

Collective Virtues

Citation for published version (APA):

Astola, M. (2022). *Collective Virtues: A Response to Mandevillian Morality*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Engineering and Innovation Sciences]. Eindhoven University of Technology.

Document status and date:

Published: 26/04/2022

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.



Collective Virtues
A Response to Mandevillian Morality

Mandi Astola

Collective Virtues

A Response to Mandevillian Morality

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van
de rector magnificus prof.dr.ir. F.P.T. Baaijens,
voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen
op dinsdag 26 april 2022 om 16:00 uur

door

Mandi Astola

geboren te Helsinki, Finland

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

Voorzitter:	prof.dr. I.E.J. Heynderickx
Promotor:	prof.mr.dr.ir. L.M.M. Royakkers.
Copromotoren:	dr. G.J.T. Bombaerts. dr. A. Spahn.
Promotiecommissieleden:	prof.mr.dr.ir. L.M.M. Royakkers dr. A. Spahn dr. G.J.T. Bombaerts prof.dr.ir. I.R. Van De Poel (Technische Universiteit Delft, Technology Policy and Management) prof.dr.ir. C.C.M. Hummels prof.dr. S. Vallor (University of Edinburgh, Edinburgh Futures Institute)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

A catalogue record is available from the Eindhoven University of Technology Library.

ISBN: 978-90-386-5499-7

The cover picture was drawn by the author.

Research for this thesis was made possible by SCALINGS, a Horizon 2020 project under grant agreement number 788359.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 788359.



Table of Contents

Acknowledgements	ix
Preface: A Musical Analogy	xiii
1 Introduction	1
1.1 Virtue Ethics and the Challenge of Emergent Phenomena in Group Dynamics	2
1.2 Collective Virtues	3
1.3 Research Questions	4
1.4 How I Answer the Research Questions	6
1.5 Publications on Which this Work is Based	8
2 Mandevillian Virtues	9
2.1 Introduction	9
2.2 What Is Mandevillian Morality?	12
2.3 Pure Virtue-Ethical Theories	16
2.4 Why Mandevillian Morality Challenges Pure Virtue-Ethical Theories	18
2.5 Three Features of a Virtue Theory that Can Account for Mandevillian Morality	19
2.6 Conclusion	26
3 Can Creativity be a Collective Virtue?	29
3.1 Introduction	29
3.2 When Does One Possess the Virtue of Creativity?	32
3.3 When Does an Innovator Possess the Virtue of Creativity?	36
3.4 Can Collectives Possess the Virtue of Creativity in Innovation?	41
3.5 Components of Creativity as a Virtue	48
3.6 Implications of Creativity as a Collective Virtue	49
4 Can Responsibility be a Collective Virtue?	51
4.1 Introduction	51
4.2 Collective Moral Responsibility	54
4.3 The Problem of Many Hands	55
4.4 Collective Responsibility as a Concept Must be Revised	57
4.5 Collective Responsibility as a Virtue	58
4.6 Concluding Remarks	65
5 The Relationship Between the Individual and the Collective	67
5.1 Introduction	67
5.2 Energy Justice and Renewable Energy Communities (RECs)	69

5.3	Selection, Method and Description of Case Studies	71
5.4	Findings: Risk, Effort and Power as Core Concerns within RECs	75
5.5	Conclusion	80
6	Do Collective Virtues have to be Scaled-Up Individual Virtues?	85
6.1	The Argument in Brief	85
6.2	Collective Virtues in Current Literature	86
6.3	These Arguments only Make Sense if the Goodness of Virtues is Grounded in the Goodness of Certain Psychological Dispositions	89
6.4	Collective Virtues can also be Grounded in Eudaimonia or Admirability	92
6.5	Collective Virtues may not have to be Structurally Equivalent to Individual Virtues if they are Grounded in Desirability or Admirability	92
6.6	A New Direction for Collectivist Virtue Ethics and Epistemology?	94
7	Conclusion	95
7.1	Relevance for Pure Virtue Theorists	97
7.2	Relevance for Ethics of Collective Action	97
7.3	Further Research	98
	References	101
	Summary	113
	Samenvatting	117
	About the Author	121
	List of Publications	123
	Academic Publications	123
	Professional and Popular Publications	123
	Simon Stevin (1548-1620)	127

Thus every part was full of vice,
yet the whole mass a paradise;
flattered in peace, and feared in wars,
they were the esteem of foreigners,
and lavish of their wealth and lives,
the balance of all other hives.
such were the blessings of that state;
their crimes conspired to make them great:
and virtue, who from politics
had learned a thousand cunning tricks,
was by their happy influence
made friends with vice: and ever since,
the worst of all the multitude
did something for the common good.

Bernard Mandeville, 1732, The Fable of the Bees: or Private Vices, Public Benefits

Acknowledgements

As you might gather from reading this dissertation, I am a *collectivist*. What I mean by this is that I think that collectives can *do things* and *be responsible* for the things they do, collectively. Naturally, as I write these acknowledgements I wonder whether I can really say that *I* wrote this dissertation. All the help, guidance and encouragement that I received has been essential for its completion. And perhaps the credit for this work should be in some way attributed to all of us collectively.

My closest collaborators have been my amazing supervisors, Gunter Bombaerts, Lambèr Royakkers and Andreas Spahn. Throughout these last three and a half years, I have looked forward to every single meeting we had. I have spoken to enough PhD students to know that that is quite special. I really appreciate the positive vibes at our meetings, the encouragement you gave me when I needed it, and also your honest criticism when I needed a healthy metaphorical slap in the face.

My dissertation turned out very different than I expected in the beginning. I began this project as a part of a Horizon 2020 project, SCALINGS. The aim of our project was to study co-creation and my part was to reflect on co-creation philosophically. I followed various co-creation projects, taking interviews and attending meetings. It was in my discussions with co-creation practitioners that I realized how morally disvaluable traits can sometimes be productive in a group context. I also saw collective creativity, collective responsibility and irresponsibility in action. This intrigued me. I first wrote a paper on mandevillian virtues in co-creation. But I soon realized that this did not apply to co-creation only. What I had seen in these projects was something that happens in all kinds of collective human affairs. This is how my thesis became more broad and theoretical than it was supposed to be in the beginning. Watching co-creation in practice opened my eyes to the wonders of collective agency.

For that reason, I am very thankful to everyone from my case study projects Brainport Smart District, Living Lab Eckart-Vaartbroek and Nautilus in Amsterdam, who welcomed me and my research. I am grateful for all the interviewees who took the time to speak with me. I am especially grateful to Wina Smeenk who welcomed me warmly to Brainport Smart District and from whom I've learned a lot about co-creation and co-design. I also thank everyone involved with the Urban Innovation Centre Think-Tank at TU/e. Discussions with all of you have made me the socially engaged and interdisciplinary scholar I am now.

I am also grateful for the three great years of collaboration that I enjoyed with all the colleagues from the SCALINGS project. I thank my economist colleagues Bożena Ryszawska, Magdalena Rozwadowska and Piotr Szymański for welcoming me so warmly in Wrocław in 2019. I also thank them as well as Sophie Nyborg, Meiken Hansen, Anja Ruess, Erik Laes and Gunter Bombaerts for the lovely research collaboration on energy justice, which made its way to a chapter in this dissertation. I am also very thankful for the countless scholarly and less scholarly discussions with Shelly Tsui and Erik Laes. Shelly, I have learned a lot from your structured and methodical approach to research that is very much the opposite of my natural style. Your dissertation will be awesome!

During the last years I saw the growth of an active and supportive 4TU Ethics PhD community. I want to thank all of you for all the fun, learning and support. I especially want to thank Anna and Leon, I somehow feel like you started all this. I also thank my peers at the OZSW PhD council for the wonderful writing retreats with great company.

I thank everyone at TU/e's Philosophy & Ethics department for providing a supportive academic environment for me to grow and flourish as a scholar. Many staff members have read my work and commented on it on multiple occasions. This I am very grateful for. I am also very thankful to Sven Nyholm for helping me and many other PhD students with writing, and also for reading my work and providing helpful comments.

I did not become a PhD student out of nowhere. My parents Laura and Aki are well acquainted with the PhD journey and have been very supportive throughout mine. They also raised me to become an environmentalist, a feminist and, most of all, a critical thinker. For this I am forever indebted to them.

To be a good philosopher, one must also live a little. Otherwise, how should one know what one is talking about when talking about life? I thank all my friends for allowing me to live a little. And I am not going to name everyone. Thank you for the love, the support, the moestuin, girls nights in Rotterdam or elsewhere, Villa Fissa, NachtiVillaFissa, Kraakpand, Bosfissa, Spaceland and all the other things we do. Thank you for the pep-talks. Thank you for the holiday in Porto and many more to come. Thank you for your encouragement, lending me your cardigan to wear to the job interview for this PhD position, as well as caring for Vriloouz when I was away on conferences. I love you all and you are all so important to me. Also thank you to The Funky Cat and everyone involved. It feels good to give something back to the Dutch

dance scene that raised me. It's quite silly that my philosophy career so often gets in the way of doing that.

Patrick, I have learned so much from you and you helped to keep me sane during this process. You have influenced my thought more than you know. I hope I have not been a bad influence on you. Also, thank you for encouraging me and, in so many ways, showing me the ropes of academic life.

Some people need 10 years of training to become a real philosopher, others just are. One philosopher in the latter category has been hugely influential on my thinking and my life. You and I share a taste for the finer things in life, as well as the highest level of integrity ever. We also share stamina. And of course we do, it's from running the hedonic treadmill of thought, defining useful and useless concepts, hour after hour. My beloved Maxim, you are simply brilliant and insane. And I am infinitely proud and grateful that you take an interest in philosophy. Don't underestimate how much of my philosophy has been influenced by you.

And finally, thank you Robert for the years of continuous support emotionally and intellectually. You are one of the reasons why I am fond of exemplarist moral theory. Because you are an actual exemplar, morally and intellectually. I admire you and I hope that I can one day be as excellent as you are. According to exemplarist moral theory, being around you should help me achieve that. If the only thing to come out of this dissertation would be that it made me sound interesting enough for you to come to Eindhoven to go on a date with me, I would do it all again a million times.

I also thank Vriloúz, a.k.a "Loaf," for not contributing anything.

Preface: A Musical Analogy

A single note played for a second is not music. Music requires rhythm and repetition and must be heard over a certain period of time. Similarly, a single action is not virtue, a virtue is developed and practiced over time, it involves repetition of good actions. Many such parallels can be drawn between music and virtue ethics. Virtuous and vicious traits are often recognized instinctively, just as quickly as one hears harmony or dissonance. Moreover, the concept of harmony is connected to both musical theory and virtue theory.

According to legend, Pythagoras discovered the principles of musical harmony at a blacksmith's workshop. He heard five hammers beating on iron and noticed that one of them was discordant. He rejected the discordant hammer and examined the four harmonious sounding ones. He found that their weights were proportional (Steadman 1964). These proportions that determined musical harmony were the key to understanding harmony in the universe and therefore also that which is *good*. This aligns with the way the Greeks often saw the human being as a *microcosm* which reflected the universe, the *macrocosm* (Wayman 1982). Perfection, according to Pythagoras, can be found in harmonizing with the music of the universe.

In ancient Greece, musical education was considered a part of moral education. Plato believed that mothers should sing to and dance with their children when they are young. Children must learn through practice and imitation what it means to be harmonious (Stamou 2002). This was considered important for the refinement of the soul, which would also allow for the development of moral virtue (Woerther 2008). Moral character, like music, is also learnt by imitation and practice, just as a lyre or flute player practices and imitates better players to learn (Zagzebski 2017).

While the ancient Greeks had flutes and lyres, they did not have synthesizers. Why is this important? Because synthesizers teach us about the beauty of distortion.

Synthesizers use an electrical signal to control movement in a speaker, which produces sound. The most basic wave is a sine wave, with a characteristically neutral sound. Often a musician will begin with the geometrical perfection of a sine wave, and work towards distorting it, to give it character. British musician and composer Brian Eno is quoted:

“The least interesting sound in the universe is probably the perfect sine wave. It’s the sound of nothing happening. It’s the sound of perfection, and it is boring. As David Byrne said in his song, ‘Heaven is a place where nothing ever happens.’ For me, perfection

represents the absence of character. If I get anything even close to a perfect sine wave from my synthesizer, I send it to an amplifier so that it starts to get imperfect and becomes distorted. Distortion is character, basically. In fact, everything we call 'character' is the deviation from perfection. So perfection to me is characterlessness."

Brian Eno speaks of character in sound, but the quote implies that he is not speaking only about sound. Heaven is a place where nothing ever happens. Therefore, imperfection can make things interesting. Imperfection can add character to a person. This is why the shenanigans of selfish and wrathful people pique our interest in magazines and movies and why we fall in love with badboys. I believe that Eno's point is not philosophically trivial. It points to a widespread intuition that is very much at odds with much of moral theory: Why is it that everyone should strive to be perfectly good, when everyone being good would make the world a boring place? Should the music of life be played with perfect sine waves only?

My dissertation provides a partial answer to this question, and this answer asks us to shift our focus from the distorted sound to the entire composition. Distortion and dissonance look different when we perceive them as a part of a whole. Pure distortion is just noise, but distortion on top of a sine wave gives the sound character. Dissonant chords can evoke emotions that harmonic chords cannot. A drunk guitar player by themselves can sound and look pathetic, but they can sound and look great in a band that accommodates for that. Similarly, individual misbehavior or vice becomes something different when we see it as a part of a whole. Individually annoying people can be wonderful group members or even group-improvers. An aggressive and obsessive visionary can push a company to achieve great success and a lying politician can make a government more productive, even though it is clear that what they are doing is wrong.

Distortion and dissonance are not necessarily opposed to harmony. In fact, distortion itself is, in the end, harmonic. When a sine wave is distorted, various harmonic frequencies are added to it, giving the characteristic richness to the sound. Dissonance too is not regarded by music theorists as opposed to harmony. Discordant harmonies are a part of the musical canon of many genres, notoriously jazz. Harmony denotes a quality which seems to transcend almost anything specific. Harmony can occur on many levels and in many ways. And music evolves when we learn to appreciate new kinds of harmony. As the playwright George Bernard Shaw is quoted "The technical history of modern harmony is a history of growth of toleration by the human ear of chords that at first sounded discordant and senseless to the main body of contemporary professional musicians." Music should be a beautiful composition of

sounds. In our world today, “ugly” sounds, like dissonance and distortion, also have their place in beautiful music.

Moral theory and morality itself, just like music theory and music itself, is always evolving to include and exclude kinds of harmony as legitimate. Many fields in moral philosophy have embraced disharmony. Many feminists, for example, have emphasized the value of discomfort, or making others feel uncomfortable, as resistance against oppressive norms, and as a vehicle for changing those norms (Ahmed 2017; Chadwick 2021). In democratic theory there has also been a shift towards agonistic democracy where disagreement and tensions are kept salient, rather than set aside by consensus (Maxwell et al. 2019; Mouffe 1999).

Human life is lived in communities, and perhaps the community, not the individual, constitutes the music of life. Vice has its place in moral harmony. We can understand that place if we take collective moral behavior as seriously as we take individual moral behavior. Rather than discarding the discordant hammer as Pythagoras does, let us think more like Brian Eno and value the character that it adds. Heaven is a place where nothing ever happens, and the music of life is played together by the virtuous and vicious alike.

1 Introduction

Virtue ethics has been guiding moral practice since ancient times (Aristotle 2015). Even today, virtue ethics continues to give helpful answers to many moral questions in contemporary life (Anscombe 1958; Foot 1978; Macintyre 1986). Various scholars have applied virtue ethics to contemporary challenges like questions around emerging technologies or environmental degradation (R. Sandler 2009; R. L. Sandler 2013; Vallor 2016). Vallor, for instance, recommends the development of “technomoral virtues” which involves, for example, cultivating attention and attentiveness, in response to the constant temptation of phone screens (Vallor 2016). Sandler’s environmental virtue ethics has focused on the importance of cultivating character traits that are environmentally justified, environmentally responsive, and environmentally productive (R. L. Sandler 2013). The work of such scholars goes to show that many of the challenges in our lives today can be fruitfully conceptualized with traditional virtue ethics.

Virtue ethics is a future-looking and development-focused moral theory. For these reasons it is seen as a very practical theory which excels at giving people concrete guidance in life (R. L. Walker and Ivanhoe 2006). For instance, when faced with a question like “How should I deal with my climate anxiety?” a virtue ethicist might answer “fear provides an opportunity for practicing courage. When faced with climate anxiety, try to develop climate-courage.” The virtue ethicist might further specify what “climate-courage” means (i.e. not avoiding negative climate news, but trying to see it as encouragement for action) and what interventions one can take to develop this virtue. The guidance of virtue ethics also focuses on long-term character development, rather than individual actions. Since many of the problems humanity now faces, like climate change, require long-term change rather than one-off decisions, one could argue that virtue ethics is particularly suitable as a guiding moral theory for our times.

However, the strength of virtue ethics, that it focuses on character development, may also be its caveat. The focus on character makes it difficult to apply the theory to collective behavior, policy and global problems. If the virtue ethicist is asked “How should a multinational corporation take responsibility for the global power it holds?” or “Should my company focus more on improving inclusivity or sustainability?” these

questions might be difficult to answer with virtue ethics. At most, virtue ethics can give guidance to individual actors, but not to collectives.

Many of the global problems humanity faces require collective behavior change as a response. Climate change demands changes in everyone's lifestyle. Problems like mass surveillance via internet also require many complex changes in systems, business models, governance and culture. It is likely that virtue ethics, which has traditionally been limited to the small and interpersonal, may face difficulties. While traditional virtue ethics may be excellent in guiding individuals in how to respond to these problems individually, it struggles to prescribe collective behaviors.

One way in which virtue ethics might be able to give guidance for collective behavior is if it could make judgements about collective agents. This dissertation is an exploration into that possibility. I propose, based on earlier work, that collective agents can, much like individuals, have character traits, which can be inherently good or bad. This dissertation justifies and refines a theory of collective virtues and vices.

1.1 Virtue Ethics and the Challenge of Emergent Phenomena in Group Dynamics

One reason to build a collective virtue theory is that it would be useful. Above I have described the uses of such a theory, like its applicability to collective behavior change. Aside from the potential usefulness of collective virtue theory, there is another problem it responds to. That problem is the difficulty of individualist virtue ethics in accounting for moral goodness and badness which emerges at the collective level.

