

HERLINDE PAUER-STUDER
UNIVERSITY OF VIENNA

A CONSTITUTIVE ACCOUNT OF GROUP AGENCY¹

ABSTRACT

Christian List and Philip Pettit develop an account of group agency which is based on a functional understanding of agency. They claim that understanding organizations such as commercial corporations, governments, political parties, churches, universities as group agents helps us to a better understanding of the normative status and working of those organizations. List and Pettit, however, fail to provide a unified account of group agency since they do not show how the functional side of agency and the normative side of agency are connected. My claim is that a constitutive account of agency helps us to a proper understanding of group agency since it ties the functional part of acting to the group agent's self-understanding and its commitment to specific norms, principles and values. A constitutive model of agency meets much better what List and Pettit seek to accomplish, namely conceiving of group agents as artificial persons, constituted by normative principles and entertaining normative relations to others to whom they are accountable.

1. Introduction

In their book *Group Agency* (2011) Christian List and Philip Pettit defend the logical possibility and factual existence of group agents. Group agents are in their view institutions such as commercial companies, trade unions, governments, political parties, churches, and universities – organizations with a tremendous impact on social life.

List and Pettit's strategy is first to formulate some general requirements of agency and then to argue that group agents can meet those requirements. Their starting point is a functional conception of agency according to which three conditions are necessary and jointly sufficient for agency: an agent must have representational states (beliefs), motivational states (desires or preferences), and a capacity to process the representational and motivational states in order „to act on their basis“ (List and Pettit 2011, 20).²

¹ For critical comments on an earlier version of this paper I would like to thank Martin Kusch, David Velleman and the members of the ERC-project “Distortions of Normativity” discussion group, Alexandra Couto, Julian Fink, Christoph Hanisch and Veli Mitova. Research for this paper was funded by the ERC-Advanced Research Grant “Distortions of Normativity”.

² List and Pettit describe a representation and a motivation as a pair of two things, namely an intentional attitude and an object towards which the attitude is held.

Such a functional model depicts the basic mechanical structure of agency, but it does not offer an account of the deliberative and normative level of agency.³ Agency, however, not only requires a capacity to process desires and beliefs, it also requires that an agent reflects on whether he or she has a reason to act on a desire. Agency needs scrutiny and assessment of the normative status of one's ends and desires.

List and Pettit are aware of the limits of the functional model of agency. In the course of their discussion they therefore augment their conception of agency with conditions such as a capacity for practical reasoning and critical reflection. But List and Pettit simply add those deliberative normative components to their functional model of agency. The functional and the normative explanation of agency thus exist side by side – whether and how exactly they are intertwined remains an open question.

This unresolved issue creates a tension in the author's account of group agency. According to the functional model of agency, a group agent is constituted primarily by a transformation procedure which allows the group agent to process representational and motivational states into actions. A functional perspective invites us to consider a group agent's transformation method simply as a mechanical device. The paramount standard for assessing a mechanical device is whether it works efficiently or not. But what interests us most in group agents are not their mere functional capacities, but their normative self-understanding. The normative standards and principles which are at the core of a group agent's self-understanding and in light of which its deliberations proceed and decisions are made, are crucial for our assessment of a group agent's activities. This entails that the merits and problems of a specific transformation method cannot be considered in isolation, i.e. abstracted from the normative identity of the group agent. It is a group agent's normative self-conception which settles what kind of transformation procedure is appropriate for the group agent, given the kind of organization it is.

³ The functional model of agency amounts to a more general description of what is commonly called the 'belief-desire model of agency'. According to that model, every action results from an agent's beliefs and desires. The agent deliberates from the recognition of having a desire for x, and a belief about the means by which x can be (best) realized, to action (i.e. taking the means for achieving x). The scope of the functional model in List's and Pettit's framework is broader than the belief-desire regime. A robot, as List and Pettit claim, can meet the requirements of a functional account of agency (the way the robot is programmed constitutes its transformation procedure and specifies the way in which the robot turns its 'motivational states' into movements). In this paper I argue that neither the functionalist model of agency nor the belief-desire model give us the full picture of agency. But I do not draw into question that the belief-desire regime depicts the causal mechanism of acting. However, as I see it, a functionalist account of the robot explains the movements of the robot, but the robot is not an agent in the full sense. We cannot - in a non-metaphorical sense - attribute beliefs and desires to the robot.

Starting from these considerations I argue that List and Pettit fail to provide a unified account of group agency; they do not show how the functional side of agency and the normative side of agency are connected. My claim is that a constitutive account of agency helps us to a proper understanding of group agency since it ties the functional part of acting to the group agent's self-understanding and its commitment to specific norms, principles and values. A constitutive model of agency meets, as I try to show, much better what List and Pettit seek to accomplish, namely conceiving of group agents as artificial persons, constituted by normative principles and entertaining normative relations to others to whom they are accountable.

The paper is structured in the following way. Section two outlines some central aspects of List's and Pettit's account of group agency in more detail. In section three, I depict the main idea of a constitutive account of agency by drawing mainly on the work of Christine M. Korsgaard and David Velleman. I defend a moderate version of the constitutive account of agency which makes no commitment to the more ambitious philosophical claims of those authors. Section four summarizes in which way a moderate constitutive account of agency is relevant for the issue of group agency.

2. List's and Pettit's account of group agency

List's and Pettit's overall aim is to show that group agents amount to „unified rational agents over and above their individual members“ (List and Pettit 2011, VII). In order to deal with them and assess their responsibilities, the authors maintain, we need to focus on their attitudes as a collective single agent, which might be different from the attitudes of their individual members. The attitudes of the group agent must, however, be connected with the attitudes of its members in some way. The authors depict this relationship as one of supervenience: the „attitudes and actions of a group agent supervene on the contribution of its members“ (List and Pettit 2011, 66).

