context of free will, accountable agency involves the wherewithal to "act for reasons" rather than "as a result of compelling forces;" (iv) within the theoretical context of social theory, personal autonomy involves the ability to "engage in critical reflection about what do with one's life" and pursue it without violating moral norms; (v) within the theoretical context of personal identity, accountable identity involves "vouching for oneself and being recognized by others for so doing."[12] For present purposes, we will focus on the tensions and parallels between moral autonomy and personal autonomy—or what Habermas calls "ethical-existential" autonomy[13]—and how these relate to the prospects of shared autonomous reasoning from the descriptive, normative and epistemic interpretations previously outlined.

II. A descriptive interpretation of shared autonomous reasoning

A straightforward approach to the "shareability" question at hand could simply be the following: to describe autonomous reasoning as shared is to describe how it actually happens—what is involved when we employ the capacity to freely and willingly engage in critical processes of reason generation and justification. But is this in fact the case? Suppose that person X is on her own, thinking about the possibility of becoming a vegan. Though this decision has clear moral and ethical implications, let us assume for now that she is thinking only of the feasibility of veganism in her current life circumstances. No one is forcing her to consider this topic: she is not being coerced or pressured; she is doing so freely—generating reasons, critically comparing them and determining which are most justifiable. At this stage, she is alone, not in dialogue with others. In this situation, is autonomous reasoning shared? It would seem, at the very least, that the contents of X's autonomous reasoning and its worth as an activity are shared: she can articulate reasons because of linguistic and

cultural parametres that she has come to adopt through her social embeddedness and her interaction with others, and she can deem the process itself as worthwhile because its value has been considered in the historical context in which she finds herself. It would seem odd of her to claim either of these features as strictly her own or solely the products of her own independent thought. To be reasoning autonomously in this case does not mean to be the originator of the contents and valuation of her activity, even though she has willingly engaged in it for herself.

And so, on a descriptive interpretation of Habermas, we can argue that autonomous reasoning is shared in that it takes places in a context of common understandings and meanings established linguistically in the lifeworld and expanded through collective language use that gives us the basis for our reasons and makes interpersonal exchanges possible. Since the lifeworld resources from which we draw when we engage in autonomous reasoning—or the "stuff" of reasoning (language, connotations, meanings, reasons themselves, etc.)—are intersubjectively constituted, when we use any of them, we are drawing from a pool of already common resources that are co-constructed with others.[14] On our own, we cannot make the rules of our own reasoning—we cannot single-handedly decide how to determine what is true, right or meaningful without recourse to others with whom these decisions are made. As Pablo Gilabert notes, "This is why Habermas prefers to talk of 'communicative reason' instead of 'practical reason,' like Kant [since it] requires practices of justification to be dialogic (or discursive) rather than monologic (or introspective)."[15] Further, the lifeworld resources that were created before us by others through their communicative action form part of the background of assumptions and significances in which we too are embedded so when we reason autonomously, we are necessarily sharing in what has already been established, even if our goal is to challenge and refine it. For example, in the very process of my writing about Habermas, I cannot cut myself off from the shared understanding of words and their various connotations, and I am aware that some terms in Habermasian philosophy will have very particular meanings in my current context of autonomous reasoning than they will in others, which enjoins me to be clear and cull from these pre-existing and evolving meanings in ways that will make sense of my own thoughts and be reasonable to others.

In a more sophisticated sense, this descriptive interpretation of shared autonomous reasoning reflects Habermas's interest in speech acts and pragmatic meaning theory. From his perspective, truth conditions are inadequate at explaining how language enables our various forms of communication and action because we do not speak purely to describe the world as it is, but also and more importantly, to make meaning intersubjectively through the giving and weighing of reasons in order to co-construct justifiable norms.[16] As James Gordon Finlayson writes,

> Habermas argues that the primary function of speech is to coordinate the actions of a plurality of individual agents and to provide the invisible tracks along which interactions can unfold in an orderly and conflict-free manner. Language can fulfill this function because of its inherent aim (or telos) of reaching understanding or

> bringing about consensus. Habermas takes it to be a fact that 'reaching understanding inhabits human speech as its telos.'[17]

Returning to the veganism scenario, if person X wants to really ascertain the reasonableness of her prospective change in diet, she would have to communicate her thoughts on the matter to others, thereby making a commitment to providing sound reasons to justify herself—or what Habermas calls "validity claims"[18]—and to having these claims evaluated by them. If, for instance, she tells her extended family that she prefers that they not serve meat dishes at their reunion dinner and that they object to her expressed wish on the grounds that traditionally they have always eaten meat dishes, this disagreement will create an impasse that cannot be settled with everyday speech-acts. For Habermas, when communicative action in the lifeworld is so interrupted, practical discourse is the more refined speech mode through which to share autonomous reasoning in the form of rational exchanges and evaluations of reasons that seek to resolve conflict, re-establish consensus and return to a mode of harmonious action. To function in this way, our discourse—or "form of argumentation" that is "norm-justifying"[19]—must meet a series of requirements: first, in terms of logic, we must ensure the products of our arguments are cogent, consistent and non-contradictory; in terms of dialectic, we must ensure the procedures of our arguments are guided by the principles of accountability and truthfulness, which helps our speech-acts reach the illocutionary aims of being transparent and understandable; and in terms of rhetoric, we must strive to ensure the process of our arguments meet the presuppositions of inclusion and equality that characterize an "ideal speech situation"—a notion to which we will return in the next section.[20]