One type of virtue ethics specifically struggles with goodness and badness which emerges at the collective level. This is *pure virtue theory*. Pure virtue theory, held by some philosophers notably Anscombe, is the commitment that virtues and vices are the fundamental basis of all morality (Anscombe 1958).¹ A pure virtue theorist generally believes that only a person's character can be bad or good. Things like actions, events, rules or policies can only be morally good or bad in the way that they relate to good or bad individual character. Therefore, pure virtue theorists believe that all goodness and badness that happen on the collective level, like bad company policy, a hostile

¹ In contrast to pure virtue theory, some philosophers see virtue theory as a helpful addition to a more fundamental consequentialist or deontological moral theory. Nussbaum's view is an example of how virtue ethics can be viewed as a useful add on, rather than a third category of moral theories next to consequentialism and deontology (Nussbaum 1999). She argues that virtue ethics focuses on personal development and is compatible with deontological and consequentialist moral theories.

culture or lack of transparency is bad because it results from the vicious or at least non-virtuous behavior of group members.

However, occasionally non-virtuous or even vicious behavior can play a positive role in groups. The desirability of some forms of vice can become evident in a social context. If a cowardly person and a foolhardy person make decisions together, they may well make courageous decisions. Because the two are each other's polar opposites, they balance each other out and act, as a dyad, much like a virtuous person would act. Flaws in moral character can also sometimes be epistemically valuable in a group context. Nosiness and epistemic self-indulgence can lead a person to deviate from commonly accepted epistemic norms. People sometimes seek out trivial information, in inappropriate or self-deceptive ways like prying, gossiping or seeking facts to support harmful and self-serving ideas. Such behavior can however enrich epistemic communities, for instance because it broadens the spectrum of information which the community is exposed to (Morton 2013). At times, epistemic failings can also enrich epistemic communities. Studies in swarm intelligence have shown that suboptimal cognitive behavior of some individuals sometimes helps groups in some epistemic tasks (Smart 2018a, 2018b). Sometimes moral failings can also improve moral communities in a systematic way. For instance, a wrathful or intemperate person can generate more valuable dialogue or punish norm transgressions of others, resulting in more virtuous collective behavior.

Pure virtue theory is premised on the idea that all moral good comes from virtue in one way or another. Therefore, vice playing a constitutive role in morally good group behavior is problematic for pure virtue theory. However, if pure virtue theory were to recognize that groups can also have vices and virtues, then pure virtue theory would be able to account for good group behavior.

1.2 Collective Virtues

We often attribute virtues and vices to collectives in colloquial speech. We say things like "Shell is an irresponsible company" and "Pixar is very creative." We also have legal constructs that protect legal personhood of corporations. While this kind of talk is common colloquially, moral theories about collective virtues and vices are still relatively new and underdeveloped. Collective virtue theories are moral theories that attribute virtues and vices to groups. Most of the work on collective virtues thus far has been conducted in social epistemology and concerns collective epistemic virtues. Social epistemology is a relatively new branch of epistemology which takes as a

premise that much of human epistemic activity occurs in groups. We often learn about our environment by hearing about it from others. Scientists often conduct research in groups. Group cognition, collective memory and co-creative ideation are all examples of groups undertaking epistemic activities which are hard to explain without reference to group-level phenomena. Some work has been done on what collective epistemic virtues should be in general and how they are to be grounded epistemically (Axtell and Carter 2008; Carter 2016; M. Gilbert 2004; Lahroodi 2007; D. Tollefsen 2004).

At the same time, moral philosophy has seen much work on collective agency, collective responsibility and collective action (Bratman 2014; M. P. Gilbert 2002; i.e. List and Pettit 2011; Smiley 2008). However, very little of this work has been integrated with virtue ethics. Virtue epistemology has begun to take collectives seriously, but virtue ethics is not there yet. How does collective knowledge, action and responsibility relate to virtue? Does a company need to consist of virtuous individuals to be responsible?

The purpose of this dissertation is to defend, refine and further the claims of collective virtue theory. Collective virtue theory can deal with vicious behavior contributing to good. This is because collective virtue theory makes it possible to distinguish between the individual vice and a collective virtue, to which the vice contributes. Collective virtue theory is also promising because the greatest challenges facing humanity at this moment cannot be remedied through individual action. Climate change, digital mass surveillance and political polarization are caused by collectives of people, and must be tackled with collective action. Collective virtue theory can prescribe states of virtuous operation for these collectives and therefore be useful as a guide the collective action that we need to ensure our flourishing as a species.

1.3 Research Questions

Collective behavior and collective life are such an important part of what makes us flourish. However, virtue ethics judges people individually only. If virtue ethics cannot adequately deal with collective behaviors, then its prospects as a moral theory may be limited. Our moral agency as individuals in a globalized society is bound up with these collective global problems. If virtue ethics is limited in prescribing action for individuals, then it might be *too limited*.

At the same time, virtue ethics has many advantages over other moral theories. It is future-looking, pragmatic and tends to provide intuitively convincing guidance.

Therefore, it should be explored in what way virtue ethics can deal with collective responsibility and collective action. The goal of this dissertation is to do precisely that, explore how virtue ethics can guide collective action and the peculiar dynamics of collectives in which vicious behavior sometimes has a positive impact. To contribute to this goal, I answer three research questions.

The first research question is aimed at specifying the philosophical problem that virtue ethics has regarding its ability to guide collective behavior. The first research question is:

1. *Do irreducibly collective behaviors of groups create problems for pure virtue ethics as a moral theory?*

Sometimes the behavior of collectives is described as “emergent.” This means that it cannot be understood, or explained only by reference to individual members of the collective. Swarm intelligence and various types of collective action are examples of this. Virtue ethics, however, is traditionally focused on the character of the individual. If the behavior of a collective is emergent, or supervenes on the behavior of individuals, then this means that it is not directly a candidate for moral evaluation within virtue ethics, given that virtue ethics locates goodness and badness in the character of individuals. Answering this requires knowing whether virtue ethics can be expanded to account for the value of collective phenomena. This yields a second research question:

2. *How can virtue ethics respond to collectively valuable phenomena which are not reducible to individual virtue?*

It may be possible for virtue ethics to respond to the challenge of collective phenomena so that it can better prescribe good conduct for collective action. This might require an expansion of traditional virtue theory. If so, then we must investigate whether this expansion retains the merits of traditional virtue theory and that it can guide action. A moral theory should be insightful, workable and broadly applicable. For this reason, the last research question is focused on the practical applicability of the adapted version of virtue ethics. Such a theory should give guidance in cases of collective and individual action. Therefore, the third research question becomes:

3. *Does collective virtue theory produce valuable insights when applied to real-life collective action cases?*

If we adapt virtue theory to make it more collective action-friendly, then it should still be workable and yield insights in practice. While moral theories are prescriptive, they also should have a descriptive component, in the sense that they should align with at

least some intuitions and practices around moral action. The last research question will evaluate this practical insightfulness and workability.

1.4 How I Answer the Research Questions

I answer the three abovementioned research questions by putting forward a collectivist virtue theory which accommodates the fact that vices are sometimes valuable. The theory takes groups as a unit of analysis. I show how a focus on collectives can explain much of human activity, from creative innovation to collective responsibility, better than individual-focused theories.

Chapter 2 begins by answering the first research question: *Does the moral value of irreducible group level phenomena create problems for, or limit the scope of virtue ethics as a moral theory?* I argue that the value of irreducible group level phenomena indeed presents a problem for pure virtue theory. This challenge is that sometimes vicious traits can be a systematic part of something inherently valuable. I call this the challenge of *mandevillian morality* for pure virtue theory.

Is this challenge detrimental to pure virtue theory? This depends on whether pure virtue theory can formulate a response to it. This brings us to the next question: *How can virtue ethics respond to collectively valuable phenomena which are not reducible to individual virtue?*

To answer this question, I formulate a possible pure virtue theoretical response in the latter part of chapter 2. I put forward a type of virtue theory with three features that allow it to account for the value of mandevillian morality. This theory has three distinctive features: a distinction between individual and group virtues, a distinction between motivational and teleological virtues, and an acknowledgement of the normativity of “vicious” roles in groups. (Astola 2021a).

Having established collective virtue theory as one of the possible responses of pure virtue theory to mandevillian morality, the subsequent chapters focus on showing that collective virtue theory can yield valuable insights when applied to real life cases, answering the third research question *Does collective virtue theory produce valuable insights when applied to real-life cases?* My approach to answering this question is to illustrate that collective action cases can meaningfully be analyzed using the concept of collective virtues. I do this for two collective action cases.

Chapter 3 focuses on the collective action context of innovating organizations and the virtue of collective creativity. I explore how the virtue of creativity often takes collective forms and show that all elements of a broadly Aristotelian definition of

creativity as a virtue can be found on a supervening level in creative collectives (Astola et al. 2021).

Creativity only produces morally justifiable behavior if it is coupled with responsibility. In chapter 4, I show that responsibility can also be a collective virtue. Furthermore, seeing collective responsibility as a virtue solves problems in the notion of collective responsibility. This is because seeing collective responsibility as a virtue can deal better with attributing responsibility in cases of “problems of many hands,” which are morally problematic situations where a collective can be held responsible for something but no individual in the group can. This chapter answers the third research question, *Does collective virtue theory produce valuable insights when applied to real-life cases?* by illustrating the usefulness of collective virtues in another collective action case, namely problems of many hands.

Having illustrated how collective virtues can make sense of collective action cases, I then turn to some limitations of collective virtue theory in this area. Chapter 5 explains how collective virtue functions as an ideal, but how some minimal conditions often need to be met in order to make group membership just for the members of the group. I use the example of volunteer run energy communities to show that in order for a group to function justly, all members must experience a balance of effort, risk and power. This chapter adds some nuance to my affirmative answer to the research question *Does collective virtue theory produce valuable insights when applied to real-life cases?* Collective virtue theory therefore can provide useful insight in many contexts, but we need more individual-focused reflection aside it, as is clear from the case of community energy projects.

After presenting a possible response for pure virtue theory, to the challenge of group-level phenomena, I reflect on whether there are more ways to argue for the response I have presented. Chapter 6 is a reflection on the current literature on collective virtues, including the claims presented in this dissertation, which tends to see collective virtues as entities that resemble individual virtues. It is often assumed that collective virtues require collective mental states, like individual virtues. I call this assumption into question and propose that with an exemplarist grounding, collective virtues need not rely on the existence of collective mental states (Astola 2021b).

In chapter 7, the concluding chapter, I summarize my findings and how they answer the research questions. I also present two directions for future research that are worth investigating: the possibility of an exemplarist theory of collective virtues and the application of collective virtue ethics to various societal collective action problems.

1.5 Publications on Which this Work is Based

The chapters of this dissertation are based on a number of publications, some of them co-authored.

Chapter 2 is based on the published article: Astola, M. Mandevillian Virtues. (2021) *Ethical Theory and Moral Practice* 24, 19–32. <https://doi.org/10.1007/s10677-020-10141-9>

Chapter 3 is based on the published article: Astola, M., Bombaerts, G., Spahn, A. Royakkers, L. (2021) Can Creativity Be a Collective Virtue? Insights for the Ethics of Innovation. *Journal of Business Ethics*. <https://doi.org/10.1007/s10551-021-04833-0>

Chapter 4 is based on the article: Astola, M. Collective Responsibility Should be Seen as a Virtue. Accepted for publication in an edited volume titled *Virtues and Values in a Changing World*. Cambridge University Press.

Chapter 5 is based on the article: Astola, M. Laes, E. Bombaerts, B. Ryszawska, B. Rozwadowska, M. Szymanski, P. Ruess, A. Nyborg, S. Hansen, M. E. Community Heroes and Sleeping Members: interdependency of the tenets of energy justice. Submitted to *Science and Engineering Ethics*.

Chapter 6 is based on the published article: Astola, Mandi. (2021) Do Collective Epistemic Virtues have to be Scaled-Up Individual Virtues? *Social Epistemology Review and Reply Collective* 10 (7): 40-47. <https://wp.me/p1Bfg0-62a>.

2 Mandevillian Virtues



2.1 Introduction

The first research question I outlined is: *Do irreducibly collective behaviors of groups create problems for virtue ethics as a moral theory?* In this chapter, I answer this question in the affirmative, at least for the case of pure virtue theory. Pure virtue theory is a virtue ethics accompanied by the meta-ethical claim that all moral value or disvalue is grounded in virtues and vices (Zagzebski 1996). This means that virtue theory must be revised and I present such a revised theory. I hereby tackle the second research question: *How can virtue ethics respond to collectively valuable phenomena which are not reducible to individual virtue?*

Virtue ethics is traditionally concerned with the character traits and flourishing of individual persons. However, an important part of moral life is the way individuals can contribute to the collective flourishing of groups or communities. Is it possible that traits that help a community to flourish can sometimes differ from the traits that help an individual flourish? Within virtue epistemology, many authors have embraced the

surprising possibility that bad traits can occasionally be good in a collective context. Recent work on epistemic traits has described how epistemic vices of individuals can sometimes lead a group to flourish epistemically.

Paul Smart (2018a) argues that the suboptimal cognitive performance of individuals within a group can actually benefit the group in carrying out epistemic tasks. Smart calls this *mandevillian intelligence*. He refers to studies involving groups of people and computer-simulated nodes. The experiments show that some tendencies make it less likely for an individual node to answer a question correctly, but increase the likelihood that the network they are a part of converges on the right answer faster. Smart describes the phenomenon as an instance of collective epistemic virtue arising out of individual epistemic vice (Smart 2018b). Other authors have also described in recent years the way individual epistemic vices contribute to emergent epistemic goods (Levy and Alfano 2019; Morton 2013). In some cases, knowledge arising out of bad epistemic character can be a lucky fluke. However, in other cases the dynamic of a group can be such that an individual who consistently displays vicious behavior, consistently leads the group to good epistemic behavior. Stable, long-term, epistemically disvaluable traits, or epistemic vices, if embedded in the right way within a collective or group, can lead the collective or group to behave in an epistemically virtuous manner (Smart 2018b).

Mandevillian intelligence has a counterpart in ethics, which I shall call *mandevillian morality*. The spirit of mandevillian morality is closer to Mandeville's message in the work *The Fable of the Bees or Private Vices Publick Benefits*, where he likens a capitalist economy to a beehive. Each bee acts out of self-interest but this has the systemic effect that the hive functions well and keeps all the bees alive (Mandeville 1714). My definition for mandevillian morality is however more specific. I mean instances where the moral vices of individuals contribute to virtuous behavior at the collective level. Mandevillian morality refers to stable, long-term character traits of individuals, which cause stable, long term behavior in groups. Mandevillian morality presents a challenge for most virtue ethical theories because it has the counterintuitive consequence that vice can sometimes be valuable or morally good. Mandevillian morality is the virtuous behavior of groups, as a result of vicious behavior of individual agents within the group.

Mandevillian morality can arise in different ways. Particularly in non-ideal situations where there is a prevalence of vice, the prevalence of the vice at the opposite extreme can make the group function at the golden middle. For example, in a group of cowardly collaborators, an overly confident member can push the group to behave

more courageously. Human resource managers often use opposing strengths and weaknesses to create a well-functioning team. In the words of Belbin, famous for developing the Belbin team roles, “What is needed is not well balanced individuals, but individuals who balance well with each other” (Belbin 2012; P. 319, Bjarne 2012).

Many virtue theorists adopt what is called a *pure virtue theory*. A pure virtue theory is a virtue theory which sees virtues as the only fundamental source of moral good (p.79, Zagzebski 1996). If one accepts a pure virtue theory, then one necessarily maintains that all moral good boils down to individual virtue. In cases of mandevillian morality collective morally valuable behavior does not boil down to individual virtues only, but also, importantly, to individual vice. This, I argue, presents a challenge for pure virtue theories. A pure virtue theorist will struggle to account for the moral good of mandevillian morality, because it cannot be explained through a connection to virtue only.

How, then, should pure virtue ethics account for mandevillian morality? I suggest that recent work in virtue epistemology can provide guidance. However, recent work in epistemology has expanded the notion of epistemic virtue in such a way that Mandevillian intelligence is more easily accounted for. Various authors have proposed theories about shared-belief formation, collective epistemic virtues and the social extension of epistemic virtues (Alfano and Skorburg 2016; Fricker 2010; M. Gilbert 1987; Levy and Alfano 2019; Spyridon Orestis Palermos 2016; Pritchard 2018) Such developments allow epistemic virtue theories to account for emergent epistemic goods, even when they arise out of vice. However, moral virtue theories have not seen this kind of developments. Hence, this paper is an exploration into the possibility of similar expansions to virtue ethics, so that mandevillian morality can also be accounted for.

Accounting for mandevillian morality would make virtue ethics, especially pure virtue theories, stronger as a moral theory, especially with regard to group, community and collaborative contexts. If a trait that is clearly a vice leads a group or community to flourish, then pure virtue ethics comes to a trilemma: either the trait is not a vice, the group is not really flourishing or that vices can sometimes lead to group flourishing. All of these are counterintuitive in true cases of mandevillian morality.

Mandevillian morality can arise in business, governance, innovation, science and other collaborative areas of life. There are plenty of true stories about flawed or devious characters leading groups and organizations to success. Therefore, if a virtue ethical response to mandevillian morality can be formulated, then this may prove to be useful in fields of applied ethics concerning such contexts. It would also make pure

virtue theories stronger against possible objections to their applicability to collective life.

I propose that virtue ethics can account for mandevillian morality if a multitude of types of virtue is recognized, much as recent virtue epistemological work has done. I put forward an account of moral virtue, with the following three distinctive components:

1. a distinction between collective and individual virtues
2. a distinction between motivational and purely teleological virtues
3. a strong role-based virtue theory which acknowledges “vicious” roles.

The three above distinctions provide three possible ways to explain how a vice, which is fundamentally bad, contributes to morally virtuous group behavior, which is fundamentally good.

In section 2, I give an extensive definition of mandevillian morality. In section 3 I explain what is meant by pure virtue theories. In section 4 I argue that mandevillian morality presents a challenge to pure virtue theories. Section 5 presents a structure for an ethical virtue theory with the three features above that allow it to answer the challenges of mandevillian morality, sometimes drawing inspiration from the field of virtue epistemology. The concluding remarks in section 6 elaborate on how far the three features align with many common intuitions about virtues and present some questions for future research.

2.2 What Is Mandevillian Morality?

To begin, the definition of mandevillian morality is inspired by Smart (2018a)’s definition of mandevillian intelligence, which is:

Cognitive and epistemic properties that are typically seen as shortcomings, limitations or biases at the individual level can, on occasion, play a positive functional role in supporting the emergence of intelligent behavior at the collective level (Smart 2018a).

I define mandevillian morality as:

A specific form of stable and long-term, ethical group agent behavior in which character traits that are typically seen as moral vices at the individual level play a structural role in constituting this ethical behavior at the group-agent level.

In what follows, I shall explain some of the terms in this definition. By “ethical group agent behaviour” I mean morally valuable behavior, such as courage or reflectiveness, manifested by a group agent. This ethical behavior can be a description of how the

group behaves towards itself, like the members respecting each other appropriately or ethical behavior regarding the shared goal of the group, like the group taking on a supererogatory courageous challenge.