List and Pettit contrast their conception with two other understandings of a group agent, namely first an „animation theory“, and secondly an „authorization theory“ (List and Pettit 2011, 8, 9). The animation theory assumes that a group agent has an ontological standing and life of its own, separate from its individual members. Characteristic of the authorization

theory of group agency is that the individual group members assign a spokesperson to whom they transfer the authority to represent and speak for them.

List and Pettit consider the animation theory as untenable because a collective entity having a life independently from individuals is dubious ontologically, let alone politically.

While certainly most of us share that skepticism, the authorization model seems initially more plausible.

Why not see group agents as institutions that are represented by a spokesperson expressing the attitudes and opinions of its group members? Does this not exactly capture what is characteristic of group agents such as firms, political parties, and universities? The CEO is speaking for the company or firm, the university president represents the university, and the spokesperson of the political party communicates the party's goals and standpoints on various issues to the public.

The authors reject the authorization theory. The model fails, they argue, because the most reasonable method of empowering a spokesperson or group representative, namely a majoritarian authorization process, faces a serious problem. Group agents cannot be created by a majoritarian process because the majoritarian aggregation of individual attitudes leads, as List and Pettit claim, to inconsistencies.⁴

Since this objection highlights a central part of List and Pettit's discussion of group agency, taking a closer look seems worthwhile. One example illustrating the problem which the authors discuss is the 'doctrinal paradox'. This paradox, first pointed out by Kornhauser and Sager, states that a court consisting of three judges cannot come to a consistent group attitude, if the group's verdict is seen as the majority aggregation of the individual attitudes of the judges. On the assumption that both premises are jointly necessary and sufficient for a verdict of liability, the majority vote supports the verdict that the defendant is liable. However, the majority vote on the conclusions supports non-liability. So, majority aggregation across the individual attitudes produces the inconsistency. (List and /Pettit 2011, 43, 44)⁵

⁴ In their assessment of majority voting as a model of aggregating individual attitudes into a group agent attitude List and Pettit assume that representational attitudes are judgments, i.e. that p is the case, and motivational attitudes are preferences, expressing a desire that p be the case.

⁵ The example by which List and Pettit illustrate the doctrinal paradox is a breach-of-contract case. A court of three judges has to come to a verdict by majority voting on the following propositions: premise 1: The defendant was contractually obliged not to do a certain action; premise 2: The defendant did that action; conclusion: The defendant is liable for a breach of contract. The legal doctrine states that obligation and action are jointly

List and Pettit generalize this result and claim that, independently from the legal case, such a paradox - they call it the *discursive dilemma* - will arise. Even when individual judgments are consistent, majority voting on interconnected propositions leads to inconsistent group judgments. They illustrate the discursive dilemma with the example of an expert panel seeking to come to a common standpoint in regard to climate change. Though the individual expert's judgments are consistent, the collective judgment generated by majority aggregation is inconsistent. (See List and Pettit 2011, 45, 46)⁶

The underlying problem is even more serious. Referring to earlier work, List and Pettit argue that a general impossibility theorem holds for the aggregation of individual attitudes. No

necessary and sufficient for the verdict that the defendant is liable, i.e. the conclusion is true if and only if the premises are true.

The inconsistency by majority voting arises if the following truth values hold for the individual attitudes of the judges:

	P1 Obligation not to do x?	P2 Action x?	C Liable for breach of contract?
Judge 1	T	T	T
Judge 2	T	F	F
Judge 3	F	T	F
Majority	T	T	F

There is a majority vote for the truth of both premises, hence the conclusion that the defendant is liable should be true as well. However, the majority vote on the individual conclusions generates the result that the defendant is not liable.

⁶ The panel consisting of three experts seeks to form a collective judgment on the following individually held propositions (the propositions p, q in List's and Pettit's example are here slightly shortened):

(p): Global carbon dioxide emissions from fossil fuels are above a certain threshold t.

(If p then q): If global carbon dioxide emissions are above threshold t, then the global temperature will increase at least by degree c over the next three decades.

(q): The global temperature will increase at least by degree c over the next three decades.

An inconsistency by majority voting arises if the following truth values hold:

	Emissions above t?	If emissions are above t, then increase of temperature?	Increase of temperature?
Individual 1	T	T	T
Individual 2	T	F	F
Individual 3	F	T	F
Majority	T	T	F

A majority of the experts agree on proposition one and two to be true, yet a majority of experts consider proposition 3 to be false. (List and Pettit 2011, 45).

The paradox can also arise in the case of preference aggregation by majority voting. (See List and Pettit 2011, 46).

aggregation function can meet four initially plausible conditions, namely universal domain, collective rationality, anonymity, and systematicity. (List and Pettit 2011, 47-50)⁷

What philosophical conclusions should we draw from that impossibility theorem? What implications does it have for the issue of group agency? It would be outlandish to read the impossibility theorem as proving that any attempt of reaching a group attitude is necessarily doomed to failure. And the authors do not interpret it that way. There are ways of avoiding the obstacle created by the impossibility theorem. The authors tackle this issue by re-thinking the formal conditions generating the inconsistency.

Relaxing systematicity provides, as they try to show, the most promising escape route (List and Pettit 2011, 54-58).⁸ The condition of systematicity has two parts: first, the group attitude on each proposition should depend only on the individuals' attitudes towards it, not on their attitudes towards other propositions (the independence part); and second, the same pattern of dependence should hold for all propositions (the neutrality condition). Relaxing systematicity, especially independence, enables the group members to decide in which way they want to align their individual attitudes and thus form a group agent's standpoint (List and Pettit 2011, 56).