So far, then, we can argue that our autonomous reasoning is shared at least in terms of its contents and worth as an activity since we draw from intersubjectively constructed lifeworld resources to critically generate and justify our reasons, and when these fall short, we use discourse to refine and re-establish the validity claims that will coordinate our actions, until we have cause to consider them afresh. This descriptive interpretation has its advantages: it does away with the illusion of an atomistic, solipsistic ego popularized by the philosophy of consciousness that Habermas rejects, and it highlights the powerful ways in which our sense of reasonableness is embedded linguistically, historically and socially.[21] But is it enough to say that our autonomous reasoning is shared because of our common language and means of communication? Could we share lifeworld resources without necessarily sharing equal participation in discursive autonomy?

III. A normative interpretation of shared autonomous reasoning

From a Habermasian perspective, it is insufficient to consider only the contents and the valuation of autonomous reasoning as shared—we also have an obligation to the people in the lifeworld with whom we intersubjectively create meaning. Given the atrocities he witnessed during the Second World War, Habermas has been steadfast in his envisioning of a deliberative form of

democracy that eschews exclusionary tactics and violations of human dignity. On a normative interpretation, then, we can argue that autonomous reasoning ought to be shared in that we are responsible for ensuring no one lives under norms they do not themselves endorse, and therefore, we must include them in the collaborative process of generating and justifying reasons, and respect their capacity to do so. Kenneth Baynes calls this the "sociality of reason" since it suggests that "reflective endorsement is not a solitary endeavour but requires social practices of justification that include other reason-givers or 'co-deliberators.'"[22] The recognition of my own capacity for autonomous reasoning and the contributions it enables must be mirrored by my acknowledgement of this capacity and potential in others, thus resulting in the shared accountability of being critically responsive to one another's validity claims.

In Habermas's estimation, this normative dimension of shared autonomous reasoning is especially pertinent in the realm of morality. He argues that moral norms are dynamic, evolving human constructions that are established and refined through moral discourse, and delimit our overall communicative action.[23] They do not originate within us, they do not exist independently of us, and they do not come from a higher power—they are the result of our attuned intersubjective exchanges as agents capable of discursive autonomy. In light of their crucial role, it is imperative that moral norms be co-constructed in ways that would be deemed valid by all those affected by them, or else they will control the actions of people who have not contributed to the reasons justifying them nor agreed to their acceptability, and whose capacity as autonomous reasoners thus risks being disregarded. For instance, if person X finds herself living under the moral norm "Thou shall kill animals for meat," because she has been excluded from the critical collaborative process of generating and justifying reasons, we cannot say that the autonomous reasoning behind the norm has been adequately shared nor that real consensus has been reached. In this case, we could say she has been intersubjectively cheated of the opportunity to challenge validity claims to rightness and provide reasons that may change the outcome of a norm's endorsement—and this lacks both respect for and responsibility to her capacity for shared autonomous reasoning, to say nothing of the possibility of her being coerced into renouncing her position.

To guard against such occurrences in discourse, Habermas specifies key requirements of "ideal speech situations" that reinforce their collaborative spirit:

> (i) Every subject with competence to speak and act is allowed to take part in a discourse.
>
> (ii) a. Everyone is allowed to question any assertion whatever.
>
> b. Everyone is allowed to introduce any assertion whatever into the discourse.
>
> c. Everyone is allowed to express his attitudes, desires, and needs.
>
> (iii) No speaker may be prevented by internal or external coercion, from exercising rights as laid down in (i) and (ii).[24]

Moreover, Habermas identifies two principles of discourse that are intended to ensure broad acceptability: first, the discourse principle (D) affirms that "Just those action norms are valid to which all possibly affected persons could agree as participants in rational discourse"[25]—if this turns out not to be the case, or even not to be anticipated, then the norm cannot be adopted; the criterion is a very demanding one that applies to moral, ethical and practical claims. Second, specific to questions of the right and the just, and more demanding still, the universalization principle (U) affirms that a moral norm is valid if and only if "All affected can accept the consequences and the side effects its general observance can be anticipated to have for the satisfaction of everyone's interests."[26] Together, these presuppositions and principles incorporate mutual respect, solidarity and responsibility into discursive reasoning by requiring that we engage in a "universal exchange of roles" that allows us to see from the perspective of what George Herbert Mead calls a "generalized other" and thereby recognize that "valid norms must deserve recognition by all concerned."[27] As Thomas McCarthy asserts, "Habermas's discourse model, by requiring that perspective-taking be general and reciprocal, builds the moment of empathy into the procedure of coming to a reasoned agreement."[28]