In order to count as mandevillian morality the group agent behavior must be stable and long-term. One-off good consequences arising “accidentally” out of vicious actions are to be excluded. I have chosen to limit the definition of mandevillian morality to applying to “group agents” rather than collectives, as Smart does with mandevillian intelligence. I see a “group agent” through List and Pettit’s definition as a collection of people which collectively possesses three characteristics. The group possesses representational states, motivational states and the capacity to process representational and motivational states, to intervene in the environment, so that the environment is brought to match the motivational state (List and Pettit 2011).

By “moral vices at the individual level” I mean disvaluable, or morally bad, traits of individuals, which in the broader context of a person’s life inhibit their flourishing. In the Aristotelian tradition, vices are generally seen as excesses or deficiencies, like foolhardiness or cowardice. In between the two extremes is a virtue, like courage. If such a trait structurally contributes to stable, long-term ethical behavior of a group agents, then it is a vice which causes mandevillian morality.

What examples are there of mandevillian morality? Individual vices can create ethical group behavior in many ways. One example is third-party punishment of norm violation, which can be beneficial for the establishment and maintenance of social norms (Fehr and Fischbacher 2004). This could mean that in some cases a disposition to spite can help maintain moral norms in a community, especially if others in the community are less disposed to punish third parties.

Imagine a company with a very good image in terms of sustainability. Now imagine that the sustainability of the company is maintained by a fiercely wrathful CEO who rules by fear. She sets high standards and threatens to fire whoever does not take sustainability seriously. The company is held in high regard as an example of it being possible for a commercial enterprise to make money without compromising its values.

In this example, the CEO possesses the vice of wrath. Because of this, she is particularly disposed to engage in third-party punishment. The punishment of employees who fail to prioritize sustainability ensures that sustainability is consistently prioritized. Hence, the wrathful can systematically and consistently keep the group behaving more ethically than they would if she were not there.

The addition of a spiteful individual to a group of benevolent and non-spiteful people can make the group more just. Vices can also add to the deliberative capacities of a group. A bigot in a group is annoying and possibly destructive to ethical group behavior. However, if embedded in the right kind of group, a bigot might cause more dialogue. This can make a group more reflective as a whole because the bigot raises discussions about morality and probably some basic assumptions that everyone in the group accepts or conforms to.

Let us consider an example. Imagine a small company of about fifteen people with a familiar and friendly atmosphere. Many of the employees share a taste for excess with regard to drinking. The social cohesion in the company is therefore greatly helped by monthly events where all employees get drunk together.

Recently, a new employee has joined the company. He does not drink for religious reasons. When planning next month's social activity, the group considers whether to continue the heavy drinking tradition now that there is a new employee who does not drink. One colleague suggests that it might be nice to plan a non-drinking activity for once in the coming month and see how it goes. Many of the colleagues present sigh and silently wish that they could just go drinking again. Some roll their eyes. Another colleague, Lisa, who was really looking forward to getting drunk together again, says: "Here in this country we go to the bar, and if you don't like it, you can go back to where you came from."

Lisa's blatant discriminatory comment is met with shock and an embarrassed silence. Soon after the others challenge Lisa, explaining to her why her comment is inappropriate. After some discussion, Lisa agrees to correct her behavior. Shocked by Lisa's comment, others reflect on themselves and their own previous eye-rolling. The desire to include their new colleague and make sure he is protected from discrimination is cemented in the collective consciousness of the group.

Like this example shows, vice can sometimes make us even more conscious of the morals we hold, by disrupting them and confronting us with their opposite. As Swierstra puts it: "'cold' morality turns into 'hot' ethics: invisible, solid, moral routines become fluid in ethics" (Swierstra 2016). There is undoubtedly something valuable about this event, where people are confronted with each other's differences and engage explicitly in moral thinking. While intuitively convincing, it has also been argued more systematically by Fabienne Peter that deliberation about ethical values is not merely instrumentally valuable, but also has procedural value. Having procedural value means that it is not merely valuable for the purpose of achieving some valuable end, like consensus or a sense of community, but rather that the

deliberation itself has some intrinsic value which makes it good regardless of the consequences (Peter 2013). Behavior resulting from vices reliably initiates these valuable events. This would be a prime example of mandevillian morality.

Sometimes epistemic emergent good also arises from individual moral vice, or moral emergent good from epistemic vice. Adam Morton (2014) discusses a phenomenon, which I would interpret as a hybrid form of mandevillian morality and mandevillian intelligence. Namely, when individual moral vices lead to better epistemic group behavior. One example that Morton discusses is that epistemic self-indulgence can, if embedded in a group in the right way, make the group flourish more epistemically. Epistemic self-indulgence is when people seek confidence or respect by deviating into fields of knowledge where they feel less challenged by epistemic authorities. For example, teenagers are known to seek knowledge about things that do not interest adults, to be in a space where they can be more of an epistemic authority than elsewhere. The presence of such individuals in an epistemic community often means more epistemic rivalry, conflict and hence more better supported views and theories as a consequence of having to defend one's position against diverse attacks (Morton 2014).

Another example of moral vices leading to epistemic benefits is the systemic benefits of dogmatic group cognition in democracies. Citizens are sometimes overly dogmatic regarding political issues that touch their own social group. Dogmatism can be seen as both an epistemic and a moral vice. Lepoutre discusses the example of city dwellers who worry about the environment and underestimate the effect of environmental measures on job creation (Lepoutre 2020). Rural populations who see a lack of jobs in their environment may overprioritize the issue of job creation over environmental protection. Both opinions may be coddled by negative perceptions about the other social group, i.e. that city dwellers are arrogant. This “dogmatic group cognition” can be beneficial systemically. In his words: “Rural inhabitants who overprioritize job creation will be more disposed to uncover evidence of how environmental regulation harms rural jobs... ..Conversely, environmentalists may, because of their contrasting epistemic standards, be more effective at finding flaws in evidence that environmental regulation harms jobs and more likely to generate new evidence of climate change” (Lepoutre 2020).

I emphasize that mandevillian morality is stable and long-term and involves a structural contribution of individual vice. This is what distinguishes it from other moral-luck-like phenomena. Older debates on moral luck and other areas have for long discussed the possibility of good things arising out of bad actions. A politician

might lie in order to influence others to get a life-saving bill passed. Walzer states in a 1973 article on *the problem of dirty hands* “No one succeeds in politics without getting one’s hands dirty” Subsequently, the author asks “But one’s hands get dirty from doing what it is wrong to do. And how can it be wrong to do what is right? Or, how can we get our hands dirty by doing what we ought to do?” (Walzer 1973). A single instance of a politician doing a good action with dirty hands, is not mandevillian morality. However, if this politician’s stable tendency to lie contributes to the parliament collectively taking greater efforts to fact-check, or helps the parliament avoid harmful stagnation, then it is mandevillian morality.

2.3 Pure Virtue-Ethical Theories

Mandevillian morality is problematic if one adopts a *pure* form of virtue ethics. In order to show this, I will first define virtue ethics and what such a pure virtue theory entails, beginning with the definition of virtue.

Virtue theories are normative theories that place the good in the character traits of people. Moral virtue theories focus on the moral character of people, which lead them to flourish as human beings (Hursthouse and Pettigrove 2018).

The central entity in virtue ethics is the virtue, or the morally good character trait, as well as its opposite, the vice. Zagzebski characterizes a virtue by listing the distinctive properties of virtues:

1. A virtue is “an acquired excellence of a person in a deep and lasting sense” (p.135, Zagzebski (1996)).
2. “A virtue is acquired by a process that involves a certain amount of time and effort” (p.135, Zagzebski (1996))
3. A virtue is not just a skill. A skill is distinguished from a virtue in the following ways:
 1. a skill need not be exercised, but a virtue must be exercised to be possessed
 2. The difficulty of skills is due to technical difficulties of the actions themselves, the difficulty of virtues arises generally from a lack of sufficient passion
 3. A virtue, unlike a skill, cannot be compatible with voluntarily giving it up. It is characteristic of having a virtue that one would never want to give it up.

4. “The behavior consisting in the exercise of a skill is not essentially connected to anything valuable” (p. 113 Zagzebski (1996)) A skill is good through the objects or acts it produces, whereas a virtue is good to have in itself.
 5. A skill is not affected if distractions or persuasion prevents a person from exercising it, but a virtue is affected by this.
 6. a skill has no contrary, or opposite, whereas virtues have vices as opposites. ²
4. A virtue has a motivational component. The person possessing the virtue has the desired emotional disposition. i.e. if a person is benevolent, they are motivated to act benevolently by feelings of duty or empathy, rather than merely by the prospect of improving one’s reputation.
 5. A virtue has a success component. This means that a person who has a virtue is also generally reliable in bringing about the desired outcomes of a virtue. I.e. if someone is benevolent, they also tend to succeed in their efforts to help others.

Zagzebski and others also make a distinction between moral and epistemic virtues. Epistemic virtue theories focus on the epistemic character of people that leads them to flourish as human beings, or at least as knowers (Code 1984; Zagzebski 1996)). Similarly, moral virtue theories focus on character traits like honesty, generosity and justice. These traits are developed through practice and help one to lead a good life. Moral and epistemic virtue theories are also popular in application to collaborative contexts like science, business and innovation (Demetriou 2013; Hicks and Stapleford 2016; Sand 2018; Steen 2013).

A pure virtue theory is described by Zagzebski as “a theory which makes the concept of a right act derivative from the concept of a virtue or some inner state of a person that is component of virtue” (p.79, Zagzebski 1996). Pure virtue theories are held, notably, by Anscombe and MacIntyre. Anscombe famously argued that deontological theories are vacuous in the absence of a divine legislator, and hence untenable in a secular worldview. Consequentialist theories, cannot be moral theories because they permit obviously immoral acts. Virtue theories are the only way in which a moral notion of the good can reasonably be approached (Anscombe 1958). MacIntyre defended the fundamentality of virtues in a similar way, arguing that other

² For a more detailed discussion on the difference between skills and virtues, see section 2.4 in Zagzebski (1996).

moral theories collapse into vacuous statements without the content of virtues and vices (Macintyre 1986).

Of course, some see virtues as additional to other moral theories. Nussbaum, for instance, argues that virtue theories are not rival theories to utilitarianism and deontology, but rather something additional to both (Nussbaum 1999). However, the pure virtue theorist rejects utilitarian and deontological notions of good completely.

2.4 Why Mandevillian Morality Challenges Pure Virtue-Ethical Theories

Let us return to the first research question: Does the moral value of irreducible group level phenomena create problems for, or limit the scope of virtue ethics as a moral theory? We can now make this question more specific: does mandevillian morality cause problems for, or limit the scope of pure virtue ethics? If we answer the latter question in the affirmative, we must conclude that at least some group level phenomena cause problems for or limit the scope of virtue ethics as a moral theory.

Mandevillian morality is widespread and important in human moral life, and virtue ethics should be able to explain and account for it. However, mandevillian morality has some consequences that seem to misalign with a typical picture of virtue. If an individual vice plays a systematic role in morally good group behavior, then it is reasonable to say that that vice is valuable. However, in virtue ethics vices are supposed to be disvaluable, the opposite of valuable. In the case of mandevillian morality, a person might reasonably be praised or rewarded for their vice. This is counterintuitive from the perspective of virtue ethics. This counterintuitiveness stems from the fact that a virtue ethicist upholding pure virtue theories will struggle to explain what is good about good group behavior, like collective reflectiveness, in a case of mandevillian morality.

Within a pure virtue theory especially, vices being systematically valuable is problematic. Because if one sees virtues as the foundation of morality, then one explains every instance of moral good with reference to the virtues. If confronted with mandevillian morality, a pure virtue theorist would be forced to either admit that the good group behavior is not morally good or that the vice in question which plays a role in this behavior is actually morally valuable.

This problem does not necessarily apply to virtue theorists who see the virtues as a supplementary concept to other moral theories. If one argues, with Nussbaum, that virtue ethics is not a “third category” but an addition, then one can use virtue ethics to account for the good of individual character and another theory to account for

collective phenomena (Nussbaum 1999). In this article, however, I aim to investigate in how far virtue ethics itself can be modified to account for these collective and individual phenomena simultaneously.

Human moral life happens interpersonally. Much of our life is spent collaborating and contributing to groups and communities. Vicious character, whether it is epistemic or moral, can have an important role in such a group or community. If we uphold a pure virtue theory, and accept that moral virtue is the source of moral good, then an explanation is needed for how vices that lead to group flourishing can still be vices. If pure virtue theories can provide such an explanation, then they are better suited for prescribing ethical conduct for group and communal behavior.

2.5 Three Features of a Virtue Theory that Can Account for Mandevillian Morality

The examples in the previous section show that mandevillian morality is irreducible to individual virtue only. Sometimes individual vices or non-virtues play an essential role in creating emergent good. As I have argued, it is hard to explain mandevillian morality in terms of traditional virtue ethics, and virtue ethics would benefit from being able to account for it.

Now we come to the second research question: *How can virtue ethics respond to collectively valuable phenomena which are not reducible to individual virtue?* It seems to me that a virtue-theoretic response can be formulated. There is at least one way in which a virtue theory could account for emergent good. As I shall argue, we can do so by accepting a type of virtue theory which has three essential components, all of which combine into a theory which accounts for the role of vice, or the *virtue of vice*. The three features, which work synergistically in both virtue ethics and epistemology, are:

1. a distinction between individuals and groups as possessors of virtues
2. a distinction between motivational and teleological virtues and
3. recognizing role-based virtues that can attach to “vicious” roles.

The first feature prescribes that groups can also possess virtues and vices in the same way individuals can. A group can develop *mandevillian virtues* which are possessed by the group as a collective agent. The second feature entails accepting that there are two types of virtue: teleological and motivational virtues. Teleological virtues are good because they lead to group flourishing or furthering the groups progress towards a moral telos. Motivational virtues are those which are good mainly because of virtuous motivations.

The third feature calls for attention to the specific virtues and vices connected to each role that people can take in collaborative activities. A group may require a role for “the dogmatist,” “the egoist,” or “the wrathful” in order to expand group creativity or maintain social order.

The three features of the view I explore make cross-cutting distinctions among epistemic or moral virtue theories. While each feature is in itself sufficient in accounting for emergent good, accepting all three allows for dealing with various kinds of mandevillian morality. The last 2 features should be seen as a re-classification of what traits count as virtues. They are categories by which a trait may be classified as a virtue. For instance, a trait which is a teleological virtue can be a procedural vice and a role-based virtue. I will show how each of these distinctions still fits virtue within the confines of Zagzebski’s characterization of virtue.

A Distinction between Individuals and Groups as Possessors of Virtues

Many accounts of epistemic virtues already entail this feature. Theories in virtue ethics and epistemology that attribute virtues to groups are often referred to as *collectivist*. Collectivism is a metaethical position which states that groups can have moral responsibility in the same way that individuals can. The position also has virtue-theoretic followers in responsibilist epistemology and ethics. A root of collectivist thinking is in most cases the *plural subject-thesis* in the work of Margaret Gilbert. M. Gilbert (1987) argues that groups or dyads of subjects, when both convinced of a belief held by a group they belong to, compose a plural subject, a single knower or single agent. Therefore, groups can have moral agency as a unit and carry moral responsibility for the actions that the group takes as a unit.

Fricker states that if all members of a group have some virtuous motive individually and are reliable in bringing about its object, then the group can be said to possess a virtue. She also argues that it is not necessary for members in the group to possess that virtue individually. Members of a group may also “go along with” a motivation, because they see it as appropriate to the role they take in a certain institution. In this case, the group can still be described as having a virtue or a vice. She contends that both Aristotelian virtues, which require a motivational component, and Platonic/Stoic virtues where the virtue is simply an excellence, can be possessed by collectives (Fricker 2010). She does not go as far as claiming that mandevillian intelligence or morality can constitute a group virtue. She discusses the example of a jury where individual, but differing, prejudices cancel each other out, creating fair-

minded decisions. This, she states, is a matter of luck rather than virtue, because the jury cannot be held creditworthy (Fricker 2010).

Fricker's example of a diversely biased jury is indeed not an example of collective virtue, nor mandevillian morality. But this is because the jury in question is a temporary group, not a group agent. A jury is generally assembled for a short period of time and hence does not have time to develop a "group personality."³ It should be noted that virtues are not merely inclinations, but long-term and stable character traits. A collective virtue, like the fair-mindedness of a diversely biased jury, can also be a long-term stable character trait, let us imagine that the jury is assembled to meet regularly over the course of a year, for example. The group dynamic, in which the dogmatism and biases of individuals play a constitutive role, can evolve to a state where the diverse biases cancel each other out in the making of decisions, which may be helped by the biases of the participants also being long-term and stable on the individual level.

On the individual level, developing a "group character" consists of developing appropriate ways of communicating with each member, developing a sense of group identity and perhaps even the individuals learning to take pleasure in certain modes of communication. All of these individual developments give rise to individual motivations that respond to the traits of other members, and keep the group in balance. They might take forms like "Robert talks too much, so I better shut him up before everyone gets bored" or "this group is full of fragile snowflakes, I better assert my opinion louder to toughen them up." If these micro motivations become systematic in the group dynamic, and contribute to group reflectiveness, then they can constitute the motivational component of the group virtue in the same way as Fricker's "willingness to go along with"- motivation can. What makes these micro-motivations count for a virtuous group motivation, is that they are in some sense aimed at group virtue, reflectiveness in this case.

It may thus be possible to account for mandevillian intelligence or morality within virtue theories by acknowledging their value and seeing them as *collective mandevillian virtues*. Then one can reasonably say that a single bigot in a group has a vice, but the group as a single collective agent has the virtue of being reflective, as a result of the bigot. One can still explain the fact that bigotry is a vice, but a group with a bigot that

³ One could also consider institutions, such as the fire department or a grand jury as collective agents, that have stable institutional traits, like courageous behavior or fairmindedness, even if individuals leave the group and are replaced. I will not consider these cases here. My focus is on groups which display certain behavior due to the constitution of the group.

functions as a reflection mechanism has a collective virtue at the same time. A distinction can be made between the individual person and the collective person, which is a collective agent that can possess virtues and vices. This distinction allows us to talk about two different agents, the individual with the vice and the group with the virtue.

If we take Zagzebski's list of properties of virtues, we see that collective virtues can also fit this description and thus that we can meaningfully speak about collective virtues.