Various strategies are now open for the group to take. Renouncing independence allows a group to prioritize some propositions over others so that its attitudes on one set of propositions can determine its attitudes on another set. Thus a group can choose a premise-based procedure. That means, the group tries to reach first a majority vote on each premise, and depending on that outcome the group forms its attitude on the conclusion. Here independence is given up since the group attitudes on the conclusion depend on the individual attitudes to the premises. Similarly the group might adopt a conclusion-based procedure,

⁷ Universal domain states that the aggregation function allows as input any possible individual profiles of attitudes, assuming that individual attitudes are consistent and complete. Collective rationality ensures that the aggregation function leads to consistent and complete group attitudes. Anonymity secures that all individual attitudes do have equal weight in determining the group attitudes. And systematicity means that the group attitude on each proposition depends only on the individuals' attitudes towards it, not on attitudes towards other propositions. So „the pattern of dependence between individual and collective attitudes is the same for all propositions” (List and Pettit 2011, 49).

List and Pettit's impossibility theorem which holds for attitudes (namely judgments on beliefs and preferences) is similar to Arrow's impossibility theorem; the latter shows that any method of preference aggregation which fulfills plausible axioms fails to generate a consistent ranking of collective preferences.

⁸ For a detailed discussion why giving up conditions like anonymity and universal domain seems problematic see List and Pettit (2011, 52-54). Giving up rationality is obviously no option.

meaning that the group forms a majority vote on the conclusions while not forming any attitude towards the premises.

For example, to avoid the doctrinal paradox, the court might decide whether the verdict should be the outcome of a premise-based or a conclusion-based procedure. In the first case the majority on the premises determines the verdict; in the second case the verdict results from the majority vote on the conclusions.

The group may also choose a ‚sequential priority procedure’. In that case some propositions are prior to others in the sense that they might constitute reasons for the other propositions. Another possibility is to relax anonymity as well as systematicity. The group can then accept a ‚distributed premise-based procedure’, meaning that a division of labor is implemented, whereby the different group members specialize on assessing only one of the premises, for example applying majority voting to that premise. Besides those procedures a group agent can resort to group deliberation for reaching a meta-agreement about an alignment of individual attitudes. (List and Pettit 2011, 56, 57)

Giving up systematicity entails that the individual judgments or propositions now differ in their relevance and importance. List and Pettit take that consequence as showing that the group agent must have a kind of autonomy and a capacity for critical reflection and reasoning. A rational and coherent group agent needs to have a reason why, for example, its majority voting focuses on the premises instead of the conclusions and why it chooses or renounces group deliberation. They claim that the premise-based and sequential priority-based procedures

„lead a group to perform in the manner of a reason-driven agent by deriving its attitudes on conclusions (or ‘posterior propositions’) from its attitudes on relevant premises (or ‘prior’ ones). Thus they lead the group to ‘collectivize reason’, in the sense of guaranteeing that the attitudes of the group are consistent as well as derived in a rational process“ (List and Pettit 2011, 58).

Assuming that agency includes those normative requirements amounts, however, to a substantial change of the conception of agency on which List and Pettit so far have relied. Recall that the authors started with a purely functional account of agency. But simply adding to such a model features like reflective rationality while still upholding a functional understanding of agency as the paramount mold of agency is incoherent. Reason-responsiveness and a capacity for deliberation cannot be given a purely functional

characterization: they are normative features. Accepting those normative features as conditions of agency commits one to adopt a different and more complex model of agency than merely a functional understanding of agency.

The authors argue as if the existence of a rationally reflective group agent was just a consequence of a certain transformation procedure. A transformation method, however, does not yet create a group agent. Conditions like reason-responsiveness and a capacity for deliberation must, I suggest, be included among the conditions of agency which constitute the group agent in the first place, in order that the group agent can then display those abilities in deciding on the aggregation function which seems most apt to promoting the group's goals.

Thus, List and Pettit fail to provide a unified account of agency, at least in the way their argument proceeds in the first part of their book. The upshot of their discussion of the inconsistencies arising from a majoritarian aggregation of individual attitudes is that a group agent needs a transformation procedure which does not set it on conflicting courses of acting. But in the framework of a functional understanding of agency, this suggestion amounts to asking for a working mechanical device enabling the group agent to move effectively.

The comparison with the robot which List and Pettit introduce to exemplify the functional model of agency helps. In the case of the robot the transformation method, i.e., the robot's program, is the defining element of agency. If its program fails, the robot cannot move. What's more, there is no robot. But in the case of personal agency or group agency the transformation method is not the defining element of agency. What's crucial here is that the individual agent or the group agent has to choose the transformation procedure in a way that characterizes it as an agent and not merely as a mechanical responder to its program. And my point is that no purely functional account of such a choice can be given.

There is a conflation between two issues in the discussion in the first part of List's and Pettit's book. The first issue is the plausibility of the aggregation method on which the group agent relies; the other issue concerns how to create a group agent in the first place. Now, the authors argue that a group agent cannot be generated by an aggregation method like majority voting on the individual attitudes since the method leads to inconsistencies. But a transformation procedure or an aggregation procedure alone does not define an agent. The aggregation

method is simply a means to realize the agent's ends. The agent must be already constituted to assess the suitability of a certain aggregation procedure.

Concerning agency, we can tell a causal story and a normative story. The functional model as well as the belief-desire model is apt for depicting the causal story. The functional model assumes that an agent must have representational states, motivational states, and a capacity to process the representational and motivational states into actions. The belief-desire model is an instance of the functional model: an agent has a desire (for end *e*) and a belief (by which means *e* can be achieved) and she chooses the means to achieve *e*. Standard examples are: *X* feels hungry, *X* believes that in the café round the corner he will get something to eat and therefore *X* walks around the corner to the café. Or, *Y* plans to be a famous actor, she believes the best way to realize this end is to go to the prominent actor school *Z*, and therefore *Y* applies for admission to *Z*.