So we can argue that beyond its intersubjectively constituted contents and valuation, our autonomous reasoning ought to be shared so that its process of contributing to norms, and by extension to communicative action, is as inclusive and equal as possible—that we embrace our responsibility to take into account the perspectives of all those affected and ensure respect for those who participate in discursive autonomy through argumentation that is open, unrepressed and non-coercive. Of course, Habermas grants these conditions are idealized; the realities of time constraints, massive populations and everyday obligations might well jeopardize our efforts towards autonomous reasoning that is shared in these ways, though he maintains they are feasible in principle.[29] One strength of the normative interpretation, as Rainer Forst and Jeffrey Flynn highlight, is that "nobody claims special privileges and everyone grants others all the claims one raises for oneself, without projecting one's own interests, values, or needs onto others and thereby unilaterally determining what counts as a good reason."[30] Yet the shared facet of autonomous reasoning is not meant to obscure the individual's freedom of thought: as Anderson writes, "having one's will determined by reason does not undermine one's self-authorship, especially once it is clear that 'listening to reason' is a matter of engaging, as a full and equal participant, in the ongoing process of giving and asking for reasons."[31] But can we really co-construct norms to which we all rationally agree? Or would some individuals inevitably end up settling, living under norms that they have accepted against their better judgment because of the influence of other overpowering factors, like the longing to belong, to fit in, to not call attention to themselves by destabilizing the status quo? Or, more plainly, would some individuals reluctantly acquiesce to certain norms because they are unable to think of better formulations for them, despite intuiting that they are not strong enough to be satisfactory? Habermas might respond that these are simply not instances of real discourse; perhaps the agents' capacity for shared autonomous reasoning is lacking in important ways. If so, however, how might levels or varieties of

this capacity lead to exclusion, despite the presuppositions and principles theoretically in place to support equal and broad participation?

V. An epistemic interpretation of shared autonomous reasoning

According to Habermas, we develop the capacity for autonomy through a learning process that combines cognitive, psychological, social and moral development, and only when we reach a certain stage of "post-conventional morality"—a phrase he borrows from psychologist Lawrence Kohlberg—can we really begin to engage in discourse about norms.[32] Learning to hone discursive autonomy involves knowing about the procedures and content of discourse—the components of rational communication and the rightness of norms—and results from an ongoing practice that conditions us to be more reflective, critical and responsible agents.[33] However, for Habermas, no matter how much we learn and come to know, we must remain perpetually open to revising our norms and related actions because we are fallible—our consensus does not entail rightness as we may be mistaken.[34] Even given our shared autonomous reasoning, "The valid moral norms legislated and internalized by morally autonomous agents thus represent our current best efforts in the ongoing process of learning to solve the moral challenges continually posed by life and raised in discourse."[35] But what else does this process of learning and knowledge acquisition involve?

Thus far, we have been defining autonomous reasoning as the capacity to freely and willingly engage in critical processes of reason generation and justification, and we have qualified it as shared in that we need co-constructed resources to undertake it as a jointly valued discursive activity, the products of which can only be deemed worthy of delimiting our actions if they meet exacting principles aimed at inclusion and equality, and result from reciprocal consideration of each other's perspectives. But upon closer inspection, the capacity for shared autonomous reasoning appears to involve faculties that are far more complex than intimated by the descriptive and normative interpretations. These faculties include but are not limited to: first, discerning the most pertinent lifeworld resources from which to draw in a given situation while doing our utmost to be aware of the assumptions that colour our justifications, notably the biases to which we know we can fall prey. Second, scrutinizing and honestly appraising whether we are forming judgments in light of our real commitment to listening to others' reasons rather than our engagement with overpowering systemic factors by which we might unknowingly be diverted and even crippled. Third, exercising very acute and wide-ranging evaluative skills that enable us to discriminate between good and bad reasons, detect fallacious thinking and manipulation tactics, and pinpoint what is missing in an argument when we are not convinced by it—and clearly articulating all of this in argument form. And fourth, exhibiting enough resilience and strength of character to employ these aforementioned faculties even in less than conducive circumstances where we might feel ill-equipped or in over our heads. In short, we need to know a lot: about ourselves, about each other, about our individual and collective strengths and weaknesses, about our current and past contexts, about the nature of reasonableness and the

myriad demands it makes on us. It seems, as Anderson suggests, that "autonomy is not something we can pull off by ourselves."[36]