1. A collective virtue can be "an acquired excellence of a person in a deep and lasting sense." In this case, the "person" is a collective agent. These ways of operating can also be stable and long-lasting.
2. A collective virtue can be "acquired by a process that involves a certain amount of time and effort." Groups take time, and collective effort to develop fruitful dynamics and ways of operating.
3. A virtue is not just a skill. A skill is distinguished from virtues in various different ways. A collective virtue, like reflectiveness fall into the category of virtues rather than skills. Group reflectiveness, for instance, must be exercised in order to be had. It does not make sense to speak of a reflective group, if the group does not take part in deliberation among members. The good of group reflectiveness is also not merely good because it leads to more informed decisions, better strategies and such. Such a "group trait" is also often described as having intrinsic value (Peter 2013).
4. As explained earlier, group virtue also has a motivational component. This consists of the virtuous motivations of individuals, and/or a "willingness to go along with a virtuous motivation" or systematic motivations towards improving or upholding a good group dynamic, even if these stem partially from vice.
5. Group virtues can also be successful in hitting their desired targets, while possessing the abovementioned qualities. For instance, a target of reflectiveness would be the ability to form new intentional states based on deliberation. In the case of group reflectiveness, these intentional states can be intentional states held by the group agent.

A Distinction between Motivational and Teleological Virtues

In addition to distinguishing between collectives and individuals as possessors of virtues, I propose distinguishing between two categories of virtue. What is referred to

as “teleological virtues” in this paper is often called “consequential virtues” in literature. The word “teleological” is preferable over “consequential” to avoid associations with utilitarianism. Making the distinction between motivational and teleological virtues means that a case of mandevillian intelligence or morality could be evaluated in the following way: the individual has a motivational vice, but this vice is also a teleological virtue. The teleological and motivational component are two different ways in which a trait can be blameworthy or praiseworthy. Motivational vices are traits that are inherently or categorically wrong. Teleological vices would be vices which are mostly bad due to the negative consequences. Motivational virtues are traits like benevolence or honesty, which are inherently good. Teleological virtues are those which are good mostly because they facilitate the end goal of flourishing.

No virtue theoretical account explicitly uses this distinction to postulate two kinds of virtues. However, the distinction seems to be implicit in many accounts when describing the aspects that constitute a virtue. Bradley, who defends virtue consequentialism, also notes that prominent virtue theorists like Philippa Foot, Linda Zagzebski and Julia Driver all embrace virtue consequentialism in some way (Bradley 2005). However, these authors (with the exception of Driver, who favors virtue consequentialism over other accounts (Driver 1989)) all also acknowledge that there is another component to virtue which has nothing to do with the consequences of one’s virtuous behavior. ⁴

Virtues are sometimes described as having a motivational requirement, as well as the requirement of reliably bringing about of valuable consequences. It seems that it is at least plausible that there are virtues where the teleological component is more important or pronounced and others where the motivational component is more important or pronounced. It also seems plausible that there are people possessing a virtuous trait that can be described as more teleological or more motivational in the specific instance of that particular person. Because of this, it is also plausible that some virtues can be described as completely motivational, referred to as motivational from here on, or as completely teleological.

To illustrate the difference, think of two versions of the virtue of honesty. Someone may have acquired a strong motivation to tell others the truth and to practice honesty whenever they have the opportunity. However, the same person may be bad at

⁴ Foot discusses the way virtue can, apart from beneficial consequences to others, be judged by “his innermost desires as well as by his intentions” (p.5, Foot 1978). Zagzebski also argues that virtues are good to have, for their own sake, even if they lead people to less good actions in certain cases, as in the case of a courageous Nazi (p. 91–94, Zagzebski 1996).

communicating, due to extreme shyness or inability to articulate what one means or judge how much another understands. This person is clearly honest, but their honesty does not “hit its target” which is to make others aware of what one thinks is true. This person is motivationally honest, but not, or at least less, teleologically honest. Think of another person, who possesses a different kind of honesty. They are not necessarily motivated to tell others the truth, but they have acquired the habit of reliably doing so, and communicating their perceptions very accurately to others, perhaps because it saves them trouble. However, even if confronted with a situation where it does not save them trouble to be honest, they reliably act honestly, because it is a part of their character and identity. Such a person also possesses the virtue of honesty. However, their type of honesty is a more teleological, and therefore “target-hitting,” type of virtue and admirable mostly for that, rather than for the motivations behind it.

In the case of a bigoted individual who nevertheless makes the group more reflective as a whole: one could say that the individual has a motivational vice. They may be motivationally arrogant or insouciant. This vice can however still be a teleological virtue, like intellectual courage, if it is an ingrained habit which tends to “hit its target” which is challenging unspoken assumptions and leading to fruitful dialogue.

Let us reflect on whether making this distinction aligns with Zagzebski’s list of properties of virtues. Both teleological and motivational virtues must be acquired excellences of the character in a deep and lasting sense as well as involve time and effort to acquire. While virtues traditionally have both a motivational component and a success component, I suggest that some virtues can have more of one than the other. This challenges the idea that having a virtue means doing the right thing, at the right time, for the right reason. However, it does seem to make virtue something more achievable for most people, answering the common critique that virtue theories are elitist.

An opponent may argue against the conception of teleological virtues by arguing that teleological virtues are merely skills. However, teleological virtue can in my view be distinguished from a skill in various ways. Firstly, a teleological virtue, like teleological honesty, must be practiced to be had. No one will be called honest if they do not show honesty. Communication skills, which may benefit one in being honest, do not need to be practiced to be had. Secondly, teleological honesty would be affected if the person was persuaded not to practice it or was too distracted to practice honesty.

Recognizing Role-Based Virtues that Can Attach to “Vicious” Roles

Lastly, I propose a third feature for virtue theories to deal with mandevillian morality: recognition of role-based virtues that can attach to “vicious” roles. Fields of applied ethics for professional contexts sometimes use a role-based ethics. This means that the right traits or actions are dependent on the role that a person has in a particular circumstance. Virtue approaches are also sometimes connected with role-based ethics. In such accounts, what is virtuous is dependent on one’s role and there can be role-specific virtues (Nuyen 2007; Swanton 2016). Levine and Cox (2016), for instance, argue, using MacIntyre’s approach, that virtues can be understood through social practices. The authors argue that due to the specific practices that are involved in academic work, we can distinguish academia-specific virtues, relating to the role of the academic as a researcher and as an educator. Similar accounts can also be imagined for virtue epistemology. For instance, the virtue of “impartiality” may be more relevant for a judge than for a defense lawyer. One’s role as a judge may therefore impose a greater normative demand on developing this virtue. Similarly, one could imagine that a good defense lawyer is precisely someone who lacks this virtue.

An account that can deal with emergent good would benefit from such a role-based conception of the virtues. In the case of mandevillian intelligence or morality, the context is an essential part of what makes the virtue. Egoism or stubbornness can only become a teleological or group-level virtue once it is embedded in the right social context. While one egoist might be essential for a group virtue, ten egoists are perhaps too many. There may be “a role” for an egoist in a specific context, but once that role is occupied, another egoist may simply be destructive to the group virtue. Importantly, an account that can deal with mandevillian intelligence and morality should acknowledge the existence of “evil roles” such as the role of the “dogmatist” or “reflection generating bigot.”

Let us consider Fricker’s example of the jury that consists of diversely biased individuals, which functions in a fairminded way. If one imagines that the jury was compiled, on purpose, to be as diverse as possible, precisely so that it would come to fairminded conclusions, then each member takes a certain role. If each has “being biased” as a part of their role description, then acting in accordance with that bias can be considered virtuous.

Accepting the possibility that virtues can be relative to roles does not expand the notion of virtue in any way that would risk making it incompatible with the definition of virtues I have been using. A role-based virtue can be “an acquired excellence of a

person in a deep and lasting sense” (p.135, Zagzebski 1996) if that role is maintained by a person for a longer period of time. Such a virtue can also be “acquired by a process that involves a certain amount of time and effort” (p.135, Zagzebski 1996) as one takes time and effort to adjust to a role. The motivational component of such a virtue can be an identification with the role, or knowledge of its significance and a passion to fulfill this significance. A “success component” would be the ability of the person to bring about the intended positive effects of having the role in the group or community. Such a role-based virtue can also be distinguished from a skill.

2.6 Conclusion

Traditional virtue theories have often revolved around a notion of balance or harmony between the virtues within the individual soul as well as between different people within society. Plato’s virtues are distributed unevenly in society. The soldier lacks the wisdom of the philosopher but makes up in courage. Together the different members of the state can form a harmonious whole (441ad, Plato 2010). In this paper, the central claim has been that virtue theories would benefit from acknowledging the way that vices too are distributed unevenly in a community, and how they can also contribute to a harmonious whole. If ones sees virtues as the right middle between two vices, then this golden mean can also be achieved between two vicious people. A group where all the vices are represented can develop a collective way of working wherein all the vices are balanced into all the virtues, so that the collective agent is virtuous.

There may be a case where a bigot causes a group to be more reflective. We may blame the bigot for her vice, but we may praise the group, including her, for having the virtue of deliberation. Imagine that her bigotry does not result in better group morality, but rather, to the fact that the group manages, partially thanks to her bigotry, to innovate a tremendously valuable innovation. Then we may blame her for her bigotry but at the same time praise her for her teleological virtue of having that same trait which pushes the group forward towards creative directions. We may understand this teleologically beneficial trait as intrinsically good by virtue of the role of “representing evil,” “expanding the collective imagination” or “triggering action” that she occupies in the group. These roles can, just like group virtues, be acquired and developed over a period of time.

Of course, a mandevillian “role” within a group can be difficult to instrumentalize

An opponent may worry that accepting collective, teleological and evil-role-based virtues is a danger to the importance of individual character development. One might

argue that if an account of virtue integrates vice, this allows people to explain away their vices as virtues and no improvement is incentivized. A twofold response can be given: Firstly, the opponent's question could be asked the other way around too: whether rejecting these kinds of virtues is not a danger to the importance of virtues arising from group dynamics. Humans live and collaborate in groups, so why should a virtue theory prioritize the importance of the individual as an agent over the group as an agent? Secondly, the advantage of the account I put forward is precisely that individual virtues remain valid and important, while emergent virtues are also acknowledged. If anything, this allows for a wider spectrum of appreciation for the valuable kinds of character development and group development that people undertake.

The approach to virtue that I recommend makes distinctions between different types of virtue, different roles and different kinds of possessors of virtues. While it still leaves many questions open, it is a plausible answer to mandevillian morality, mandevillian intelligence and similar phenomena that pose a challenge to virtue ethics. Other possible responses also merit investigation.

3 Can Creativity be a Collective Virtue?



3.1 Introduction

In the previous chapter, I argued that vicious people can be good when they are embedded in a group. This poses problems for the pure virtue theorist, who thinks all moral good originates in virtue. I proposed a solution for the pure virtue theorist. This solution includes ascribing virtues and vices to groups as well as individuals. This seems to provide an answer to the puzzle of mandevillian morality. But can it also say

something insightful about other situations? Does the framework of collective virtues have any merits apart from its answer to mandevillian morality? Are collective virtues plausible? And do we see evidence of them in our moral lives? This leads us to the third research question: *Does collective virtue theory produce valuable insights when applied to real-life cases?*⁵

In this chapter, I provide evidence that collectivist virtue theory is superior to mere individualist virtue theory in explaining one important virtue of collectives, namely creativity.⁵ Let us begin with an example.

Pixar has quite exceptionally managed to maintain a high standard for the movies it releases. In 2021 the animation studio holds 16 academy awards, 10 golden globes, and many more awards and nominations that it has been winning steadily since 1995 (Wikipedia 2021). Leadership scholar Linda Hill, who studied the leadership at Pixar and other companies, notes that Pixar productions are products of a special kind of innovation process. At Pixar, there is no leader with a vision who leads others to execute that vision. Instead, Pixar's working culture is conducive to something Hill calls "collective genius" (Hill 2014). Every employee, regardless of rank, has the opportunity to give the director of a movie notes. The team is diverse and arguments and conflicts happen often and are even encouraged (Catmull 2008; Hill 2014). It is often said that the best ideas can come from unexpected sources. John Stuart and Harriett Taylor Mill famously argued that an idea becomes better when exposed to a rich multitude of perspectives. The best way to gain knowledge is also by "studying all modes in which it can be looked at by every character of mind" (Ch. 2, Mill 1956). Additionally, recommendations for diverse teams, for productivity reasons, are often found in management literature (Bassett-Jones 2005; Belbin 2012; Paulus 2003). Work in design studies suggests collective creativity plays, or can play, a role in many different areas of product and service innovation (Sanders and Stappers 2012). This type of literature shows that creative innovations do not always originate in the heads of individual geniuses. Often genius ideas are co-created and their origins can best be described as interpersonal. Hill even argues that visioneering is generally not an appropriate way to lead innovation. She favors Pixar's approach to facilitating collective genius (Hill 2014).

Just like the conception of "genius," innovation is heavily associated with creativity (Sawyer 2012). And creativity is seen by many philosophers as an inherently valuable

⁵ The article on which this chapter is based was co-authored. However, I am writing in first person for reasons of better readability. I do this with permission from the co-authors Andreas Spahn, Gunter Bombaerts, Lambèr Royackers.

character trait, or virtue (Kieran 2018; Zagzebski 1996). Recent work in ethics of innovation has therefore rightfully called for attention to the way innovation processes, rather than products, contribute to human flourishing. This work has, in particular, recognized the inherent value of creativity as a virtue (Sand 2018; Steen 2013). Virtue theories see inherently valuable character traits, or virtues, as fundamental units of moral good. The good life consists of the development and practice of the virtues (Aristotle 2015; Zagzebski 1996). Approaches to innovation ethics, like value sensitive design and technology assessment, often equate ethical innovation with the products of innovation and their societal consequences. However, if innovation processes are bound up with creativity, then the processes of innovating may contribute to human flourishing just like the products of innovation do.

However, none of this work has considered creativity as a collective phenomenon. I supplement the existing work on the virtue of creativity in innovation by exploring whether creativity can be a collective virtue in organizations that engage in product and service innovation. The purpose of this article is to lay down a possible groundwork for a collectivist account of creativity as a virtue, in response to the interest that other fields have taken to this phenomenon. Recognizing this collective virtue, I argue, is a valuable addition to existing accounts of creativity as a virtue in innovation. Recognizing creativity as a collective virtue will make virtue accounts more accurate about how much successful innovation happens. If one takes our argument seriously, then one will recognize the moral value of facilitating collective creativity within innovation practices, in addition to nurturing the individual creative genius.

Our argument proceeds as follows: first, I outline a definition of creativity as a virtue. Since the most complete account of creativity as a virtue has been constructed in the context of aesthetics and philosophy of art, I will use such an account, namely, that of Matthew Kieran. Second, since Kieran's definition of creativity as a virtue is context-dependent, I construct an account of what it means to have the virtue of creativity in the context of product and service innovation. Thirdly, I explore whether it is possible for that conception of creativity to apply to an innovation group. I ground this argument in work on collective virtues, and in our conception of what creativity means in the context of innovation.

I conclude that creativity in the context of innovation can be a group virtue. I am not claiming that the virtue of creativity cannot be possessed by individuals, but that it can also be possessed by collectives. The collective creativity described in management literature can in fact be seen as an inherently valuable group trait.

Philosophy of innovation will benefit from devoting more attention to group traits like creativity driving innovation. For one, this will make innovation ethics more accurate regarding how innovation actually happens. Secondly, there is a normative argument to be made for recognizing collective creativity. Often only the leaders of successful projects are acknowledged for collective efforts, when the creative outcomes are not necessarily products of their creativity only. There is a sense in which this is an unfair practice of credit-giving. It could be argued that the idolization of individuals, leaders, or visionaries also leads to an individualistic and competitive culture which has a well-known set of disadvantages. Again, I do not claim that innovation ethics should focus only on collective creativity, and that individual creativity is irrelevant. I think that the collective virtue perspective is an often-overlooked one that deserves more attention.

3.2 When Does One Possess the Virtue of Creativity?

To begin, it must be established what is meant by creativity as a virtue. What we make of creativity partially depends on the kind of virtue theory we endorse. I base my account on a framework where the virtues are seen as “character virtues” rather than strictly either moral or epistemic.⁶ Within such an account, the virtue of creativity would also be seen as an excellence of character with a relationship to all transcendentals, rather than focusing on how creativity relates to the true or the good in particular. If one sees creativity as a trait in service of either the good or the true, one may define it differently. For instance, a philosopher investigating creativity as an epistemic virtue may be more inclined to define it as a form of insightfulness (Carter 2017). We consider creativity as a character trait with a relationship to the good, the true and the beautiful.

The most complete account of creativity as a character-virtue, with a relationship to all transcendentals, is that of Matthew Kieran. He distinguishes creativity as a capacity which most people have, from *creativity as a virtue*, which only exceptionally creative people possess (Kieran 2018). A person with the virtue of creativity, he writes “is someone who has acquired a certain degree of mastery and knows what she is doing in coming up with novel and worthwhile ideas or artifacts. In doing so, she is motivated by the values internal to the relevant domain and chooses what she does for reasons that hook up with those values in the right kind of ways” (Kieran 2014).

⁶ MacIntyre and Anscombe also exemplify such a view of the virtues as I describe (Anscombe 1958; MacIntyre 1986).

I distinguish three components in Kieran's definition. I will refer to these three components as teleological, procedural and motivational. Each of these components is necessary but by itself insufficient for possessing the virtue of creativity. A virtuously creative person must possess them all. I will explain these briefly and how this dissection of components of virtue can be supported by other work in virtue ethics.

Teleological

The teleological part of Kieran's definition is the requirement: "coming up with novel and worthwhile artifacts." This refers to the successful reaching of the end-goal or *telos* of creativity. The *telos* of creativity is the ability to bring about novel and worthwhile artifacts. Kieran takes this component from Boden's definition of creativity, "the ability to come up with ideas or artifacts that are new, surprising, and valuable" (pp. 1, Boden 2003).

Most virtue theories acknowledge a "teleological component" in the definition of virtue. Virtue theories are sometimes even categorized together with utilitarianism as teleological theories, as opposed to deontological theories (Macdonald and Beck-Dudley 1994). This is because virtue ethicists generally accept that cultivating all the virtues leads to a morally valuable goal, a flourishing life, or eudaimonia.⁷ Driver makes the goal of hitting the target the focus of an account of consequentialist virtues (Driver 1989). However, for each virtue individually, there is also a certain goal. Zagzebski phrases this as the virtue "hitting its target" (Zagzebski 1996). The target of benevolence, for instance, is helping another. She argues that we would not, and should not, praise someone as being benevolent if they meant well but kept acting in such a way that others are disadvantaged by them rather than helped. Such a person may have a virtuous motivation, but does not hit the target, and hence does not possess the virtue.⁸

If we follow Kieran's definition, the target of the virtue of creativity is "the ability to come up with ideas or artifacts that are new, surprising, and valuable" (Kieran 2014). The meaning of "new, surprising, and valuable" requires some reflection. Kieran does not go into this and investigating exactly what this means would be a whole other article, hence, I make some assumptions. "New" I will take to mean that the artifact or idea did not exist before. "Surprising" I take to mean that there is

⁷ Philosophers however disagree on whether virtues are grounded in the way they lead to eudaimonia, or whether eudaimonia is simply a common consequence of practicing the virtues.