The normative story goes beyond the functional model and the belief-desire model. The person is here confronted with a normative question: should she or he go to the café? Being hungry is a reason to do so – not just a cause. The same holds for the second case: the person needs to reflect whether she should apply for admission at the actor's school. The deliberation might be more complicated since additional reasons might come into play generating considerations like: would it not be better to cook and eat at home and avoid the unhealthy food in the café? Would it not be better to go to law school instead of aiming to be an actor? The deliberation might then result in forming the intention to go to the café, or, in the other example, in forming the intention to pursue an actor's career and apply for admission at the school.

Agency requires connecting those two levels. An account of agency must cover the causal and the justificatory, or normative, side of agency, and it must show how those two are internally related. Not only that: the normative and the causal story must be connected by the agent herself; that is to say, the agent's intending to do *X* or *Y* must be the source of the bodily movements constituting the realization of *X* or *Y*. The normative reasoning resulting in an intention must cause the behavioral outcome (going to the café, applying for admission at the school).

The story of agency is not that there is a body moving around the corner and independently from these movements there exists somewhere a mind in which the reasons for visiting the café are weighed. The reflections in the mind, resulting in an intention, are the source of the person's moving in the direction of the café. Similarly, the person applying at the actor's school is not simply driven by her 'program' to transform her desire into a particular movement; she applies due to her forming the intention to do so. At least this is the story if the agent comes into the picture.

And my objection is that List and Pettit do not highlight the role of the agent in the proper way, namely showing that it is the agent who establishes the connection between the causal and the normative level of agency. To avoid misunderstanding, my claim is not that we should do away with the belief-desire model. The point is that the model merely captures the functional, i.e. mechanical or causal side of agency.

Constitutive accounts of agency try to close the gap between the causal and the normative level of agency. Let us take a look.

3. Constitutive Accounts of Agency

Agency involves, as mentioned, two levels: the causal mechanism initiating certain movements and the reflective mind capable of making decisions triggering the causal sequence of movements. However, those two levels must be connected in the agent in order for the agent to act. Constitutive accounts of agency try to integrate those two levels into a unified conception of agency. According to the proponents of a constitutive account, unifying the causal level and the justificatory level, or the motivating and the normative side of agency, requires constitutive standards of agency.

A basic assumption of a constitutive model of agency is that agents are constituted by certain standards – normative standards. Being an agent requires that the standards constituting the practice of agency are in place. To get a clearer grasp of what is meant by 'constitutive standards of agency', let us take a look at the accounts Korsgaard and Velleman offer.

According to Korsgaard, constitutive standards are those "normative standards to which a thing's teleological organization gives rise" (Korsgaard 2009, 28). Korsgaard's example is the

way a house is built. A house needs walls and a roof, but also a certain form. The different parts need to be organized in a way so that the house serves its purpose, namely to provide shelter. A house can be built in a shabby way, but without a minimal structure of roof and walls we cannot even talk of a house. A ‘good house’ fulfils its function well; a ‘bad house’ serves its function poorly.

Similar conditions hold, Korsgaard claims, for agency. Her argument in short: In order to be an agent, one must be constituted and guided by the principles which make agency possible. These are the principles of practical reasoning. Without those normative standards we cannot be agents, namely someone deliberating and then deciding which incentives she or he takes to be reasons for action.

Action is defined by Korsgaard as a form of self-determination: you determine yourself to be the cause of something. In addition, she claims, there must be *a part of yourself* which is separate from the incentive and which chooses on which incentive to act. As Korsgaard notes:

“When you deliberate, when you determine your own causality, it is as if there is something over and above all of your incentives, something which is *you*, and which chooses which incentive to act on. So when you determine your own causality you must operate as a whole, as something over and above your parts, when you do so” (Korsgaard 2009, 72).

Velleman’s account of agency shares those assumptions. He notes:

“What makes us agents rather than mere subjects of behavior – in our conception of ourselves, at least, if not in reality – is our perceived capacity to interpose ourselves into the course of events in such a way that the behavioural outcome is traceable directly to us” (Velleman 2000a, 128).

Agency requires a ‘me’ in the sense that there must be someone directing the causal order of behavioural events and looking at them through “the lens of critical reflection” (Velleman 2000a, 139).

The constitutive standards of agency become apparent by reflecting on the requirements of practical reason. The basic moment of practical thought is according to Velleman a desire “to act in accordance with reasons”, and this requires acting in a way that makes what happens intelligible to the agent herself. As he notes:

“(R)ational agents have a desire to do what makes sense, or what’s intelligible to them, in the sense that they could explain it; (...) reasons for a particular action are considerations by which the action could be explained and in light of which it therefore make sense” (Velleman 2000a, 141).

The defining element of agency is thus a mental state, namely the drive for self-understanding and making sense. Making sense of one’s doings is what establishes the connection between the motivational force and the justificatory force of reasons. (Velleman 2009, 123)

Velleman draws an analogy between theoretical and practical reason: whereas belief consists in the truth-seeking acceptance of a proposition, action consists in behavior that follows considerations that make it intelligible to the agent. Action is thus behavior aimed at intelligibility, just as belief is acceptance aimed at truth. (Velleman 2009, 132, 133)

Korsgaard and Velleman pursue a more ambitious philosophical program than just giving an account of action. They claim that the constitutive standards of agency provide a foundation for normativity.⁹ The idea is that the normative standards indispensable for agency amount to universal normative standards of regulating our social relations, relations to ourselves and to others.