And so, given this daunting endeavour, on an epistemic interpretation, we can argue that autonomous reasoning benefits from being shared in that the combining of our efforts in terms of both procedure and content can increase knowledge and advance learning in ways that expand the scope and integrity of our collective communicative agency. With regard to scope, it seems likely that we will be better at autonomous reasoning and ensure it covers larger ground if we learn from others and grow from their knowledge, divvying up the responsibilities of communicative action and discourse so we complement and make up for each other's shortcomings. Yet these benefits in the form of epistemic gains seem to also entail a particular obligation. The fallibilist nature of our judgments (the reality that we know nothing for certain) and our own fallibility (the reality that we are error-prone) do not imply that we are incorrigible in how we share our autonomous reasoning. Indeed, with regard to integrity, our learning from others and access to their knowledge bases should arguably impel us to recognize them as epistemic agents who contribute to our autonomous reasoning not only by virtue of their discursive faculties but also, more controversially for the Habermasian agenda, because of their particularities as individuals with their own sense of what is good—or their own ethical-existential autonomy—which is experienced as concretized not generalized.

This epistemic interpretation goes further than Habermas would probably condone but it is worth exploring. If we return to the example of person X and her issue of veganism, even if her autonomous reasoning is shared in the descriptive and normative senses previously examined, something may still be amiss. The contents and valuation of her reasoning may be intersubjectively co-constructed; she may be competent as an autonomous reasoner and allowed to take part in discourse about the morality of animal consumption in the form of questioning and introducing assertions, and expressing her attitudes, desires, and needs, all in a seemingly non-coercive atmosphere. She may even end up agreeing to a norm different than the one she set out to defend. But could she still somehow be settling or acquiescing prematurely for reasons that have more to do with factors influencing her capacity for autonomous reasoning than with her apparent equal inclusion as a discursive agent? Put differently, could there be threats to the "sharedness" of her autonomous reasoning that impair its faculties in ways that increase the risk of her exclusion?

Conceivably, her fellow interlocutors could be including her in discourse but not fully recognizing her as an epistemic agent and thus not benefitting from her knowledge to the extent outlined above. Beyond her discursive faculties, there may be a host of particularities relevant to the norm in question that are being overlooked but could enrich the collective critical processes of reason generation and justification—say, as random examples, her cultural heritage growing up in remote and fragile mountainous ecosystems, her applied research training in zoology, her childhood experience with animal-assisted therapy due to a congenital disability, her avid interest in alternative and sustainable approaches to nutrition, her skills of synthesis and her spirited sense of compassion.[37] These distinctive details that have in part formed and in other ways resulted from her ethical-

existential autonomy—or recalling Anderson's definition, her ability to "engage in critical reflection about what do with [her] life"[38]—surely must be affecting her faculties of autonomous reasoning on some level. To bracket them from consideration would seem to discount a significant part of her learning and knowledge. And yet certain threats to the "sharedness" of autonomous reasoning may do just that, as this section's examination of unacknowledged conformity and unstated privilege will seek to show, with reference to critiques of Habermas from within deliberative political philosophy and feminist theory.

a) The threat of unacknowledged conformity

The presuppositions of discourse that Habermas proposes are in part aimed at guarding against external and internal coercion, but their formulation assumes that we are aware and realize when such coercion is happening. We may, at times obliviously, live under norms that we have accepted not because of their validity but due to the influence of other overpowering systemic factors affecting the way we learn and know—or, in short, due to the threat of unacknowledged conformity. Here, knowing about the particularities of people and their circumstances, and how these connect to their exercise of ethical-existential autonomy, may reveal how we can better share autonomous reasoning. For instance, citing the example of stoic slaves, Joshua Cohen notes how unfavourable, unjust conditions for autonomy can produce "accommodationist preferences" in people who deliberately choose to subordinate themselves because they have no other alternative.[39] Sharing the task of reflecting on how such preferences are formed and how they may exclude people in ways that go unnoticed may help us to prevent their status as epistemic agents from being undermined. Similarly in terms of power relationships, Johanna Meehan argues that dynamics of domination may be ingrained in how we form our very identities before we even engage in moral discourse, making members of marginalized groups all the more vulnerable to those who have been in a sense "raised" to exclude them: "When the fabric of a child's relationship to self and to other is woven in threads of domination, the seeds of disrespect and domination are sown, and children may grow to be adults whose very construction of others undercuts the possibility of respect."[40] For those being marginalized, the result may be a tendency to conform to norms rather than accept them, due to a distorted, discrediting sense of their worth as epistemic agents or out of a need to not call attention to themselves by destabilizing the status quo.