⁸ For more discussion on this, see Zagzebski (1996).

something about the artifact or idea that makes people feel surprise. Some virtue ethicists see the emotion of admiration as a way of recognizing moral virtue (i.e. Zagzebski (2017)). Perhaps the emotion of surprise is a way of recognizing originality in an artifact or idea which hits the target of creativity. Kieran suggests that the meaning of “valuable” or “worthwhile” should be determined in accordance with the nature of the activity one is undertaking. For the context of art, most would agree that to be valuable, an artwork should be beautiful or powerful. However, what is valuable as an artwork is not necessarily valuable as a scientific hypothesis. Hence, for each context, the meaning of “valuable” should be evaluated differently.

One may question whether creativity should be assessed in such a context-dependent way. After all, creative people often show creativity across a range of different contexts and activities. However, if we hold the view that virtuously creative people produce valuable outputs, then we must have some basis for what counts as valuable. This context-specificity of what counts as valuable aligns with the kind of virtue theory that emphasizes tradition and culture as determining the substance of the virtues, in the tradition of MacIntyre (Macintyre 1986).

Procedural

In addition to producing new, surprising and valuable artifacts or ideas, Kieran argues that creativity as a virtue also requires “a certain degree of mastery” and that the agent “knows what she is doing in coming up with novel and worthwhile ideas or artefacts.” This, I see as the procedural component of Kieran’s definition. As an illustration, Kieran discusses the hypothetical case of a stroke patient who tries to write but due to neural damage, ends up producing sequences of beautiful and artistic markings on paper. Although the person may be acting from a virtuous motivation, is exhibiting perseverance, and is producing something novel, we would not consider them creative in this action (Kieran 2018). Hence Kieran requires that creative persons have the kind of control over their creations that would separate them from this hypothetical stroke patient.

Kieran outlines different types of control that creative people often have over their creations. Firstly, a creative person can be agential in fulfilling her intentions when creating. For instance, one expects that a painter making strokes on a canvas intended the strokes to look as they do, and that she or he would be able to paint different looking strokes if she or he decided to. Secondly, a creative person can make judgements about whether to keep or discard “happy accidents” or creations which were unintended (Kieran 2018). A painter can accidentally splash paint on a painting

and decide to keep the stain, because she or he decides that it contributes to the artwork. The point is that this kind of judgement requires mastery and relevant understanding. The painter uses his or her aesthetic understanding to decide whether the stain can become a part of the painting or not.

The conception of mastery seems to imply authorship. There are various discussions in the philosophy of art and aesthetics about what exactly counts as authorship.⁹ These discussions reach far outside the scope of this paper. For our purposes here, I accept Kieran's suggestion that mastery requires a match between the artist's intentionality and the new, surprising and valuable output, in such a way that the artist ultimately "knows what they are doing."

Motivational

Most virtue theorists, including Kieran, recognize the need for a virtuous motivation as a part of any virtue. The motivational component, also needed for the virtue of creativity, is encompassed by the last part of Kieran's definition: "In doing so, she is motivated by the values internal to the relevant domain and chooses what she does for reasons that hook up with those values in the right kind of ways." A motivational component is necessary for the trait to be a virtue, Kieran states citing Aristotle (Kieran 2018). For each virtue, there is a set of proper accompanying motivations. For instance, we would not praise a person as benevolent, if we knew for a fact that their benevolence was motivated by egoism, and the ambition to improve one's reputation. A benevolent person must be motivated by a motivation proper to the virtue of benevolence, like love, duty or empathy (Zagzebski 2017).

Kieran argues that the right motivation for the virtue of creativity is an *intrinsic motivation for that particular creative activity*. Think of the delight that poets or artists take in the process of creating their art and of expressing themselves through it. This is wholly different from someone making a painting only in order to make money from it, or other extrinsic motivations. Kieran defines intrinsic motivation as being driven by "the values internal to the relevant domain." Therefore, a creative artist can be intrinsically motivated by aesthetic usage of color and form, expression or any other value inherent to art. Kieran discusses the example of a creative hedge-fund owner. The virtue of creativity in hedge-fund management seems paradoxical because the relevant values in hedge-fund owning and managing have to do with gaining money, a typical external reward. However, Kieran argues that a creative hedge-fund owner

⁹ For a discussion of authorship in art, see Irvin (2005).

should not be motivated by earning money in the sense that he or she is just doing it to have money, but he or she must take pleasure in the practice of money-making itself. If one is motivated by an interest in hedge-funds and finance, or enjoys thinking outside the box when it comes to strategy, then this can be an intrinsically motivated hedge-fund owner (Kieran 2018). Again, what intrinsic motivation means for a hedge-fund owner is substantially different from what it means for an artist. The “relevant values” are different in different creative contexts. The motivational component is therefore also context-dependent.

Kieran does not take a strong stance on the place of emotions in creative motivations. Other philosophers have however described the role of emotions in virtue in more detail. Zagzebski describes the virtuous motivation as “a persistent tendency to be moved by a motive of a certain kind” a motive being “an emotion or feeling that initiates and directs action towards an end” (pp. 131–132, Zagzebski (1996)). We can model the “intrinsic motivation” requirement with Zagzebski’s view. Then one could see an intrinsically motivated person as someone with a persistent tendency to be moved by motive x. Motive x can be an emotion like inspiration or love of the creative medium, whether it is the language of a strategy document or paint on a canvas.

To summarize, the virtue of creativity requires a teleological component of producing new and valuable output, a procedural component, having a degree of mastery and a motivational component, intrinsic motivation. Yet in some artists, musicians or scientists, not all of these components are present. For instance, someone may be able to create novel and worthwhile artifacts through mastery, but be solely extrinsically motivated. In these cases, Kieran states, we can attribute creativity to the person, but it is not creativity as a full virtue of character (Kieran 2014). To possess the virtue, an agent must fulfill all requirements of this definition. Various virtue ethicists have distinguished virtue from capacity or skill in this way. A virtue is much more demanding and much more valuable, although skills and capacities also have value or are desirable.

3.3 When Does an Innovator Possess the Virtue of Creativity?

The previous section details how a substantive account of each component of the virtue of creativity is context-dependent. The virtue of creativity looks different in a painter than it does in a hedge-fund owner or football player. Therefore, for the

context of product and service innovation, the virtue of creativity will have its own qualities that align with the values and the purpose of the context.

When, then, does a product and service innovator possess the virtue of creativity? To answer this question, I apply the framework based on Kieran's definition to the context of innovation, to develop an innovation-specific definition of creativity as a virtue. I will treat the different components of Kieran's definition separately for the sake of clarity.

Teleological

The teleological component is the ability to create novel and worthwhile ideas and artifacts. What is "worthwhile" is context-specific. Therefore we must reflect on the purpose of innovation and the values inherent to it to determine what the teleological component of creativity in innovation is. I begin by using some philosophical work on the purpose of innovation, supplementing and nuancing it with work from other fields.

Van den Hoven, Lokhorst and van de Poel have argued that the purpose of innovation should be to create the ability to fulfill an obligation which could previously not be fulfilled. The authors discuss problems of moral overload, which means that there are two duties which cannot both be fulfilled practically (Van den Hoven et al. 2012). For instance, a parent can face a problem of moral overload if a sleeping baby prevents him from attending a new neighbor's housewarming. On the one hand, a parent has a duty to respond if their child cries, on the other, he may have a friendship or community-related duty to attend the housewarming. A solution to this problem of moral overload is the baby phone. The baby phone allows the parent to fulfill both obligations, one of which could not be fulfilled without the baby phone (Van den Hoven et al. 2012). Innovations that make it possible for us to fulfill more moral duties than we otherwise could are definitely worthwhile.

However, requiring that an innovation solves a problem of moral overload seems overly restrictive. Many innovations are clearly creative and ingenious, even if they do not solve a moral problem.¹⁰ Definitions of innovation found in innovation studies tend to be more inclusive in terms of purpose, but more restrictive in terms of context. Here is one such definition: "Innovation is the multi-stage process whereby organizations transform ideas into new/improved products, service or processes, in order to advance, compete and differentiate themselves successfully in their marketplace (Baregheh et al. 2009)" This definition suggests that the purpose of

¹⁰ This point was raised by two anonymous reviewers.

innovation is to allow organizations to differentiate themselves in the marketplace. This can certainly be a purpose of innovation, much like solving moral problems, but it is not a defining feature of all successful innovation.

What seems to be a defining feature of successful innovation is usefulness. Usefulness seems to be an underlying assumption in both characterizations of innovation. Usefulness I take to mean that it allows an action which was previously not possible or not possible in a certain situation. The baby phone is a good innovation because it makes it possible for a parent to attend a party next door and be responsive towards his baby at the same time. The innovation is therefore useful. In order to differentiate an organization in the marketplace, an innovation must also be useful. Some may argue that to be novel and worthwhile an innovative object does not need to be useful. An object can also be a novel way to use a material or an idea never executed before. However, I would maintain that such a useless but original artifact can be novel and worthwhile, for its aesthetic value, inasmuch as it is an artistic object. A useless object which was meant to be a useful innovation is, however, a failure. An innovator who produces aesthetically novel and worthwhile objects, which solve no problems, lacks the mastery to be virtuously creative. The ideas or artifacts she or he produces also lack the worthwhileness to be creative as innovations. They may, of course, qualify as the result of artistic creativity, if they can be regarded as a work of art, and if she or he intended them that way.

If “worthwhile” in the context of innovation means “useful,” then readers are bound to raise the question “useful for what?” If someone innovates a new and useful way of torturing people, should this count as worthwhile? ¹¹ Many philosophers have discussed whether a virtue can be misguided and aim at the wrong goals. A Nazi’s actions can be made much worse if the Nazi is courageous. Does a courageous Nazi really possess the virtue of courage? Our position on this is that a virtue is still a virtue, with all the inherent value of a virtue, even if it is bad-making. Creativity in a torturer may make their actions much worse, just like courage in a Nazi can. However, the badness in both cases is not a result of the creativity or courage, but because of other vices of character, like cruelty, insensitivity or close-mindedness. Had the Nazi or the torturer not possessed these vices, then their character would have been good in virtue of their creativity or courage. I am aware that this is a contested claim and some virtue

¹¹ This point was raised by an anonymous reviewer.

theorists would disagree with us. However, this discussion goes outside the scope of this paper.¹²

Procedural

In addition to the creation of novel and useful innovations, to be creative, an innovator should possess enough mastery to know what she or he is doing in the relevant discipline of innovation. What level of competence counts as mastery should be evaluated according to appropriateness per case. An innovator doesn't need to know the precise details of how self-confidence works in people, down to the neurochemical level, in order to create an app which helps people develop self-confidence. Rather, the person should understand apps and how they are used, and have a sufficient understanding of self-confidence and the lack of it. However, in the case of a chemist creating new drugs for mood disorders, creativity should be accompanied by mastery which includes an understanding of mood at the neurochemical level.

Kieran discusses the role of “happy accidents” in the creative process of artists (Kieran 2014). These accidental discoveries can also happen in product and service innovation. In this context, they can also count as an exercise of creative mastery, as long as the agent is capable of recognizing the innovative potential of the happy accident and decides to act on this recognition. Safety glass was invented by a chemist who knocked over a flask that had contained nitrocellulose. The flask was still intact. The creation of the nitrocellulose coating inside the flask, and the dropping of the flask were not done intentionally to create safety glass. However, the chemist noticed the flask and saw its innovative potential, reverse engineered what had happened and eventually patented safety glass (Ewing 1932). The chemist's mastery of chemistry was crucial in him recognizing the fallen flask as a source of discovery.

Motivational

Now I come to the final component of creativity, intrinsic motivation. In the context of product and service innovation, this involves being motivated by the values inherent to innovation. Which values, then, are inherent to product and service innovation?

¹² For a discussion on the case of courage as a “bad-making” virtue, see pp. 15 Foot (1978) for an argument that courage is not always a virtue. See also pp. 92–93 of Zagzebski (1996) for an argument that a courageous Nazi in fact does possess the virtue of courage.

What distinguishes the values inherent to art from the values inherent to product and service innovation is that good products and services must be useful. As stated previously, product and service innovation has the telos of making an action possible which was previously not possible, or alternatively creating useful products. Van den Hoven et al. outline one more specific telos of high-tech innovation, making it possible to fulfill two duties in a situation of moral overload (Van den Hoven et al. 2012). A more business-oriented telos of innovation is differentiating oneself in the marketplace (Baregheh et al. 2009). Making useful products or services, ones that solve moral problems, or ones that have features that make them sell well can all be relevant sources of value for the domain of product and service innovation. An intrinsically motivated product or service innovator may therefore be motivated by the thought of making a good or successful product. Steve Jobs, former CEO of Apple, famously had precisely this kind of motivation (Isaacson 2013). Whereas someone who designs phones just because the salary is good, without being interested in the product itself, is likely not intrinsically motivated.

The values required for intrinsic motivation can be the values internal to the sub-discipline of innovation that the innovator is working with. An innovator may, for instance, be intrinsically motivated by the creation of more sustainable cars through a fascination with cars or through taking pleasure in the design of intricate machines. Since innovation is a broad label that covers many different types of activities, it follows that many different types of motivations can count as intrinsic to product and service innovation processes. Sand suggests that at the root of a driven innovator's character often lies "an appetite for change" (Sand 2018). However, an appetite for change by itself is not sufficient to constitute intrinsic motivation. The appetite must also be connected to the values inherent to product and service innovation. A burning appetite for a change in how one is viewed by peers is not connected to these values. An innovator motivated by such an appetite is not intrinsically motivated. If an innovator has an appetite for change in the sense that one wants to see people do previously impossible things, that is connected to one's desire to create a product that allows this, then such a person is intrinsically motivated for product and service innovation. This appetite must also be strong and motivating enough to drive the person to reliably exercise their creativity.

If we add the above reflections to Kieran's definition of a virtuously creative person, we can formulate a description of a virtuously creative innovator:

Someone who has acquired a certain degree of mastery and knows what she or he is doing in coming up with novel and worthwhile ideas or artifacts. Worthwhile innovations

are those that are useful, or solve some practical problem. In doing so, the virtuously creative innovator is motivated by the values internal to innovation. These values can be specific to the craft or sub-domain of innovation, or values inherent to innovation itself. Having an appetite for change qualifies as being motivated by values inherent to innovation, if she or he chooses what she or he does for reasons that hook up with those values in the right kind of ways.

This definition applies to individual creative innovators. But this individualistic definition does not describe the kind of creativity that is exhibited in Pixar's movie production process. As I outlined earlier, collective forms of creativity are important in many innovation practices. In a collaborative context like innovation, a definition of creativity which only applies to individuals is at risk of becoming redundant. Having characterized what it means to have the virtue of creativity in innovation, let us turn to the following question: Can groups possess the virtue of creativity in innovation?

3.4 Can Collectives Possess the Virtue of Creativity in Innovation?

It is not immediately obvious that the virtue of creativity in innovation can be a proper group virtue. Various arguments have been made for the existence of group virtues in general (Beggs 2003; Fricker 2010; Lahroodi 2007; Sandin 2007). These accounts have not treated very extensively the structure of specific virtues in collective form. The virtue of creativity as a collective virtue is yet unexplored by philosophers. Contemporary innovation, in the style of Google and Pixar, presents an excellent test-case for collective creativity (Hill 2014). In order to answer the question "Can collectives possess the virtue of creativity in innovation?" let us begin by investigating whether collectives can be agents that possess the teleological, procedural and motivational components of the virtue of creativity.

Teleological: The Collective is Coming up with Novel and Worthwhile Ideas or Artifacts: Worthwhile Innovations Are Those That Are Useful or Relieve, or Contribute to Relieving, Problems of Moral Overload

At first glance, it seems unproblematic that collectives can fulfill the teleological component of the virtue of creativity. Companies that solve practical problems by innovating are examples of collectives coming up with novel and worthwhile artifacts. However, one can still question whether talking about collectives inventing things is just short-hand for talking about the individuals inventing things that contribute to

the innovation in question. At the heart of this is the question: Can a collective be an agent?

While the exact nature of collective agency has been disputed for a long time in philosophical literature, many philosophers have argued that collective agency exists and can be meaningfully analyzed (Huebner 2016; List and Pettit 2011; Searle 1990). List and Pettit's influential account sets three conditions for collective agency, which they call "group agency". The authors argue that a random collective of people, such as a collection of people inside a supermarket at a given moment, do not, however, constitute agents. Collective agency only emerges when a collective has representational states, motivational states and the capacity to intervene in the environment, hence the term "group agency" rather than collective agency. These components of group agency are aggregations of representational and motivational states, and actions of individual group members. The group agency, however, supervenes the individual contributions of which it is aggregated. The fact that the actions of the group are ultimately caused by individuals is not seen as problematic for group agency (List and Pettit 2011).

Settling the debate about group agency, and various dependent concept, like group responsibility, is outside the scope of this article. I acknowledge that to those who find group agency and group responsibility unconvincing, the conception of collective virtue, and also collective creativity, will be unconvincing. It may be possible to argue that group virtues and a virtue of group creativity are possible without the need for group agency or group responsibility. However, that too is outside our scope. Therefore, I will make the assumption that groups can take actions which supervene upon the actions of individuals in the group, and for which the group, as a whole, is responsible. I also assume with List and Pettit that group responsibility can in some cases supervene upon the moral responsibility of individual group members, so that the group, rather than the members, are responsible for something. The question I attempt to settle here is not: "Can groups be agents and can they be held responsible as groups?" but rather "Given that groups can take actions that supervene upon individual actions, and reasonably be responsible for actions, in a way that supervenes the responsibility of individuals, can creativity as a virtue be the kind of thing that a group practices on this supervening level?"

If we accept the abovementioned assumptions and continue this line of questioning, then we may ask: "can a group come up with novel and worthwhile ideas or artifacts?" Intuitively, it seems that a group can do such things. The development of complex ideas or artifacts often requires the collaboration of many people. In fact,

policy-makers have recognized that finding solutions to many of the practical problems that society struggles with today requires collaboration between different disciplines and sectors of society (European Commission 2020; Mauser et al. 2013). The idea that addressing societal challenges requires a diversity of parties already implies that collaboration is needed for innovation to be worthwhile. If this claim is true, then groups can have capabilities superior to individuals with regard to the teleological component. Groups are in some cases better capable of creating novel and worthwhile innovations. If groups are in some cases capable of things that individual members are not capable of, as in the case of innovation, there is good reason to accept that groups can come up with novel and worthwhile ideas or artifacts. The teleological component of the virtue of creativity can be fulfilled.

Let us, therefore, turn to the procedural component.

Procedural: The Group Has Acquired a Certain Degree of Mastery and Knows What It Is Doing

Possessing the procedural component would mean that the group has mastery over the craft at hand. The idea that groups can master something is relatively common in colloquial speech. People often talk about a football team or a band as being excellent at playing football or music. But can we truly speak of a group mastering something?