Korsgaard’s argument is straightforward. In order to be agents, we must be guided by the principles of practical reason, and these principles are the hypothetical imperative and the categorical imperative. (Korsgaard 2008)¹⁰ The hypothetical imperative helps us to means-end coherence; the categorical imperative is the principle of autonomy. In other terms: An agent must be an autonomous will which means to be guided by a self-given law. The principle of

⁹ Constitutive accounts of agency have been criticized in this respect by David Enoch (Enoch 2011). Enoch draws into question that Korsgaard’s and Velleman’s constitutive standards provide foundations for normativity. As I argue in this article I share that scepticism in respect to Korsgaard’s attempt to ground morality directly in agency; but I think that the objection does not undermine Velleman’s account.

In an earlier paper (Enoch 2006), Enoch challenged constitutive accounts of agency more radically by asking: Why should an agent not reject the constitutive aim of agency by simply claiming to be an agent in a different sense, namely a “shmagent”? This challenge is kind of rhetorical because Enoch does not depict the conditions holding for a “shmagent”. For Velleman’s answer to Enoch’s 2006 paper see Velleman (2009), 142-144, and 204. Enoch also attributes to Velleman the claim that his account of agency solves the internalism-externalism dispute. But Velleman does not argue for that.

¹⁰ In the 2008 afterword to that article which appeared first in 1997, Korsgaard modifies this thesis; her more recent claim is that the only normative principle of practical reason is the categorical imperative. Her argument is motivated by her downplaying the normative force of the principle of instrumental rationality, which is what Kant means by a ‘hypothetical imperative’. See also footnote 16 of this paper.

an autonomous will is henceforth a law; the feature characterizing a law is universality, and this condition of universality is exactly fulfilled by the categorical imperative.¹¹

The reason why Korsgaard attempts to justify the principles of morality via the constitutive standards of agency is that she attributes to them a special normative standing. It simply does not make sense to draw them into question. Just as one cannot ignore the constitutive standards of a house in case one is engaged in building a house, one cannot ignore the constitutive standards of agency when one is being an agent. Since self-constitution, i.e., trying to turn ourselves via our choices into unified selves, is inescapable for human beings, the constitutive standards of agency are therefore according to Korsgaard normative and “unconditionally binding” (Korsgaard 2009, 32).

The idea seems clear and striking: If it can be shown that the categorical imperative is a constitutive standard of agency, then we have a philosophically more convincing justification of this central moral principle than just claiming that the a priori truth of the categorical imperative is evident to all rational beings. This is not the place to discuss Korsgaard’s program in detail, but it seems sufficient to say that it is too ambitious. The constitutive standards of agency do not amount to fully-fledged moral principles.¹²

Velleman’s claims are more modest - and more plausible. He argues that the normative standards which are indispensable for agency set restrictions which drive us towards morality, but do not amount to moral norms. His argument proceeds as follows: The constitutive aim of action, namely trying to make sense of what we are doing, establishes the standard of intelligibility as an objective norm. This requirement is “inescapable for rational agents”, holding not only for particular actions but for a person’s way of life as a whole (Velleman 2009, 82). In order to lead a life in which we will be more intelligible to ourselves, we may have to take practical steps of social self-reform. The attempt to make sense of ourselves thus “pushes us in the direction of a moral way of life” since bringing coherence into our reasons and doings presupposes standards like universality, transparency, and mutuality (Velleman 2009, 150, 151).

¹¹ There are various places in her work in which Korsgaard develops this argument in more detail. For a succinct depiction see Korsgaard (1996, 97, 98) and Korsgaard (2007, 17, 18).

¹² The main problem is how the account copes with the problem of bad action. Autonomous agents simply need not be moral persons; they are still agents even if they act in a morally bad or evil way. But a constitutive account of agency need not necessarily be a moralizing account of agency.

To see the strength of Velleman's approach one needs to reflect for a moment on how to avoid the problems of Korsgaard's moralizing account of agency. Recall: Korsgaard identifies the constitutive standards of agency with principles of morality by endorsing a Kantian conception of autonomy. Autonomy is indispensable for agency, but the Kant-Korsgaard form of autonomy is tied to the categorical imperative. An obvious strategy to save the basic idea of a constitutive account of agency, while giving up the more presumptuous aim of anchoring morality directly in agency, would be to fall back on a non-moralizing conception of autonomy, just associating autonomy with critical reflection and having control. An agent would thus be someone having the capacity for reasoning, but there would be no specific normative direction and guidelines for the reflections of the autonomous agent.

Exactly here Velleman tries to fill the void by suggesting a non-moralizing way of defining the normative standard associated with the constitutive aim of action. The normative standard has to be spelled out more specifically, yet in a form that avoids untenable moral assumptions about agency - and claiming that the constitutive aim of action sets the standard of intelligibility is a way to do so.¹³

Though Velleman highlights a very important insight about agency, more can be said in respect to 'making sense'. We can still ask: making sense in which respect? Intelligibility and making sense to ourselves is, as Velleman himself emphasizes, tied to a way or form of life we endorse and want to lead. But the way of life we chose to lead is connected with who we are or aim to be – it is tied to our normative identity. Our normative identity is shaped by the more particular values, norms, and projects we endorse. And they provide further standards in light of which we then have to ask ourselves whether our actions and ways of acting make sense and help us to agential coherence.

Let us take stock. Korsgaard's and Velleman's claims that agency requires a deliberating 'me' guided by the constitutive standards of agency reveal important insights about agency. But

¹³ Velleman's views have undergone a certain development. In his article „The Possibility of Practical Reason“ (2000b) he held that autonomy is the constitutive aim of action, but he came to revise that position, arguing that intelligibility is the constitutive aim of action. The reason for this revision is exactly the one outlined above, namely that identifying autonomy with taking a critical stance is not sufficiently specific. The idea is that aiming for intelligibility sets restrictions on an individual's behavior. Certain ways of acting do not make sense, given the aim of an overall coherence in one's way of leading a life and in one's self-understanding.