Moreover, Habermas has claimed that during discourse, in principle, "nothing coerces anyone except the force of the better argument."[41] But the force of the better argument may come from a very forceful argumentator whose feigned interest in the epistemological vantage points of others is tokenistic at best. According to the universalization principle, if an individual thinks a norm is right, she must anticipate that others will agree—if not, she ought not to be rationally convinced by it herself. But where does this leave the perspectives of disenfranchised people whose specific particularities may yield normative considerations that ruffle the feathers of the more forceful voices

and stand no chance of motivating consensus? It seems an overemphasis on commonality may also breed unacknowledged conformity. Concerned about the eclipsing of differences, Seyla Benhabib has recommended that we recognize not only the perspective of "the generalized other" as worthy but also that of the "concrete other," which "requires us to view each and every rational being as an individual with a concrete history, identity and affective-emotional constitution."[42] In so doing, we will move away from construing shared autonomous reasoning as an abstracted, existentially disconnected process of reason generation and justification that is only successful if it reaches consensus, and instead see it as a practice that trains us to be more reflexive and comprehensive so as to acknowledge our commonality as well as our multiple differences, and promote a more genuine reciprocal recognition. In her words: "The emphasis now is less on rational agreement, but more on sustaining those normative practices and moral relationships within which reasoned agreement as a way of life can flourish and continue."[43]

b) The threat of unstated privilege

On a related note, Habermas's presuppositions of discourse state that every subject "with competence to speak" can participate, while his take on moral development contends that this competence is defined in no small part by our having reached a "post-conventional" moral consciousness.[44] But does this stance exclude people based on their age, their upbringing, their maturity, their education, their faith, their mental acuity, their psychological stability, their moral compass—in short, the host of particularities that have shaped what they have learnt and come to know, and how? If autonomous reasoning is indeed a capacity that we develop, do we need a certain amount of it before we can even be considered as epistemic agents with validity claims worth examining? More troubling still, do we all equally share in the capacity, or is there unstated privilege that exists as a result of some individuals being more predisposed to autonomous reasoning or having more opportunities to enhance it? Here, Philip Pettit's distinction between virtual and actual capacities is useful. He describes an actual capacity as "a capacity that is ready to be exercised" and a virtual capacity as "a capacity that is yet to be fully developed," using the example of an individual who does not play the piano but might discover musical gifts if he tried.[45] While the would-be pianist's virtual capacity should not be dismissed, he cannot reasonably be said to be capable of piano-playing— nor be evaluated for this aspect of his agency—until he can actually play, learn to play or declare himself musically inept. Correspondingly, regardless of our virtual potential for faculties of autonomous reasoning, we can really only benefit from sharing it if it is "ready to be exercised." Even supposing we do all have the same virtual capacity for autonomous reasoning—a contentious assumption—do we all actualize it to the same degrees, or do some of us by some privilege get more use out of it and thus more epistemic clout? It seems all too possible that those of us without the actual capacity for autonomous reasoning, by no fault of wanting or trying, risk being ineligible for

sharing in critical processes of reason generation and justification, which affects the scope of the collective knowledge that drives discursive autonomy and affects its integrity as an "inclusive" practice.

Further, even if we do actually possess faculties of autonomous reasoning like those listed earlier in this section, the modes in which we are required to use them might themselves incorporate unstated privilege—we might be presuming that given the same chance and access, we will all communicate in the same ways. Iris Young has argued that Habermasian-type discourse approaches "assume a culturally biased conception of discussion that tends to silence or devalue some people or groups" while elevating others, and affect the "internalized sense of the right one has to speak or not to speak.[46] In her view, the kind of discourse Habermas envisions tends to be assertive, competitive, combative, dispassionate and disembodied, and use direct, literal language rather than "speech that is tentative, exploratory, or conciliatory."[47] Accordingly, those individuals whose culture, gender, socio-economic status and education favour this mode of argumentative communication are likelier to thrive while the rest have to adapt to alien speech styles or risk having their epistemic vantage point excluded. To prevent these "powerful silencers of speech" that privilege some "strong" voices at the expense of other "weak" ones, Young maintains we must expand what counts as valuable communication forms beyond argumentation to include speech characterized by figurative language, emotion, humour and camaraderie, like rhetoric and storytelling. By revealing the meanings of the particularities that characterize people, narrative has the power to expand the scope and integrity of our collective communicative agency by giving us "social knowledge from the point of view of that social position," helping us "understand why the insiders value what they value" and recognize that "values, unlike norms, often cannot be justified through argument, but neither are they arbitrary."[48]

On the epistemic interpretation, then, for autonomous reasoning to really benefit from being shared, it seems we need to not only learn from others and their knowledge so we get epistemic gains that expand the scope of our collective communicative agency; we also need to be concerned with matters of integrity, addressing and correcting threats to the "sharedness" of our autonomous reasoning—like unacknowledged conformity and unstated privilege—that get in the way of our recognizing each other as epistemic agents whose worth as fellow interlocutors stems from both discursive faculties and concrete individual particularities. As Cristina Lafont stresses, the justice of a norm does not depend on whether we all agree on it since we could be wrong. Even if we reach unanimous agreement, "we still need to be vigilant to the (ever-present) possibility of undetected injustices and powerful ideologies that such agreements may contain."[49]