Let us begin with whether groups can acquire mastery. If a group agent, such as an institution, exists over a longer period of time, then it can also develop capacities. The ability to collaborate efficiently is often, if not always, an acquired capacity of groups. Groups often need time to function optimally. This is a fairly familiar experience from most people's working life, or personal relationships. While innovating requires various capacities, group work, or innovating with others is in itself a capacity that is acquired through experience.

And does this acquired mastery resemble individual creative mastery? I believe that both intentional creation and judgements regarding happy accidents can be made by collectives of people. A group like a company or a research team can have a shared intention. List and Pettit would call such a group a "jointly intentional group agent" (List and Pettit 2011). The group can carry out its intentions in different ways. For instance, a software company can have the intention to make an intuitively designed app with a certain purpose. The company can employ a designer, who is able to carry out the intention of the intuitive design. The writing of the app is carried out by those with skills in development and a manager can ensure that the app fulfills its purpose

effectively. The group's joint intention is brought about by mastery distributed over different people.

One could argue that such instances of group mastery can be reduced to individual mastery. Development processes for new products and services are complex and require many people. But does the "group mastery" not simply boil down to individuals possessing mastery over their own contributions, making "group mastery" just short-hand for collection of individual mastery? I would argue that in some cases, this is indeed the case. However, there are special types of mastery which truly supervene upon the mastery of individuals.

The kinds of truly supervenient mastery are described well by Palermos and Tollefsen's conception of "group know-how." The authors argue that the difference between truly supervenient "group know-how" and a mere collection of individual capabilities is the interactions between those activities. "When individual members coordinate on the basis of reciprocal interactions, they adapt mutually to each other by restricting their actions in such a way so as to reliably—that is, regularly—achieve ends that they would only luckily, if ever, bring about were they to act on their own" (S. Orestis Palermos and Tollefsen 2018).

Collective creativity of the sort that Pixar and Google show seems to involve exactly this kind of know-how. The routine of dialogues, discussions and conflicts that Pixar employees are encouraged to engage in requires coordination and each individual action within the routine incites a response from others. The routine of lateral communication and sharing that Google employees engage is also coordinated to be such that it yields results which individuals could not achieve.

One group member can cause a "happy accident" and it may take a discussion between two other group members to recognize it as a possible good idea, while yet another pair is able to reverse engineer the happy accident, in response to the recognition that it may be a good idea. In group agents, control and mastery can be distributed over different members and united by the joint intention of the group and the joint nature of the capabilities of the group. It seems then, that the conception of group mastery describes well what goes on in many innovation practices.

Motivational: The Group Is Motivated by the Values Internal to Innovation. These Values Can Be Specific to the Craft or Sub-domain of Innovation, or Values Inherent to Innovation Itself. Having an Appetite for Change Qualifies as Being Motivated by Values Inherent to

Innovation. It Chooses What It Does for Reasons That Hook up with Those Values in the Right Kind of Ways

In contrast to the requirement of the teleological component, intrinsic motivation, seems, at first sight, to be the most difficult when applied to groups. Even if we accept that a group can possess “group know-how” or collective mastery, can it possess something like a supervenient group motivation? And if so, is it possible for a group be intrinsically motivated to do something?

The focus here is the question “Is it possible for a group to possess a supervenient group motivation?” I will propose an account of how this may be possible, using Miranda Fricker’s work regarding collective motivations. Fricker constructs an account of collective motivations to argue for the existence of institutional virtues (Fricker 2010). I will use insights from her work to argue for the possibility of a group analogue of intrinsic motivation. I do not wish to adopt her exact view of institutional virtue but merely to adopt concepts and argumentation she uses. Showing that groups can possess this analogue of intrinsic motivation will provide evidence for the motivational component needed for the existence of creativity as a collective virtue.

Fricker introduces the concept of a *practical identity* which members of an institution share and which can prompt certain virtuous or vicious actions. She introduces the concept in order to argue for the possibility of institutions possessing virtues, even in cases where not all members of the institution are virtuous individually. She discusses the virtue of fairmindedness in the case of a jury. I will use her conception of a practical identity to show that it can serve as a basis for group motivation. I shall first explain Fricker’s use of the conception.

With regard to the possibility of institutions having virtue, Fricker outlines three views: the summativist view, the collectivist view and an in-between view, which she argues for. The summativist view sees collective virtues as a summation of all the individual virtues within a group. Within this view, an institution being virtuous would mean that all members are individually virtuous. Fricker rejects the summativist view for the reason that it is too demanding. We should still be able to call a jury fair-minded if it behaves fair-mindedly, even if there is one un-fair-minded person in the jury (Fricker 2010)

The collectivist view states that if a group functions reliably in such a way that it manifests the virtue, then it possesses the virtue. According to this view a jury which is diversely biased, and comes to a fair-minded conclusion due to the cancelling out of biases, can be called fair-minded. Fricker also does not agree with the collectivist view, because she does not see a possibility of there being a virtuous motivation present in

such a group. While the summativist view is too demanding, the collectivist view, according to Fricker is too lenient (Fricker 2010) Thus, she argues for an in-between view: a form of summativism with the possibility of collective motivations.

This in-between view is that collective virtues can be the summation of individual virtues and the practical commitments that influence virtuous group behavior. Fricker defends this view. According to Fricker, we cannot speak about a jury being virtuously fair-minded if the only reason it acts fairmindedly is that members' biases cancel each other out. However, in a jury where a few people are fair-minded, and others, though not fair-minded, have a willingness to go along with the fair-minded ones, out of a sentiment of professional loyalty, for instance, we can meaningfully speak of an institutional virtue. The institutional virtue consists of the virtues of some individuals and the commitments of the others to go along with their virtuous decisions (Fricker 2010). Fricker's view is that if enough people in the group have a virtuous motivation or are willing to "go along with" a virtuous motivation, or take a virtuous practical identity, then a group virtue can be ascribed to the collective.

Practical identity refers to the identity that a person adopts as a part of their commitment to membership in a group or institution. The distinction one makes between speaking personally and speaking as a member of a group, or from a particular role, is the distinction between personal and practical identity. Practical identities can be significantly different from personal identity, also in morally relevant ways. Lahroodi discusses the following example. A church committee consists of members that are individually open-minded about gay rights. However, when they get together as the church committee, they become close-minded. They might do this because they feel it is their duty to take the close-minded *view of the church* within their practical identity as church committee members (Lahroodi 2007). Similarly, group members may take on a practical identity which is morally superior to their personal attitudes. In such cases, Fricker argues, the practical identity can provide a substitute for virtuous motivations of individuals. If a practical identity containing a virtuous motivation is taken on by members, then it is fair to attribute a virtue to such a group or institution (Fricker 2010).

The question I am interested in is: Can intrinsic motivation be a practical identity and hence function as the motivational component of the virtue of creativity in groups? I suggest that it can because intrinsic motivation can be a part of a practical identity. People working at Pixar, who might not themselves be virtuously motivated, can "go along with" the intrinsic motivation of others in various ways. One may help uphold a culture of enhancing intrinsic motivation. One may recognize the intrinsic

motivation of others and put measures in place to keep it up, even while being extrinsically motivated oneself. Google is notorious for fostering a culture that is conducive to innovation by maintaining an open and reflexive culture and rewarding “Googley” behavior (Steiber and Alänge 2013). One may also “go along with” a creative idea for the sake of exploration even though one does not necessarily delight in that activity, but knows that it is important to explore creatively. An employee of Google may go along with certain decisions, because doing so would be “Googley.” If a group takes on such a practical identity, then, just as in the case of a fair-minded jury, a creative company can include people with no virtuous motivation but a practical identity which favors creativity and creative activity for its own sake. This way, a group can have an intrinsic motivation without any individual even having one. The nature of this practical identity is inherently social. Again, possessing the virtue requires besides motivation also the other two components, procedural and teleological. If a group intrinsic motivation, as described here, is combined with supervenient socially construed and coordinated mastery over the process of coming up with novel and worthwhile ideas, then the group possesses a group virtue of creativity which cannot be reduced to traits of its individuals.

3.5 Components of Creativity as a Virtue

	Teleological component	Procedural component	Motivational component
A creative person (Based on Kieran, 2018)	Comes up with novel and worthwhile artifacts	Has acquired a certain degree of mastery knows what she or he is doing in coming up with novel and worthwhile ideas or artefacts	In doing so, she or he is motivated by the values internal to the relevant domain and chooses what she or he does for reasons that hook up with those values in the right kind of ways
A creative individual innovator	Coming up with novel ideas or artifacts that are useful or relieve, or contribute to relieving, problems of moral overload	Has acquired a certain degree of mastery in innovation or a relevant sub-field of innovation, and hence knows what she or he is doing	Is motivated by the values internal to innovation. These values can be specific to the craft or sub-domain of innovation, or values inherent to innovation itself
A creative innovation collective	Engages in collective ideation towards ideas or artifacts that are useful or relieve, or contribute to relieving, problems of moral overload	Has acquired mastery in harnessing different skills distributed over members and a capacity to collaborate	Is driven by aggregated intrinsic motivation of members for innovation or various sub-domains of innovation and/or a shared practical identity which upholds and stimulates intrinsic motivation

3.6 Implications of Creativity as a Collective Virtue

Collective creativity is a much-discussed topic in organizational psychology, design and management literature. Collective creativity is something which innovation groups can and should develop and foster (Hill 2014). Considering the amount of attention in other fields that deal with innovation, surprisingly little philosophical work has been devoted to the moral value of collective creativity. In this article, we have analyzed the possibility of regarding collective creativity as a collective virtue.

Our approach has been to dissect Kieran's definition of creativity as an individual virtue into three components, teleological, procedural and motivational, and see whether each can be applied to a collective. One may contest this approach and argue that collective creativity should not be investigated with an individualistic definition as a starting point. I do not claim that the originally individualistic definition of creativity is the only correct one and that more research into virtuous group creativity would not yield a better definition. Showing that Kieran's definition can be applied to groups merely demonstrates that it makes sense, philosophically, to attribute a virtue like creativity to groups. This can likely be done with other virtues too.

If one accepts that group creativity is a virtue, then one accepts that it is *prima facie* valuable. This has substantive consequences for various ethical decisions in innovation. Firstly, it means that organizations should facilitate collective creativity and that it is worth pursuing in its own right. Collective creativity may come into conflict with other values. For instance, facilitating an inclusive process can take more time than developing a much-needed product quickly. If collective creativity is recognized as a virtue, and thus inherently good, then abstaining from bringing it about should be seen as a sacrifice. The nature of the sacrifice would be giving up on developing something that is inherently good. I intend our account of group creativity as a virtue to be an addition to the existing literature on the virtues in innovation, the existing literature being largely individualistic (Brenkert 2009; Sand 2018; Steen 2013). Accounting for the good of creativity in innovation as an individual virtue only, and not as I suggest, as a collective virtue, also has some limitations.

Secondly, a more practical consequence of recognizing collective creativity as a virtue is that it may provide new reasons for adopting more collective types of intellectual property frameworks, or crediting practices more broadly. If a group can be credited with the virtue of creativity, then perhaps this means that collectively created products should be credited to collectives, rather than compartmentalizing and crediting individual contributions.

There is also a normative argument to be made for why collective creativity should be recognized as a virtue. Only recognizing individual creativity incentivizes a perverse obsession with individual visionaries and creative geniuses. This, one could argue, can lead to a competitive atmosphere which discourages cooperation and collective or community-oriented thinking. Accepting that collective creativity is just as excellent and inherently valuable as individual creativity is likely to motivate the creation of more collaborative and less competitive environment for innovation to happen.

Of course, we should not think that innovations used for torturing people are good if the process of innovating them was creative. However, we should also not ignore the fact that a huge part of the population is in some way involved in product or service innovation, and the meaningfulness of one's work is a big part of one's wellbeing. Innovation is not merely a means to an end, but a valuable human activity in itself. Even with the risk of justifying foolhardy innovators conduct, the value of innovating as an activity should be taken seriously.

Additionally, some organizational psychologists would argue that it is simply not true that innovation is always driven by creative geniuses (Hill 2014). An individualist virtue approach to creativity is likely to emphasize the contribution of gifted individuals even in cases where the contribution of the group played an essential role. Additionally, there is often no single genius behind every good innovation. In fact, scholars like Hill would even argue there should not be (Hill 2014). Co-creation and participatory innovation processes are often praised because they allow different parties contribute with their ideas and experience, not to realize the vision of one or two geniuses. The co-creation process taps into the collective genius of a diverse group (Sanders and Stappers 2008). Therefore, in many cases, an individualistic account of creativity as a virtue in innovation is simply not an accurate description of what happens in an innovation practice. To be empirically accurate, a virtue approach to innovation should recognize collective creativity. I have shown that it is, at least whilst making some assumptions, possible.

4 Can Responsibility be a Collective Virtue?

4.1 Introduction

In the previous chapter I illustrated an application of collective virtue theory. Collective virtue theory, as I have argued, can make sense of the excellence of companies and organizations with a long-term and stable tendency to innovate creatively. Such groups possess the collective virtue of creativity. Collective creativity is not the only collective virtue. In this chapter I make a case for another important virtue, namely responsibility. Collective responsibility as a virtue is a useful concept for making sense of collective action problems. It can also be a useful contribution to philosophical questions about collective responsibility in general. This is because traditional notions of collective responsibility suffer from a serious flaw, which is their inadequacy in dealing with “problems of many hands.” To begin, let us consider the concept of collective responsibility.

We often praise and blame groups of people, just like we praise and blame individual persons. Blaming Shell for climate change or a government for war is often as natural as blaming a friend for betrayal (D. P. Tollefsen 2003). It makes sense that we blame groups. After all, wars, environmental degradation and pandemics are never one person’s fault. Attributing responsibility to collectives is also often effective in driving change. Because being held responsible is often what drives us to fix the problems we have caused and to demand that others do so. For instance, the oil industry is responsible for much environmental



degradation. This is why we may easily feel comfortable demanding that the oil industry take the largest steps to protect the environment. Group responsibility justifies globally important group action. But what does it really mean to blame a group?

Philosophers have long discussed what it means for a group to be morally responsible for something. Does it boil down to individuals being responsible for their individual contributions? Or is there something like blaming a group, *as a group*, which is different from blaming individuals? Many theorists agree that group responsibility does not always boil down to individual responsibilities (Björnsson 2020; M. P. Gilbert 2002; Giubilini and Levy 2018; List and Pettit 2011; Smiley 2008). Groups can be held responsible but only under certain conditions. Generally, what these conditions are depends on which aspects of the group are seen as supervening, or in some way “over and above” the individual members (Giubilini and Levy 2018). Characteristics that are “over and above” other characteristics and often referred to as supervening. The supervenience relation is described in the following way by Lewis “We have supervenience when there could be no difference of one sort without differences of another sort” (p.14 Lewis, 1986). The behavior of a group changes when the behavior of individuals changes, because the group supervenes over the individual members.

Sometimes supervening properties are also emergent, meaning that the supervening properties cannot be explained solely with reference to the individual properties. Just like the elaborate shapes created in the sky by a flock of birds, many collective human behaviors emerge from many individual actions. These emergent phenomena however have their own characteristics, which are best described at the level of the collective. Some philosophers have argued that a group’s actions supervene on the actions of members (Bratman 2014). Others argue that a group’s intentions or beliefs do so (M. Gilbert 1987; Searle 1990). The supervening qualities are what give the group a kind of unity. If this unified part of the group fulfills the conditions of moral responsibility, knowledge, control and moral competence, then it seems intuitive that the group can be held responsible as a group. This means that when one is praising or blaming a group for collective and emergent behavior, then one is praising or blaming the collective as a whole, not the individuals. Philosophers disagree on whether this praise or blame is always shared by individual members too.

Collective responsibility is an important concept to have. Many times, when for example a big corporate scandal is revealed, it is not clear which individuals were responsible for the transgressions. Often, each member may have played a small part, partaking only in a minor transgression or no transgression at all. In such cases, we

must blame the collective rather than individuals for the crime. However, exactly these kinds of groups tend to be internally fragmented in the kind of way that renders them incapable of responsibility. Such cases, often referred to as a “problem of many hands,” is a situation where a collective has done something wrong, but no individuals can be blamed. Usually, this is seen as a problem because it means we cannot hold individual responsible. I argue that in such cases, it also becomes difficult to hold the collective responsible.

Should we do away with the idea of collective responsibility completely in such cases? I think that would be premature. I propose instead that we adopt another way to see collective responsibility, which does not run into this problem. I suggest we abandon the idea of collective responsibility as a relationship between the collective agent and the action in question. Instead, we should see collective responsibility as a virtue. A virtue is a long-term and stable positive character trait of an agent, in this case, a group. If we see responsibility as a virtue possessed by a collective then it becomes easier to attribute moral praise and blame to groups. Instead of thinking “is this collective responsible for x” let us think “is this a responsible collective? Or is it irresponsible?” Changing our conception of what responsibility means makes it possible to praise and blame groups for their actions, even if the group is fragmented, like in the case of an example I treat in section 3, the Deepwater Horizon disaster.

How can a group be responsible or irresponsible? To answer this, I use philosophical literature to show that groups can have a character, much like individuals. A group can be said to be, just like any of us, benevolent, foolhardy, responsible or irresponsible. This means that we can distinguish virtues and vices in the character of groups. According to virtue ethics, vices (bad character traits like foolhardiness or irresponsibility) are inherently blameworthy and virtues (good character traits like benevolence and responsibility) praiseworthy. Virtue ethics can, therefore, give guidance in praising and blaming groups. The guidance of virtue ethics, I will argue, is better than the guidance given by the more traditional philosophical concept of collective responsibility.

What then does a virtuously responsible collective look like? My answer to that is derived from the negative opposite of collective responsibility, which is the vice of collective irresponsibility, of which the problem of many hands is expressive. I outline some causes of the problem of many hands and show how these traits can be categorized as collective irresponsibility as a vice. Collective responsibility is the long-term and stable acquired trait of collectives to avoid these pitfalls.

Section 2 outlines what is needed for a collective to be held collectively responsible for something. Section 3 explains the problem of many hands. In section 4, I show how the notion of collective responsibility runs into problems in cases of problems of many hands, and why an additional way of holding groups responsible is needed. In section 5, I propose an alternative way of holding collectives responsible, based on the idea of groups having a character, which can be virtuous or vicious. The concluding remark reflects on the significance of collective responsibility as a virtue.

4.2 Collective Moral Responsibility

Let us begin with the question: what is collective moral responsibility? Two questions are important for understanding collective responsibility: “what is *collective* about collective responsibility?” and “How is collective responsibility *moral responsibility*?” Let us briefly consider each.