Korsgaard's program to anchor morality directly in the constitutive standards of agency is implausible. Velleman's account seems more helpful: striving for self-understanding, i.e., intelligibility, is crucial. One needs to find and give an explanation and a justification for why one sets those steps and not others. But the standard of intelligibility can be fleshed out by tying it to the agent's normative identity.

To complete the argument, I would like to address a possible objection. One might argue that the belief-desire model (on which List and Pettit implicitly base their account of agency) amounts already to a normative model of agency by relying on the normative requirement of instrumental rationality. The argument in more detail: Desiring an end, and believing that some means is necessary to achieve the end, does not yet lead to action. One must be guided by a norm of reasoning that takes one from the belief and the desire to the conclusion (intention) to take the necessary means. The crucial norm here is, so the objection goes, the norm of instrumental rationality, namely rationality requires (if you desire e and believe that doing m is a necessary means to e , you should do m).¹⁴ Thus, one might claim, the belief-desire model already adds a normative dimension to agency.¹⁵

Instrumental rationality, however, is only a necessary condition of rationality. Someone, X , could fulfill the requirement of instrumental rationality by taking the necessary means or by giving up the end without any further reflection on whether he or she has actually a reason to take the necessary means or give up the end. We would have doubts whether X amounts to an agent in the full sense. To be an agent, we also need to assess the reasons we have for pursuing our ends in the first place and accordingly for taking the means to our ends.¹⁶

¹⁴ Note that the requirement of instrumental rationality is here formed in terms a wide-scope requirement, i.e. it can be fulfilled by either giving up the end or taking the means. This does not undermine the point I want to address, namely that according to some philosophers the principle of instrumental rationality has normative force.

¹⁵ In some of his articles Pettit seems to subscribe to that thesis. He claims, for example, that the belief-desire model requires of me to deliberate my "way to action" not by "mechanically forming and acting on beliefs and desires", but by deliberating whether a desire or preference gives me a "deliberative reason" to perform a particular action (Pettit 2007a, 40, 41). In the same article Pettit concedes that the belief-desire model does not amount to a "comprehensive depiction of agency" (Pettit 2007a, 46). Note: Pettit's argument in this article proceeds via a discussion of the rational choice paradigm. Pettit objects that the rational choice paradigm offers only a partial account of agency. Since he also holds that the goal-maximizing rational choice schema is equivalent to the belief-desire regime of folk psychology, he thereby admits that the belief-desire model is at best only a partial account of agency. See also Pettit (2002).

¹⁶ Some philosophers have argued that the normative force of the instrumental principle is merely derivative, it depends on the normative force of the ends. See, for example, Korsgaard (2008). Joseph Raz has argued that the instrumental principle is merely a "facilitative principle" enabling someone to function properly as an agent. If one is instrumentally irrational, one is, as Raz claims, "not functioning properly" (Raz 2005, 18).

4. Implications for group agency

How is all this relevant to the issue of group agency? I have discussed the basics of a constitutive account of agency with respect to individual agents. However, I think it is fairly easy to see that the constitutive conception of agency is highly relevant to the issue of group agency.

Group agents are artificial entities; they come into existence by being founded. A decision to launch a company, establish a cartel, found a university or establish a charity fund are all examples of forming a group agent in List's and Pettit's sense. But those institutions cannot be invented without defining their legal and their organizational structure. Institutions need a normative identity that underlies their practices and undertakings. And this normative identity, partly made up by the legal form of the group agent, partly made up by the organizational structure and the specific institutional aims of the group agent, yields the normative standards to assess whether the practices of a group agent allow for a coherent self-understanding.

As I argued in the last section, the specific normative identity of an agent provides more particular normative principles in light of which the basic constitutive standard of intelligibility can be applied and used for an assessment of the group agent. In the case of group agents their particular normative identity is shaped by their 'constitution' which defines what kind of a group agent they exactly are. The constitution, defining the legal form of an institution, determines whether the group agent is, for example, a commercial company, a university, a trade union, a church, a political party, a military organization or a charity organization. And these particular identities then give rise to more specific normative standards which guide the group agent in defining its decision procedures, but also in asking whether its decision procedures and undertakings make sense and are intelligible, given that the group agent is, for example, a commercial company, a cartel, a trade union, a church, a political party, and so on.

The claim in this paper has been that a constitutive account of agency helps connecting the normative and the causal level of agency. The same holds for group agency: The causal chain, we might say, is triggered by the group agent when the group agent amounts to a 'deliberating me', given its normative identity and decision procedures. In the case of group agency the

normative level of agency must be tied to the constitution of the group agent. The constitution defines the specific identity of a group agent and its legal rights and obligations. The individual members of a group agent thus have to define the decision procedures and the specific aims of the group agent in light of the group agent's constitution and the corresponding legal rights and obligations. And in the case of a well-functioning group agent there will be coherence between its legal and organizational design, its decision procedures and its ways of acting.¹⁷ What happens at the level of the actions of the group agent must make sense in light of the normative structure and commitments of the group agent. The constitution and the self-definition of a group agent provide the normative standards by which to assess the behavior of the group agent in terms of coherence and intelligibility.

The standard of intelligibility thus applies to group agents. Group agents like companies, cartels, political parties, and so on, must design their decision procedures and practices in a way that makes sense – given what they are and try to accomplish. The kind of decision procedures a group agent chooses has to cohere with its particular constitution and self-understanding. And we can easily see that a mere aggregation of the preferences of a group agent's members by majority voting might not be the appropriate decision procedure for most organizations – certainly not for group agents like governments or university departments. Those institutions need decision procedures which do not merely take into account preferences, but also allow them to come to justified outcomes which answer to the standards constituting their normative identity.