V. Shared autonomous reasoning in practice

Through a descriptive, normative and epistemic interpretation of Habermasian discourse theory as well as some of its critiques, this essay has claimed that autonomous reasoning is, ought to be and benefits from being shared, and has problematized what faculties it may involve as a capacity and the threats it may face. In closing, it is worth considering how shared autonomous reasoning

might be honed through an applied practice, namely the Community of Philosophical Inquiry (CPI) pedagogical model. To cultivate a sense of integrity that expands the scope of the sharedness of our autonomous reasoning and deepens our recognition of others as epistemic agents, we not only need to be able to critically generate and justify reasons, but to do so with epistemic virtues. As Baynes has argued, Habermas's principles of moral discourse cannot "be guaranteed by specifying formal features—the rules of argumentation—alone; they depend upon many other cognitive and empathic skills as well."[50] And within the realm of ethical-existential concerns, as Anderson observes, "The expansion in possibilities for choice brings with it an expansion in the responsibilities for choosing well," which must involve joint efforts toward being "maximally open to relevant considerations."[51]

A CPI may be the ideal setting for fostering the epistemic virtues that can help make autonomous reasoning truly shared and thereby support communicative action. Originally developed by educational philosopher Matthew Lipman as a philosophical practice for children, this pedagogical model aims to develop responsible, relational autonomy through multidimensional thought (or combined critical, creative and caring thinking), by challenging us to confront the contestable questions we deem central to our lives and seek reasonable judgments through structured group dialogue.[52] The CPI model shares many features in common with Habermasian discourse, notably its pragmatist roots, its fallibilist view of knowledge, its commitment to intersubjective meaning-making, its use of dialogic argumentation, its emphasis on communicative rather than instrumental rationality, its principles of equality, respect and inclusion, its concern over similar social "pathologies," and its desire for real-world relevance as a practice that can help people to interpret and understand the complexities of life. As CPI scholar Barbara Weber has noted, because this model "genuinely aims for understanding and simultaneously makes us aware of our differences as well as of our own prejudices," it can "provide the missing link to make Habermas's concept of communicative rationality more practicable by cultivating a natural illocutive intention in children"[53] and, for that matter, adults.

While the CPI model is not immune to threats that affect the "sharedness" of autonomous reasoning like unacknowledged conformity and unstated privilege,[54] its method builds in ways to avoid them: inquiry members are invited to consider the philosophical dimensions of a stimulus they experience together (like a story, art work or exploratory project), generalize from these to formulate open-ended questions that address issues of overall concern to humanity, deliberate over what is reasonable to think with respect to these questions, and bring in concrete examples from everyday life that can problematize the positions under consideration to make them more nuanced and applicable. In so doing, the CPI model seems primed to promote the "enlarged mentality" that Benhabib extols by cultivating judgment that "involves the capacity to represent to oneself the multiplicity of viewpoints, the variety of perspectives, the layers of meaning which constitute the situation."[55] If the CPI succeeds in its efforts, it is in no small part because of its focus on developing self-correction, which involves crucial epistemic virtues like intellectual humility, attentiveness, discernment, comfort with uncertainty, acceptance of fallibility, resistance to bias and a willingness to freely change

positions when reasonableness demands it. By helping us address questions of the right and of the good, and by fuelling both our moral and ethical-existential autonomy, the CPI and its epistemic virtues can contribute to what Richard Bernstein has called our "democratic ethos."[56] If we understand autonomous reasoning as a capacity comprising a range of faculties driven not only by our commitment to establishing justifiable norms but also by a sense of integrity that recognizes others as epistemic agents whose worth stems from both their discursive aptitudes and concrete particularities, we may bring the "sharedness" of discursive autonomy to new heights.