As for the first question, Giubilini and Levy (2018) have provided a coherent answer that summarizes the views of most collective responsibility scholars. Imagine a group of schoolchildren who are directed by their teacher to each draw themselves on one large shared piece of paper. Each enthusiastically picks out colored pencils and soon becomes absorbed in drawing, mumbling to themselves and blissfully ignoring everyone else. Of course, the children are creating something together, namely a collectively drawn class picture. However, if we follow Giubilini and Levy, the pupils are not collectively responsible for the drawing. Rather, the responsibility for the drawing is better explained by saying that the teacher is responsible for the idea and each pupil for their individual drawing. However, if the teacher asked a group of pupils to write a report together, then this would force the pupils to make agreements, form shared intentions and devise their contributions based on what others are contributing. Such an assignment would probably qualify as something the pupils are collectively responsible for.

As for the second question, “How is collective responsibility *moral responsibility*?” let us take a closer look at what is required for moral responsibility in general. One comprehensive view, which serves our purposes here, and captures well the most dominant intuitions about moral responsibility, is Fischer and Ravizza’s account.

Fischer and Ravizza’s theory of moral responsibility focuses on the internal mechanism that brings about a certain action. For example: If a person knocks over a glass because of a seizure, then the mechanism that brings about the action is not reasons-responsive. A seizure cannot be reasoned with, it just happens. However, if a

person knocks over the glass because they are angry, then this mechanism can be reasons-responsive. Because they may refrain from knocking the glass over if they are told that their behavior is scaring their friends. In the case of a seizure, the person is not morally responsible for knocking the glass over, in the latter case they are. The authors also specify an epistemic condition for responsibility, meaning that to be responsible one must have knowledge of what one is doing. For instance, if the person knocking the glass could not have known that it was glass, not plastic, and would break when knocked, then they would not be morally responsible for the glass breaking, although they would be responsible for knocking it over (Fischer and Ravizza 1998 (The example is my own)). The mechanism by which the person acts must also be “the agent’s own” meaning that the agent: “(a) sees herself as the source of her behavior (which follows from the operation of K [the mechanism]); and (b) believes that she is an apt candidate for the reactive attitudes as a result of how she exercises her agency in certain contexts; and (c) views herself as an agent with respect to (a) - (b) based on her evidence for these beliefs. (Fischer and Ravizza 1998)”

We can distill three important components of moral responsibility from Fischer and Ravizza’s account, which also align with other accounts of moral responsibility. Therefore, for a collective to be morally responsible for an action or event, it must be the case that:

- The collective agent has knowledge of what it is doing.¹³
- The collective agent’s mechanism of control over the action is reasons responsive.
- The mechanism of control over the action is the collective agent’s own in the sense that the agent believes to be in control of it.¹⁴

Armed with this definition of collective responsibility, let us move to what the problem is with applying this definition.

4.3 The Problem of Many Hands

The problem of many hands, henceforth PMH, is a morally problematic situation which arises when a collective is morally responsible for something but no individual in that collective can reasonably be held responsible (Poel et al. 2015). For instance, take the Deepwater Horizon disaster, the oil spill that happened in 2010 in the Gulf

¹³ For a discussion on how a group can know things, see Smith (1982). For a discussion of mutual know-how, see S. Miller (2020).

¹⁴ For a discussion on how a group can hold joint beliefs, see M. Gilbert (1987).

of Mexico. This is a case that has been described as a PMH (pp. 1-11, Poel et al. 2015). The part of the oil rig that extended into the earth under the sea was not sealed off properly. No one intervened early enough because of various epistemic failures. For instance, British Petroleum's internal tests had shown the cement to be unstable. However, this was not communicated to the contractors implementing the cement seal. The site leaders at the oil rig even dismissed obvious signs of leakage. This was because they accepted a theory held by other oil rig staff, about the signs of leakage being caused by something harmless. The oil rig staff had not been trained properly. The necessary information did not reach the right people, because of a lack of necessary information, protocol and training at other places (Poel et al. 2015).

The PMH is problematic because it seems to cause a responsibility gap (Poel et al. 2015). A responsibility gap emerges when there is an event for which someone should be responsible, like an oil rig disaster, but no one can be held responsible. One could argue that each member of the collective that caused the disaster can be attributed a small amount of the responsibility, which all adds up. However, distributing responsibility in a piecemeal fashion still causes a responsibility gap, because a small piece of responsibility for "not noticing that one's boss was not properly trained" or "not taking great efforts to find out what was wrong" is something very different than the heavy responsibility of a disaster involving deaths and pollution (Poel et al. 2015). The heavy responsibility, for which we feel someone should be blamed, disappears, unless we attribute it meaningfully to the collective as a whole.

Groups with PMH can however also not be attributed collective responsibility. To illustrate why, let us consider the definition of collective responsibility from the previous section. To be collectively responsible, a group must fulfill all three conditions of moral responsibility: epistemic condition, control condition and moral competence condition.

The epistemic condition is: "The *collective* has *collective* knowledge of what it is doing." Think again of the Deepwater Horizon disaster. Information which could have been knowledge about the impending disaster was distributed in such a way that it was merely distributed information. No common knowledge of the danger was present until after the disaster. Therefore, although the necessary information may be distributed among members, the possibility to form the knowledge is not there in that state. The fact that crucial information is not distributed properly is generally a sign that the group suffers from an underlying collective behavior problem, like a lack of transparency, secretive culture, or insufficient collective attention to signaling problems. In such a state, it is group-psychologically impossible for the information

distributed among group members to be used in a functional way to improve the actions of the group. Therefore, there is no collective knowledge, meaning that the collective agent does not meet the collective knowledge condition for moral responsibility. We also cannot resort to merely holding the individual members responsible for their lack of knowledge. In such cases, the individuals *do not know that they should know* something. Therefore, they do not fulfill the knowledge condition for individual responsibility either. In this case, the group not fulfilling the epistemic condition is what led to the PMH and the tragedy.

4.4 Collective Responsibility as a Concept Must be Revised

Problems of many hands, make it impossible to attribute a responsibility properly to individuals. They also make it impossible to attribute collective responsibility. We may be tempted to bite the bullet and admit that the concept of collective responsibility does not apply to cases like climate change or engineering disasters, which involve PMHs. The alternative is to call these cases a combination of some individual fault, and bad luck. Engineering disasters and climate change would then be in the category of unfortunate events, like natural disasters ¹⁵. Then we must argue that only epistemically and structurally well-organized groups can be held responsible for their mistakes.

However, there are good reasons not to bite the bullet. Climate change, corporate scandals and engineering disasters often happen precisely because a group is not a well-functioning collective agent. If we accept that all such cases are nobody's responsibility then we may lose the motivating force of responsibility attributions which is practically very important for remedying such situations. If a theory of collective responsibility fails to attribute collective responsibility in a case like the Deepwater Horizon disaster, then it seems intuitive that we should use another theory to justify the practice of attributing collective responsibility.

This does not mean that we must eschew the notion of collective responsibility completely. It works in many collective responsibility cases, after all. However, in cases of PMH, collective responsibility should be attributed differently. Because the

¹⁵ This would be different from what van de Poel et al argue in their work on PMH. They argue that in a case of PMH, the collective is morally responsible, but individual responsibility attributions are impossible to make (Poel et al. 2015). This is where I disagree with the authors, because I argue that a group with a PMH cannot be collectively responsible either.

traditional way of attributing it is not metaphysically sound for that particular situation.

One way to avoid biting the bullet might be to try to find ways of attributing individual responsibilities to members of the group. Fahlquist has suggested such an account (Nihlen Fahlquist 2015). She re-conceptualizes responsibility by viewing it as a virtue, meaning an acquired, stable, long-term disposition of actively taking responsibility. A person with this virtue is a *responsible person*. If each individual in an organization takes the responsibility upon themselves, even when they are not assigned it, then this can fill responsibility gaps that may have been left unaddressed by job descriptions and responsibility divisions. What is important here is that taking responsibility means that one makes it so that one can reasonably be held responsible for something [Fahlquist, 2015; Poel et al. (2015)]. A responsible person involves themselves with issues in such a way that others can hold them responsible for it, thereby eliminating responsibility gaps [Fahlquist, 2015; Poel et al. (2015)].

This approach, however, cannot escape the limitation that it needs to break apart a large responsibility into small pieces, which means that the large responsibility, being of a different nature, disappears. If a group causes deaths, but each member has only contributed to it in a minimal way for which they are hardly blameworthy, then the piecemeal way of distributing responsibility fails in a significant sense. For this reason, I choose to further the concept of collective responsibility, rather than individual responsibility.

My proposal is a collective version of Fahlquist's. I propose seeing responsibility as a virtue, which is *the disposition to make oneself fulfill the conditions for moral responsibility* for important matters. Collective responsibility, therefore, refers to a desirable characteristic of groups. Instead of asking whether a group is responsible for something, we can ask whether a group is *responsible*. This provides an alternative possibility for philosophically justifying the praising and blaming of groups.

4.5 Collective Responsibility as a Virtue

Groups that regularly meet and undertake joint actions, can be said to have a character. And group characters can contain positive and negative traits. Just like individual character, group character can be developed to become virtuous or vicious. Instead of asking the question "Is Coca-Cola responsible for sugar consumption-related health problems?" we should ask a different question "Is Coca-Cola an irresponsible company?" This kind of thinking is not new in philosophy. Many

philosophers have written about collective virtues, or what it means for a group to be, for example, clever, open-minded or benevolent (Astola 2021a; Byerly and Byerly 2016; Fricker 2010; Lahroodi 2007; Palermos and Pritchard 2013; Palermos 2020; Pritchard and Palermos 2016).

My proposal is to create a collectivist version of Fahlquist's argument. We should attribute collective responsibility as a virtue to collectives that act responsibly. Responsibility as a virtue has so far only been characterized as a trait of individual people (Williams 2008). However, as various authors have argued, virtues and vices can also be possessed by collective agents (Astola 2021a; Beggs 2003; Fricker 2010; Lahroodi 2007; Sandin 2007). I will argue for a view of collective responsibility which includes the following claims, which I will treat in subsequent subsections:

- A responsible group has a responsible identity, constituted by the practical identities of members
- The problem of many hands is generally expressive of the collective vice of irresponsibility, which is the inverse of responsibility as a collective virtue
- Collective responsibility as a virtue is the long-term and stable, acquired trait of a group to make itself meet the three conditions for collective moral responsibility

A Responsible Group has a Responsible Identity, Constituted by the Practical Identities of Members

Collective responsibility must attach to a mechanism of action that is truly collective and does not boil down to actions of individuals (Giubilini and Levy 2018). What then is this mechanism? I believe Fricker's work on institutional virtue provides an answer. Fricker argues that group membership often entails the development of a *practical identity*. If one becomes a member of a company or an association, then this means that one must sometimes speak and act *as a member of that group*. These practical identities are mutually dependent, because their existence in members depends on the commitment of other members to similar practical identities. When one acts in their capacity as a member, they sometimes act in ways that are different from how they would act if they were not in that capacity. According to Fricker, this explains why people sometimes act against their own views in professional, or other membership-related, contexts (Fricker 2010).

When the streets surrounding a fast-food restaurant are littered with straws and burger-wraps, we tend to hold the customers individually, not collectively, responsible. They are not, after all, organizing and coordinating their littering activities. We might

however hold the fast-food restaurant (perhaps even the franchise) as collectively responsible in their part in the littering. The restaurant and the chain coordinate their activities so that littering may be reduced, for example by handing out less packaging, providing more bins and reminding customers not to litter. They are, therefore, collectively responsible, because they are more organized and have a unifying group identity consisting of the practical identities of members.

What operationalizes this possibility for coordination, are the mutually dependent practical identities. If the fast-food staff decide that they must take action on this problem, this stance will become a part of every employee's practical identity as an employee of that restaurant. And the mutually dependent practical identities are the object of the responsibility attribution. When groups have practical identities, they also become candidates for blameworthiness if they are not responsible collectively. Because adopting practical identities means making a certain promise, to act in a certain way to internal or external parties.

It should be noted that if we see groups as entities that develop character, we must not judge the character of groups that have had no time to develop their character. If a violent mob bands together randomly in the span of an hour, we should attribute responsibility for the actions of the mob to individual members rather than the mob as a collective. The group is nobody yet. However, groups that exist over a long period of time and establish their own forms of operating, like organizations, can be held responsible as groups.

The Problem of Many Hands as Indicative of the Collective Vice of Irresponsibility

The negative opposite of a virtue is a vice. It will be useful therefore to add some detail to our conceptualization of responsibility as a collective virtue by describing what its opposite is. If we think about collective responsibility as a virtue, we may think of its negative opposite as being collective irresponsibility. The problem of many hands is expressive of exactly that.¹⁶

PMHs should not be seen as mere events of instances where no responsible persons can be found. It is not merely by the unlucky circumstance that the Deepwater Horizon disaster occurs without a clear responsible person. The lack of responsibility was already there for a long time before the disaster. The procedures, attitudes and working practices within the collective were conducive to a lack of transparency,

¹⁶ For a discussion on collective vices, see Fricker (2010) and Baird and Calvard (2019).

internal communication and perhaps also a lack of care for safety. The PMH was a symptom of this collective irresponsibility. The PMH developed and worsened over a long period of time, and the disaster in turn was a tragic symptom of it. PMHs are then indicative of collective irresponsibility. I propose to treat collective irresponsibility as a collective vice, which often leads to PMHs.

PMHs may not always be a symptom of underlying collective irresponsibility. It is conceivable that a virtuously responsible group fails to take responsibility appropriately for some given object in a specific situation, without doing so structurally. However, in such cases, as soon as the failure is signaled, the group will be equipped and capable of carrying out the correct response. It is in cases where a PMH is signaled but the group fails to take appropriate actions that one sees proof of an underlying collective vice of irresponsibility. An irresponsible group has a character that does not allow them to take responsibility, in the same way that an irresponsible person fails to take responsibility. In both cases, the appropriate response is to blame the agent for having developed an irresponsible character.

A vice is by definition blameworthy. But a collective irresponsibility is not always blameworthy. For group irresponsibility to be blameworthy, it is necessary that the group occupy a role where collectivization is required of them. Not every group needs to be a collective agent. Some groups, however, like companies, governments or other organizations whose members adopt practical identities related to their group membership should be considered collective agents that can have praiseworthy or blameworthy traits. Adherence to law or protocol and prevention of harm to society are generally tasks that require organizations to perform joint actions. These are the kinds of collectives for whom collective irresponsibility is a vice.

The collective vice of irresponsibility, which expresses itself in the PMH, is realizable in multiple ways. There are three distinctive ways in which a PMH can occur, the epistemic, structural and motivational. Each of these categories shows a particular trait, all of which can be categorized as collective irresponsibility, because they culminate in a situation where the collective agent does not meet the three criteria for collective moral responsibility. This generally means that the collective agent generally does not have the disposition to make itself meet the conditions for moral responsibility or even the active disposition to avoid making itself meet the criteria for responsibility. When this tendency is negligent enough, we can speak about a collective vice of irresponsibility. Let us discuss each type of irresponsibility in turn.

Epistemic irresponsibility

Many stories exist of organizations compromising their ability to fulfill the epistemic condition. A group may fall into an epistemic problem of many hands if it begins to exercise collective habits that lead it away from transparency and information sharing, which prevents the collective agent from forming collective knowledge about its responsibilities. Just as an individual can adopt bad habits that compromise their epistemic agency and make them forgetful or ill-informed, a group can adopt habits that similarly compromise their epistemic agency. As in the Deepwater Horizon disaster, where knowledge was not spread efficiently enough for necessary conclusions to be drawn. Such a group fails to be virtuously responsible, because it operates in a state in which it cannot meet the epistemic condition for collective responsibility.

Structural irresponsibility

A PMH can also arise as a result of structural issues. Dysfunctional bureaucracies are an example of this. Inefficient policies or a lack of policies can however also constitute structural collective irresponsibility. Many PMH that have been investigated thoroughly show this kind of trajectory. For instance, the Volkswagen Diesel scandal involved probably a few hundred employees from Volkswagen. This group of people had concealed a portion of the emissions from Volkswagen vehicles. This was done by creating a test-rig setting for the automobiles that would give much lower than accurate emission readings when tested for emissions (p.219 -220 Davis 2020).

We might say that everyone involved in the creation and implementation of the cheat code was individually responsible for their misconduct. To what extent this holds for the hundreds of people involved is unknown. However, the problem of many hands here rests on the rest of the organization not inquiring, signaling and intervening in this illegal practice. It is, after all, quite unusual that hundreds of employees in an organization take part in an illegal and fraudulent activity without anyone intervening. One analysis of why this happened is because of the lack of preparedness for misconduct. There were no internal processes that would allow for whistleblowing. And because of the lack of organizational structure and attention to whistleblowing, no whistles were blown by employees, perhaps out of fear of social sanction (Davis 2020). VW had a long-standing collective habit of trusting upper management too much and not considering the possibility of misconduct. This eventually created structures which did not support responsible behavior within the

organization. This makes it difficult for the company to meet the control condition for moral responsibility.

Motivational irresponsibility

A problem of many hands can occur due to motivational traits. As experience often shows, demotivation is contagious, and it is hard to stay motivated to do a task in groups where this contagion has already spread. One example of a motivational PMH is “innovation trauma” in organizations. Companies that experience backlash following an innovation, or an attempt to innovate, sometimes develop a pessimistic company culture that is wary of trying anything new. Innovation trauma has been defined in management literature as “the inability to commit to a new innovation due to severe disappointment from previous innovation failures” (Välikangas et al. 2009). The demotivation caused by a previous failure easily spreads even to those employees who did not experience the traumatic failure. And the pessimism persists even if there is a good reason to innovate (Välikangas et al. 2009). A company may stick to old and inefficient ways for much too long, because of such collective pessimism. This can happen even if the company possesses all the shared knowledge and capabilities to innovate. Such a state of inaction can lead to a problem of many hands, and a purely a motivational one. For example, innovation trauma can discourage members from proposing much-needed improvements to the rest of the group, or make it shameful for anyone to get excited about them once they are proposed.

In some cases, motivational irresponsibility can go so far that it becomes irresponsible to reason. In a case of heavy collective trauma, it can become impossible for the collective to change its motivational patterns. In such cases, we can speak of the collective not meeting the collective moral competence condition. Because in a case of severe innovation trauma, the mechanism of control over the *joint* action, which is the fear of innovation in this case, does not respond to reason. This means that the collective does not fulfill one of the conditions of moral responsibility, indicating irresponsibility, or at least a lack of collective responsibility as a virtue. One sometimes hears about organizations that are “beyond saving” because the collective culture is such that no new manager or no voice of reason can inspire change. In other cases, a collective lack of motivation for something can prevent the group from revising epistemic norms, information flows or structures that need revision. This can in turn lead to epistemic or structural problems which prevent the group from fulfilling the epistemic and control condition for moral responsibility, meaning that the group is irresponsible.