The individual group members cannot simply impose their individual aims and strategies on the group agent. They first have to look to the constitution of the group agent which gives the group agent a specific identity and defines its legal rights and obligations, and then the group members must find ways of making group decisions which are intelligible in the light of the normative identity of the group agent. Thus, the connection between the causal and the normative level is made by a 'me' who is the group agent. But this 'me' of the group agent does not require anything like the formation of a mystical 'we'; it is created by the normative elements of the constitution of the group agent, and the decision procedures and aims of the group agent have to cohere with those constitutional norms.

¹⁷ A similar claim is made by Carol Rovane when she argues that agency is tied to a practical commitment to reach overall rational unity (also in a commitment to unifying projects). See Carol Rovane, *The Bounds of Agency. An Essay in Revisionary Metaphysics*, Princeton University Press 1988, esp. 179-182.

The standard of intelligibility is tied to the normative identity of a group agent. Take the case of commercial companies. Firms are usually aimed at making profit. But there are vast disparities in the way to do so. Not all practices of accumulating profit are in line with a company's identity. If a company, for example, claims to uphold ethical standards such as respecting the workers' rights and welfare, then outsourcing their production to countries with poor standards of workers' protection and exploiting those conditions undermines its credibility, but also its coherence and intelligibility. Those practices only make sense if a firm commits itself to ruthless profit maximization. Another example where a company's practices are challenged by the standard of intelligibility is when a company sets up different branches which then start to fight each other on the market. This can quickly lead to an instability which might even endanger the existence of the firm.¹⁸

Institutions like commercial companies, political parties, churches, etc. do not exist and operate in isolation, but are part of the social world. As such they have to explain and justify their undertakings to the public, and their public standing and reputation is dependent on how well they perform in that respect. Their public credibility depends to a large extent on the coherence between their particular constitution, defining their normative identity, and their actual practices and ways of acting. And living up to the requirement of making sense is a way to achieve coherence. Intelligibility is a normative requirement driving a group agent towards a behaviour that meets publicly shared ethical standards.¹⁹

Exactly this situatedness in a shared normative space which entails accountability to others often drives group agents like commercial companies and firms towards implementing a firm-internal ethical code and announcing to others that they define their identity and self-

¹⁸ A famous case is the Chicago-based accounting firm of Arthur Andersen which in 1989 split into two separate branches, an accounting firm and a consulting firm which then harshly competed with each other. In 2000 the two companies split, and the accounting branch set up itself a consulting wing. The firm was then involved in the Enron scandal which finally ruined its reputation, but the dispute between the two branches had already weakened the company. Here we can say that the actions of the group agent simply do not make sense given what the group agent tries to accomplish, namely being an economically successful company.

The list of examples of institutional incoherences, i.e. practices which do not make sense given the constitutional commitments of the group agent, can be extended: political parties showing a remarkable mismatch between the values and ends shaping their party program and the actual behavior of their leading political representatives; churches committed to high ideals of respecting others and (sometimes implicitly tolerated) sexual abuse of minors by priests; trade union leaders helping themselves to high salaries and a lifestyle far from the life of those workers whose interests they should represent.

¹⁹ One might object that the standard of intelligibility is merely relevant to the internal self-assessment of a group agent. Crucial, however, one might claim, is to test the practices of a group agent in light of publicly shared normative standards, such as, for example, rules of ethical economic behaviour. But this criticism rests on a misunderstanding. Ethical standards have to be impartial in the sense that they do not merely hold for some particular group agents, but not for others. But this does not imply that ethical standards have to be external standards, lacking any connection to the standard of intelligibility.

understanding in terms of such an ethical code. But as soon as group agents do so, they are committed to a forceful normative standard in light of which they have to ask whether their undertakings make sense. In this way the intelligibility of the practices and actions of group agents is tied to ethical norms.

Let us come back to List's and Pettit's account of group agency. I have claimed that the authors fail to provide a unified account of group agency. Although List and Pettit make clear in the first part of their book that they are working within the framework of a functional account of agency, the authors draw attention to the normative level of group agency in the third part of their book.²⁰ Group agents need to be considered as persons in the "performative" sense, namely having the competence of performing "effectively in the space of obligations" and accordingly being able to recognize their obligations and to answer the claims of other agents (List and Pettit 2011, 173).²¹ Group agents must be able to normatively assess the options they have and must respond to recognizing their obligations "in a manner characteristic of reasoning subjects" (List and Pettit 2011, 177). As I argued, these normative elements are not integrated into the authors' account of agency.

In the third part of their book List and Pettit actually come closer to some of the ideas at the basis of a constitutive account of agency.²² They claim, for example, that agency requires self-identification. The agent must not only be aware of certain attitudes which are hers, but must be able to ascribe her attitudes to herself.

²⁰ In an overview of his work, Pettit defends a normatively reduced account of agency, namely a so-called deflationary understanding of minded agency. Accordingly, "(t)he agent will act for the realization of certain goals, according to evidentially tuned representations as to how things stand at any moment. It will be a system of intentional states like desire and belief, as it is more often put, where desires select goals, and beliefs constitute evidentially sensitive representations" (Pettit 2007b, 229). Pettit assumes that our being rational, i.e. capable of reasoning, involves higher-order or meta-propositional beliefs such as 'p' and 'if p then q' entails 'q'. But he is skeptical to include standards of rationality other than consistency and closure. (Pettit 2007b, 233-238). Compared with that account, his common book with List on *Group Agency* shows a much stronger awareness of the normative elements of agency, though the full implications for an account of agency are not yet taken into account.

²¹ List and Pettit call this the performative model of agency, connected to the conception of the legal person in the Roman law tradition where a person is defined by having legal rights and duties, for example, the right for owning and selling property, for making contracts and also for defending their interests in court.