Endnotes

[1] This article will focus primarily on Habermas's mature positions as developed in the two volumes of his *Theory of Communicative Action* (1984, 1987) and related mid-career works.
[2] Dworkin, 1988, 7.
[3] Habermas, 1997, 38.
[4] Finlayson, 2005, 66.
[5] Habermas, 1996, 468.
[6] As Habermas writes, "the reconstructive sciences explain the presumably universal bases of rational experience and judgment, as well as of action and linguistic communication." Habermas, 1990, 16.
[7] Habermas, 1984, 70.
[8] Habermas, 1987, 130.
[9] Habermas, 1984, 332.
[10] Habermas, 2004, 32.
[11] As such, moral norms have universal validity whereas ethical values have only relative validity.
[12] Anderson, 2011, 91, 108.
[13] Habermas, 1993, 11.
[14] In this sense, we can say we have a linguistic dependence that translates into an intersubjective dependence: As Anderson writes, "when one acts for reasons, a full explanation of what one is doing must make reference to the cultural and linguistic background in virtue of which certain noises count as giving reasons." Anderson, 2011, 93.
[15] Gilabert, 2005, 408.
[16] Habermas, 1984, 277. Habermas also borrows from Richard Rorty, who writes: "Saying things is not always saying how things are." Habermas, 1990, 10.
[17] Finlayson, 2005, 34.
[18] Habermas, 1984, 9.
[19] Habermas, 1990, 19, 70.
[20] Referencing Aristotle, Habermas writes: "Rhetoric is concerned with argumentation as a process, dialectic with the pragmatic procedures of argumentation, and logic with its products." Habermas, 1984, 26. He offers logical-semantic rules as the departure for argument (Habermas, 1990, 87), ethical rules that promote mutual recognition (88) and rules for communication that avoids coercion (89).
[21] As Finlayson outlines, Habermas is suspicious of many features of what he calls the "philosophy of consciousness," notably Cartesian subjectivity and subject-object metaphysics. Finlayson, 2005, 28.
[22] Baynes, 2015, 91.
[23] As Baynes writes, "Moral rightness, according to Habermas, is 'constructed' not discovered." *Ibid*, 105.
[24] Habermas, 1990, 89.
[25] Habermas, 1996, 107.
[26] Habermas, 1990, 65.
[27] *Ibid*, 65.
[28] McCarthy in Habermas, 1990, viii.
[29] As Habermas writes, "The need to act in the lifeworld, in which discourses remain rooted, imposes temporal constraints on what is, from an internal perspective, 'an infinite conversation.' Hence it requires highly artificial measures to insulate rational discourses against the pressures of the lifeworld and to render them autonomous." Habermas, 2003, 253.
[30] Forst and Flynn, 2002, 66.
[31] Anderson, 2011, 96.
[32] Habermas, 1987, 174.
[33] Habermas distinguishes between the "empirical and analytical knowledge" of purposive-rational action and the "moral-practical knowledge" of value-rational action—the latter being the focus of our analysis. Habermas, 1984, 174.
[34] Habermas writes about the fallibilism of knowledge: "An expression satisfies the precondition for rationality if and insofar as it embodies fallible knowledge and therewith has a relation to the objective world (that is, a relation to the facts) and is open to objective judgment" Habermas, 1984, 9. He also writes about the fallibility of agents: "Members know that they can err, but even a consensus that *subsequently* proves to be deceptive rests to start with on uncoerced recognition of criticizable validity claims." Habermas, 1987, 150.
[35] Anderson, 2011, 98.

[36] *Ibid*, 104.
[37] To be clear, she may also be introverted and slightly agoraphobic, prone to flights of fancy, mistrusting of authority and overly obsessed with koala bears—the point is not that her particularities are positive or negative but that regardless, they somehow affect her autonomous reasoning.
[38] Anderson, 2011, 91.
[39] Cohen, 1997, 78.
[40] Meehan, 1995, 244.
[41] Habermas, 1990, 198.
[42] Benhabib, 1992, 159.
[43] *Ibid*, 38.
[44] Habermas seems to have a very specific idea of socialization and identity formation in mind, claiming that "anyone who has grown up in a reasonably functional family, who has formed his identity in relations of mutual recognition, who maintains himself in the network of reciprocal expectations and perspectives built into the pragmatics of the speech situation and communicative action, cannot fail to have acquired more intuitions." Habermas, 1993, 114.
[45] Pettit, 1996, 580.
[46] Young, 1996, 120.
[47] *Ibid*, 123.
[48] *Ibid*, 128, 131.
[49] Lafont, 2004, 49.
[50] Baynes, 2015, 121.
[51] Anderson, 2011, 102.
[52] For more on the nature of the CPI as method, please see Matthew Lipman's *Thinking in Educatio*n (Cambridge, MA: Cambridge University Press, 2003).
[53] Weber, 2008, 5.
[54] If not well facilitated or understood, a CPI can also become exclusionary by overemphasizing commonality and consensus at the expense of difference, and allowing instrumental and strategic action to parade as genuine intersubjective concern.
[55] Benhabib, 1992, 54.
[56] Bernstein, 1995, 1134. Bernstein argues against Habermas's stark division between morality and ethics because it risks compromising the dispositions required for democracy, which stem from both our moral judgments *and* our ethical convictions. His description of this ethos very closely resembles the aims and virtues that Lipman envisioned for the CPI model: "When Dewey speaks about 'debate, discussion and persuasion,' he is not simply referring to formal rules of communication, rather his major concern is with the ethos of such debate. For democratic debate, ideally, requires a willingness to listen to and evaluate the opinions of one's opponents, respecting the views of minorities, advancing arguments in good faith to support one's convictions, and having the courage to change one's mind when confronted with new evidence or better arguments. There is an ethos involved in the practice of democratic debate. If such an ethos is violated or disregarded, then debate can become hollow and meaningless." *Ibid*, 1131.