²² In a reply to Kornhauser and Sager (2004) List and Pettit tend to a personification of groups by arguing that the discursive dilemma poses an integrity challenge for a group, namely holding inconsistent attitudes. This integrity challenge arises when a group treats its members not merely as bearers of preferences or as mere "instruments for testing the truth of some proposition", but when the group aims at being "an entity that has to answer for its judgements in the manner of a person" (List and Pettit 2005, 387). Then the group must be sensitive to inconsistencies. List and Pettit add that this personification needs self-regulation on the side of the group. My point is that this "personification" is exactly the question which norms and standards are constitutive for the group agent and its self-regulation.

But the way the authors discuss the issue of avoiding the „identification gap“ shows again that their account of agency does not bring in the agent as an active ‘me’ – neither with respect to individual agency nor with regard to group agency. Self-identification consists, the authors write, „in the formation of beliefs and desires that implicate the agent at crucial points, pinpointing the agent as the one those attitudes activate“ (List and Pettit 2011, 188). Here the agent comes in rather as the object of attitudes that operate on her, instead of being a ‘me’ that decides whether and to which extent she allows those attitudes to „activate“ her.

Self-identification needs more than merely ascribing one’s attitudes to oneself. It includes an endorsement of those attitudes. A constitutive standard like intelligibility provides a deeper form of self-identification by requiring the agent to justify the coherence of those attitudes that are traceable to her.

What about the closing of the identification gap in the case of a group agent? The gap arises, as List and Pettit see it, when the group’s attitudes and the attitudes of the individual members of the group fall apart. The authors argue that an alignment between the group’s attitudes and the individual attitudes in the form of a mere coincidence of interests is not sufficient to close the identification gap. One way to close the gap is when the individual members adopt the group’s viewpoint so that „their individual attitudes are under the automatic guidance of the group” (List and Pettit, 192). This amounts according to List and Pettit to individual attitudes of the group members „in whose propositional expression the group figures as *we*” (List and Pettit 2011, 192). Note: the authors do not discard their earlier skepticism about an animation theory of group agents, arguing that this expression of the ‘we’ in the ‘me’ only holds if there are no “red lights“ that stop individuals to align so deeply with the group. Nevertheless the authors depict the identification with the group agent as a state in which „(m)etaphorically speaking, the group’s mind is then instantiated in each of us and leads us to act automatically, as required by the group“ (List and Pettit 2011, 192).

The constitutive standard of intelligibility helps here to draw the line. When the group agent acts in a way that does not make sense, given its normative identity and its ends, then the individual members have to react with criticism. But the choice agents have to make should not be between identification with the group agent in the form of acting “automatically, as required by the group“, on the one hand, and keeping away from the group altogether, on the

other hand. We can still identify with a group even when we have reason to criticize its ways of acting. Social life often demands and entails that we are members of group agents – but in doing so, we have to ask ourselves whether this makes sense, from our individual perspective and from the perspective of us as part of the group agent.

Literature

Enoch, David (2006) “Agency, Shmagency: Why Normativity won’t Come from what is Constitutive of Action”, in: *Philosophical Review* 115, 169-198.

Enoch, David (2011) “Shmagency Revisited”, in: Michael Brady (ed.): *New Waves in Metaethics*, Basingstoke: Palgrave Macmillan, 208-233.

Kornhauser, Lewis A. and Lawrence G. Sager (2004) “The Man as One: Integrity and Group Choice in Paradoxical Cases”, in: *Philosophy and Public Affairs* 32, 3, 2004, 249-276.

Korsgaard, Christine M. (1996) *The Sources of Normativity*, Cambridge. Cambridge University Press.

Korsgaard, Christine M. (2007) “Autonomy and the Second Person Within: A Commentary on Stephen Darwall’s *The Second-Person Standpoint*”, in: *Ethics* 118, 1, 8-23.

Korsgaard, Christine M. (2008) “The Normativity of Instrumental Reason”, in: Korsgaard: *The Constitution of Agency. Essays on Practical Reason and Moral Psychology*, Oxford: Oxford University Press, 27-68.

Korsgaard, Christine M. (2009) *Self-Constitution. Agency, Identity, and Integrity*, Oxford. Oxford University Press.

List, Christian and Philip Pettit (2005) “On the Many as One: A Reply to Kornhauser and Sager”, in: *Philosophy and Public Affairs* 33, 4, 2005, 377-390.

List, Christian and Philip Pettit (2011) *Group Agency. The Possibility, Design, and Status of Corporate Agents*, Oxford: Oxford University Press.

Pettit, Philip (2002) “Decision Theory and Folk Psychology”, in: Pettit: *Rules, Reasons, and Norms. Selected Essays*, Oxford: Clarendon Press, 192-221.

Pettit, Philip (2007a) “Construing Sen on Commitment”, in: Fabienne Peter and Hans Bernhard Schmid (eds.), *Rationality and Commitment*, Oxford: Oxford University Press, 28-48.

Pettit, Philip (2007b) “Joining the Dots”, in: Geoffrey Brennan, Robert Goodin, Frank Jackson, and Michael Smith (eds.), *Common Minds. Themes from the Philosophy of Philip Pettit*, Oxford: Clarendon Press, 215-344.

Raz, Joseph (2005) "The Myth of Instrumental Rationality", in: *The Journal of Ethics and Social Philosophy* 1, 1, 2-28.

Velleman, David J. (2000a) "What happens when someone acts?" in: Velleman: *The Possibility of Practical Reason*, Oxford: Oxford University Press, 123-143.

Velleman, David J. (2000b) "The Possibility of Practical Reason", in: Velleman: *The Possibility of Practical Reason*, Oxford: Oxford University Press, 170-199.

Velleman, David J. (2009) *How We Get Along*, Cambridge: Cambridge University Press.