Bibliography

Anderson, Joel. "Autonomy, Agency and the Self." In Habermas: Key Concepts, edited by Barbara Fultner. Durham: Acumen, 2011.

Apel, Karl Otto. "The a priori of the Communication Community and the Foundations of Ethics." In Toward a Transformation of Philosophy. Milwaukee: Marquette University Press, 1998.

Baynes, Kenneth. "Public Reason and Personal Autonomy." In Handbook of Critical Theory, edited by David Rasmussen. Hoboken, NJ: Wiley-Blackwell, 1999.

Baynes, Kenneth. Habermas. London: Routledge, 2015.

Benhabib, Seyla. "In the Shadow of Aristotle and Hegel" and "The Debate over Women and Moral Theory Revisited." In Situating the Self: Gender, Community, and Postmodernism in Contemporary Ethics. New York: Routledge, 1992.

Bernstein, Richard. "The Retrieval of the Democratic Ethos." Cardozo Law Review. 17, 1995, 1127-1146.

Cohen, Joshua. "Deliberation and Democratic Legitimacy." Deliberative Democracy: Essays on Reason and Politics, edited by James in Bohman and William Rehg. Cambridge, MA: MIT Press, 1997.

Dworkin, Gerald. The Theory and Practice of Autonomy. Cambridge: Cambridge University Press, 1988.

Finlayson, James Gordon. Habermas: A Very Short Introduction. Oxford: Oxford University Press, 2005.
Forst, Rainer and Jeffrey Flynn. "Ethics and Morality" in The Right to Justification. New York: Columbia University Press, 2012.
Gilabert, Pablo. "A Substantivist Construal of Discourse Ethics." International Journal of Philosophical Studies, 13(3), 2005, 405-437.
Gilabert, Pablo. "The Substantive Dimension of Deliberative Practical Rationality." Philosophy and Social Criticism, 31(2), 2005, 185-210.
Habermas, Jürgen. The Theory of Communicative Action: Volume 1. Boston: Beacon Press, 1984.
Habermas, Jürgen. The Theory of Communicative Action: Volume 2. Boston: Beacon Press, 1987.
Habermas, Jürgen. Moral Consciousness and Communicative Action. Cambridge, MA: MIT Press, 1990.
Habermas, Jürgen. Justification and Application. Cambridge: MIT Press, 1993.
Habermas, Jürgen. Between Facts and Norms. Cambridge, MA: MIT Press, 1996.
Habermas, Jürgen. "Modernity: An Unfinished Project." In Habermas and the Unfinished Project of Modernity: Critical Essays on the Philosophical Discourse of Modernity, edited by Seyla Benhabib and Maurizio Passerin d'Entrèves. Cambridge: MIT Press, 1997.
Habermas, Jürgen. "A Genealogical Analysis of the Cognitive Content of Morality," In The Inclusion of the Other: Studies in Political Theory. Cambridge: MIT Press, 1998.
Habermas, Jürgen. "Rightness versus Truth: On the Sense of Normative Validity in Moral Judgments and Norms" in Truth and Justification. Cambridge, MA: MIT Press, 2003.
Habermas, Jürgen. "The Moral and the Ethical: A Reconsideration of the Issue of the Priority of the Right over the Good." In Pragmatism, Critique, Judgment, edited by Seyla Benhabib and Nancy Fraser. Cambridge: MIT Press, 2004.
Lafont, Cristina. "Procedural Justice? Implications of the Rawls-Habermas Debate for Discourse Ethics." Philosophy and Social Criticism, 29(2), 2003, 163-181.
Lafont, Cristina. "Moral Objectivity and Reasonable Agreement: Can Realism be reconciled with Kantian Constructivism?" Ratio Juris, 17(1), 2004, 27-51.
Lipman, Matthew. Thinking in Education. Cambridge, MA: Cambridge University Press, 2003.
McCarthy, Thomas A. The Critical Theory of Jürgen Habermas. Cambridge: MIT Press, 1981.
McCarthy, Thomas A. "Legitimacy and Diversity: Dialectical Reflections on Analytic Distinctions," Cardozo Law Review. 27, 2005, 1083-1126.
Meehan, Johanna, (ed). Feminists Read Habermas: Gendering the Subject of Discourse. Routledge: New York, 1995.
Pettit, Philip. "Freedom as Antipower." Ethics, 106 (3), 1996. 576-604.
Weber, Barbara. "The Practicability of the Ideal Speech Situation," Analytic Teaching. 28(1), 2008.
Young, Iris. "Communication and the Other: Beyond Deliberative Democracy." In Democracy and Difference, edited by Seyla Benhabib. Princeton: Princeton University Press, 1996.

Address Correspondences to:
Natalie M. Fletcher
Concordia University
Montréal, Canada
nataliefletcher@gmail.com