Collective Responsibility as a Virtue is the Long-Term and Stable, Acquired Trait of a Group to Make Itself Meet the three Conditions for Collective Moral Responsibility

Having described the vice of irresponsibility, and how it can emerge in a collective through epistemic, structural and motivational issues, let us again turn to its positive opposite, collective responsibility as a virtue. The virtue of collective responsibility must be a trait to actively avoid the epistemic, structural and motivational pitfalls described in the previous section.

Possessing the collective virtue of responsibility means that the collective avoids vices of irresponsibility. However, it is also useful to characterize the virtue of responsibility in a positive sense, by what it is, rather than only what it avoids. Fahlquist's characterization of responsibility as virtue provides such a positive description, which aligns with avoiding the pitfalls which lead to PMH. Namely, a responsible agent actively takes responsibility (Nihlen Fahlquist 2015). A responsible agent acts in such a way that it fulfills the knowledge, control and moral competence conditions.

The collective virtue of responsibility can be defined as:

The collective long-term and stable trait of a group to actively make sure that it fulfills the three conditions of moral responsibility. A responsible group makes sure it has collective knowledge of what it is doing, ensures that there is a feeling of collective control over what it is doing, and the process whereby the group acts is reasons-responsive.

Virtues are traditionally acquired traits. Some philosophers also specify that a virtue must be acquired through significant effort (Zagzebski 1996). The same goes for collective responsibility as a virtue. Similar to the way that a PMH is acquired, or created, over a long period of time, it takes time to fix it. It also takes significant effort.

Acquiring the collective virtue of responsibility is not easy and requires time and effort. This can be seen in how complex it is for companies to take on this kind of "self-improvement" projects. For a company, particularly a big one, to change its personality usually requires many different effortful long-term interventions. Take for example the banana company Chiquita's corporate responsibility plan, which was executed in response to the negative attention on the company, which involved the following actions:

- hiring a corporate responsibility officer;
- selecting an external measurement standard (in Chiquita's case SA8000) to assess performance;

- creating a code of conduct, training employees in corporate social responsibility;
- “mentoring and coaching managers to build decision-making skills that integrate CR-criteria in evaluating options”(Werre 2003).

As this plan shows, the required changes are not merely effort-requiring for the top management, but for all levels in the company. Managers need to develop their skills and all employees need to be trained. New structures are created and new collective knowledge and joint action are introduced. The process of becoming a more responsible company can be characterized as effortful joint action based on collective knowledge about the company’s goals.

Epistemic, structural and motivational elements can also be found in Chiquita’s corporate responsibility program. For instance, epistemic functioning of the collective is improved by providing training for all levels of employees ensures that everyone is aware of corporate responsibility goals and responsibilities. This creates collective knowledge and a collective intention, making it easier for the group to fulfill the knowledge condition. Making corporate responsibility criteria a part of the evaluation of managers is a structural change, which makes it easier for managers to prioritize corporate social responsibility tasks. Motivation towards becoming a more socially responsible organization was also enhanced by training, setting incentives and mentoring programs. A collective that develops itself in this way is more likely to collectively fulfill the conditions for moral responsibility.

4.6 Concluding Remarks

Whether we can attribute moral responsibility to groups matters. Many of the greatest challenges humankind is facing are in some way caused by groups. Being able to assign responsibility in a manner that is philosophically justified is important for legitimizing action to fix the problem. If it is true that such fragmented groups cannot be held responsible, then this means that we cannot blame groups like British Petroleum and the other partners for a mistake like the Deepwater Horizon disaster at all. Since it is also hard to hold individuals responsible for this disaster, the result would be that we cannot ascribe the responsibility to anyone. This is very counterintuitive and likely to lead to practical impasses where no individual or institution feels compelled to take responsibility for the problem. Our ascriptions of responsibility seem to fail us in such cases, indicating that they may need replacement. Using a virtue-based one allows us to judge that the conglomerate of parties involved at Deepwater Horizon was

collectively irresponsible and hence have the duty to develop their collective character to a more responsible one.

Furthermore, groups often commit wrongful acts precisely because of internal fragmentation, disorganization or bad information flows. This means that the problem that applies to the Deepwater Horizon case, is likely to apply to other such problems too. Asking whether a group is responsible, rather than whether it is responsible for x , may also be a more productive way of approaching collective responsibility. Because it focuses on the long term operation of the group and prescribes realistic advice for the underlying cause for problems. Changing the way we think about responsibility as a system and a group character trait is more holistic than seeing responsibility as only *responsibility for x* . It is likely to be the kind of thinking that we need to solve humanity's greatest problems.

5 The Relationship Between the Individual and the Collective



5.1 Introduction

I have argued in the previous chapters that a focus on the collective level is important in understanding moral behavior and virtuous traits of groups like creativity and responsibility. These examples show that we have good reason to answer the research question *Does collective virtue theory produce valuable insights when applied to real-life collective action cases?* in the affirmative.

However, we should not forget the interests of the individual group member. There are likely many examples of groups that seem to possess virtues like creativity or responsibility collectively, but where the individuals are not treated justly by the group agent. Think of an institution that performs exceedingly well, improving the world or generating brilliant works. Imagine that the good performance rests on a high work pressure that consumes the personal lives of the members. As I argued in chapter 2, collective agents' virtues do not always require virtuous states of mind in individual members. Neither do they always require individual members to flourish. Insisting on collective virtue only may run the risk of losing sight of the interests of individuals.

Collective virtue theory, therefore, needs to be supplemented with conditions for fair group membership for individuals. If we are pure virtue theorists, the good ness of such conditions may be explained in terms of a collective virtue of justice. However, I will approach formulating these conditions from the perspective of the individual.

I carried out empirical research together with other researchers. We interviewed members of various community energy projects, and tried to determine whether they had a just relationship with the collectives that they were a part of. Community energy projects are an interesting case study on the relationship between the collective and individuals because there are little formal rules and expectations dictating how much members should contribute to the project, compared to for example employees in a company.

This case study revealed that members found it important that there was a balance in how much effort they put into the project, how much power they had within the project and how much risk the project imposed on them. If participants were exposed to a large risk, or took large burdens, they could be compensated by giving them a lot of influence, or power, inside the community energy project. If you take risk and put in effort, you also get to decide what happens in the community energy project. Similarly, a lack of power was not seen as a problem for participants in those RECs where participants were exposed to little risk and burden. We see here a clear connection between procedural and distributional justice. Exploring this connection reveals two justice equilibria in RECs, one of which we call the high risk high effort high power equilibrium, and the other the low risk low effort low power equilibrium. Some RECs function according to a high risk high effort high power equilibrium, and others according to a low risk low effort low power equilibrium. Individual members also often choose to take high risk, effort and power, as a community-hero, or to settle with low power but also low risk and low effort, as a sleeping member. These factors could be understood as the fulfillment of interdependent principles of justice: procedural and distributional justice. We used a justice framework, popular in energy ethics, to frame the importance of risk, effort and power for individual members.

Three conclusions can be drawn from these findings. The first is that energy justice literature must take the interdependence of these dimensions of justice into account. The second is that the governance and policy surrounding community energy projects can and should support the negotiation of interdependency of tenets. The third and more general conclusion, which touches upon one of the central questions of this dissertation, is that a balance of risk, power and effort is a useful tool for conceptualizing when members of groups have a just relationship with the group

agent that they are a part of. This contributes to answering the research question: *Does collective virtue theory produce valuable insights when applied to real-life collective action cases?* Because it highlights an important dynamic between the individual and collective which is better analyzed in practice through an understanding of the individual's needs. Collective virtues, therefore, are not enough to understand all questions of ethics in collective action.

5.2 Energy Justice and Renewable Energy Communities (RECs)

Let us begin at what justice means in the particular context of this case study, energy communities. Energy justice has become an substantial field of research (Heffron et al. 2015; Jenkins et al. 2016; C. A. Miller et al. 2013; B. K. Sovacool et al. 2016; B. Sovacool and Dworkin 2014). The dominant approach in energy justice literature is to refer to three tenets of energy justice: distributional, procedural and recognition justice. This approach was first proposed by Mccauley et al. (2013) and developed further by others (Jenkins et al. 2016; Mccauley et al. 2013; B. K. Sovacool and Dworkin 2015).

Distributional justice is geared at just outcomes in terms of the distribution of benefits and disadvantages (Jenkins et al. 2016). Distributional justice aligns with various philosophical ideas about the importance of distributing goods and bads in a just society.

The second tenet, procedural justice, is the sense of justice which is concerned with the justice of the process by which benefits and disadvantages are distributed. It is concerned with how decisions are made and whether the procedures to make decisions are just (Jenkins et al. 2016). Generally, energy justice literature favors democratic processes as being more procedurally just (Mccauley et al. 2013; B. Sovacool and Dworkin 2014; Stephens et al. 2018; Stephens 2019). Recognition justice places importance on the inclusion of people, especially marginalized groups. An example of a recognition justice issue would be a failure to recognize the energy needs of people, some of whom, for instance, depend on a higher indoor temperature for health reasons (Jenkins et al. 2016). Sometimes recognition justice is seen as falling under procedural justice.

According to some, renewable energy communities, henceforth called RECs, are a promising way to bring about a just energy transition (McCabe et al. 2018). One could say that RECs redistribute power over the energy market to citizens. However, recent literature has also been critical.

For example, RECs redistribute risks of the energy transition to individual participants or households (Adams and Bell 2015). More democracy is generally praised as making the energy system more just. However, in reality inequalities of power can exist within RECs which are not immediately visible in the formal structure of the REC (Van Veelen 2018). Mundaca et al. (2018) describe cases of RECs where the procedures of decision-making are fair, but the end-result tends to benefit some people more than others. Localist energy policy has also been criticized for being unrealistic, deferring responsibility from the state on to unwilling citizens and making unfounded assumptions about local scale being more democratic (Catney et al. 2014; Purcell 2006). Other literature speaks about the green energy transition including a sacrifice of communities who are disadvantaged by green energy technologies (Finley-Brook and Thomas 2011; Scott and Smith 2018). The critical literature on localism and inequality in RECs hints at the idea that RECs could become, in the worst case, “sacrifice zones” that bear a larger than average share of the risks of the energy transition. This highlights the urgency of understanding what makes a REC just.

Reviews of energy justice literature show that energy justice literature tends to use the three tenets of energy justice, although work using other approaches exists (Bommel and Höffken 2021; Heldeweg and Séverine Saintier 2020; Horst 2014; Lacey-Barnacle 2020; Rasch and Köhne 2017). While literature on the three tenets of energy justice and RECs exists, there is little engagement with the question of whether some tenets are more important than others and what the minimal conditions for justice are. For instance, Mundaca et al argue that their cases scored well on procedural justice and less well on distributive justice (Mundaca et al. 2018). Van Veelen makes a similar point, that accounting for the tenet of procedural justice with local control and autonomy can sacrifice other areas of justice like equality (Van Veelen 2018). It has also been shown that distributive and procedural justice are important in determining the amount of citizen support for energy projects (C. Walker and Baxter 2017). No conclusions are drawn about whether REC cases as a whole are just or not. The tenets of justice are often treated as separate aspects which contribute to justice. Van Bommel and Höffken (2021) approach energy justice in RECs in this way too. The authors highlight some procedural, distributive and recognition justice issues within RECs from literature, but remain neutral on their interconnections and the respective normative weight of each issue. Literature remains agnostic about how much democracy or equality is enough and whether the fulfillment of all tenets must be experienced by REC members.

Philosophical theories of justice often contain multiple principles, just like the energy justice framework. However, generally a ranking of the principles and a description of their interrelations is given. For example, a crucial part of Rawls's theory of justice, in addition to the principles, is a prioritization of the principles. The principle of fair equality of opportunity comes before the difference principle: meaning that it is more important that inequalities are attached to offices open to all, than that the inequalities are to the benefit of the least advantaged. These two principles make up the second principle of justice, which is lower in priority than the first principle (Rawls 1971). This kind of scheme does not exist for the energy justice framework, making it difficult to understand when exactly a REC can be deemed just. Creating such a scheme, where the tenets of energy justice are related to each other, requires understanding the interrelations between the tenets in RECs. Our findings on these interrelations emerged from an empirical study of RECs.

5.3 Selection, Method and Description of Case Studies

Selection and Method

We collected qualitative data from 6 energy communities. These were located in Belgium, the Netherlands, Poland, Denmark and Germany. We approached the cases and set up interviews with members. We conducted semi-structured interviews with members, asking about the energy community they were involved in, what they think about it and what problems and benefits they experience. We collected from each case study distinctive justice issues. Through analyzing interviews and hosting workshops and sharing case study narratives, we were able to identify core concerns in the cases and patterns in the way those core concerns manifested.

Three basic REC characteristics were used to compile a pragmatic set of European case studies to analyze with optimized representativeness and validity how the tenets of energy justice emerge in RECs. We adopted a purposeful sampling method: we identified cases by relevance to the topic of interest (Creswell 2006). We used maximum variation sampling: we selected cases which are as diverse as possible with regard to the topic of interest in order to document a broad range of dimensions to a topic (Creswell 2006). In our case, the topic of interest was justice and energy community governance.

We selected our cases, for maximum variation, from European countries with a strong energy community movement or tradition like the Netherlands, Germany and

Denmark (Beuse and Organisationen for Vedvarende Energi 2000; Forsman et al. 2020; Holstenkamp and Stier 2018; Karnøe and Jensen 2016; Nyborg and Røpke 2015) in contrast to countries with a weak or emerging energy community movement or tradition, such as Poland (Adamczyk and Knyszewska 2017; Producent et al. 2012) with Belgium being somewhere in between. We also selected some RECs aimed at low-income residents and some aimed at high income residents. Some of our cases experiment with cutting edge technological developments (e.g. virtual power plant like in the Belgian case, peer to peer trading in the Danish case) in contrast to cases that rely on more traditional technology and business models (like the German case, with its traditional cooperative structure and renewable energy sources). Most cases were cooperatives which were already running, but we selected one (Brainport Smart District) which is in the planning-phase, in order to observe the setting-up of a cooperative.

Each case was studied and analyzed by one of us. We collected from each case issues which participants of the REC were dissatisfied about. The dissatisfaction or satisfaction of participants was assessed by the researcher who spoke to them. We collected points of dissatisfaction, which clustered into 3 themes. We labeled these themes as core concerns. The core concerns, we discovered, were linked to each other and together explained the dissatisfaction or satisfaction of participants.

Table 1 shows the six selected cases and an overview of the empirical methods used. Central and overarching in all cases were the participant interviews to enable comparison, complemented with specific methods based on contextual boundary conditions to collect empirical data.

Table 1: Empirical material for each case study

Energy community	Empirical methods
Housing Cooperative South in Wroclaw	11 interviews, survey research, panel, workshops, participatory observation.
Brainport Smart District in Eindhoven	Interviews with 9 participants, observations at meetings and public events, participatory observation
Nautilus in Amsterdam	Interviews with 3 individual participants and one group interview with 3 participants simultaneously. Observation at one meeting.
Erneuerbare Energien Rottenburg (eER)	Interviews with 3 individual participants, analysis of 261 documents reduced to 24 core documents, participatory observation.
EnerGent cooperative (Flanders region, Belgium)	6 interviews, meetings with researcher doing case study research on EnerGent in the context of the Interreg cVPP project
The Energy Collective/Svalin (DK)	7 qualitative interviews with participating researchers, 1 group interview with 3 inhabitants, participatory observations in the community

Housing Cooperative in Wroclaw (Poland)

The Housing Cooperative Wroclaw in Poland, henceforth called HC Wroclaw was begun to decrease the costs of living and carbon emissions. The project covers 102 buildings and nearly 11,000 dwellings, in which about 30,000 inhabitants live. The first stage of the project consists of investments in a solar power plant for the highest buildings. The second involves producing energy with photovoltaics for running air source heat pumps. The solar power plant was developed in consultation with inhabitants.

Brainport Smart District in Helmond (the Netherlands)

Brainport Smart District (BSD) is a cooperation between Eindhoven University of Technology, the municipality of Helmond and various other organizations to build a livable, healthy, energy-neutral and sustainable district which serves as a testbed for new technologies. One of the design principles of the project prescribes renewable

energy generation in the district. The first inhabitants have been invited to take part in designing the part of the district where they will live. The future inhabitants must also collectively make decisions about how and whether they will produce their own energy. The infrastructure of the district is still being planned simultaneously. Hence, the residents are furthering their own plans, occasionally with the help of the BSD organization, within a risky and changing context. The process is still ongoing at the time of writing this article (2021).

Nautilus in Amsterdam (the Netherlands)

Nautilus is a housing cooperative in Amsterdam that built their own housing complex co-creatively. Getting permission to install a heat pump with photovoltaics required a legal battle. Now inhabitants have been living in the complex for 3 years. They run the complex and the maintenance of the heat-pump through voluntary initiatives. The residents have extensive democratic procedures for decision-making which diverge from legally defined cooperative voting procedures. Each resident has veto power.

Erneuerbare Energien Rottenburg (Germany)

Founded in 2009, Erneuerbare Energien Rottenburg eG (eER) aims at advancing regional small-scale renewable energy projects as well as shaping the German energy transition from the grassroots, mainly through solar PV and wind power projects. eER currently counts more than 200 members. Financing is provided by the members' shares, external financing and internally generated capital. The community sells the electricity generated to the property owners who provide their rooftops to eER, mostly the city of Rottenburg. Their individual consumption is turned quasi-renewable, as the electricity generated in eER equals the approximate amount of electricity consumed by the total of its members.

EnerGent in Gent (Belgium)

The EnerGent cooperative originated from social initiatives organized by volunteers in the Macharius neighborhood, which is a disadvantaged neighborhood in the city of Ghent in Belgium. Around 2013, the idea arose to invest in renewable energy in Ghent through a citizen cooperative. The EnerGent energy cooperative was founded as a place-based community driven by the goal to make society more sustainable, but also to reduce inequality by making energy generation more accessible for low-income

households (Summeren et al. 2020). For the time being, they still account for around 50% of their income from subsidies, but the intention is to reduce this percentage in the future.

The Energy Collective/Svalin (Denmark)

The Energy Collective is a “living lab” project that involves the housing cooperative Svalin, near Roskilde in Denmark. The focus is mainly to develop ways to share/trade energy among people living in communities – in other words, they are looking at “new approaches for all to directly exchange energy in a truly consumer-centric manner” (The Energy Collective 2021). Blockchain is one of the main designs that is explored. The project website reads: “Houses and shared infrastructure there were designed to accommodate solar panels, a geothermal heat pump and electric vehicles. The community as a whole is energy positive, meaning that it produces more renewable energy than it consumes, on a yearly basis” (The Energy Collective 2021).

5.4 Findings: Risk, Effort and Power as Core Concerns within RECs

Three core concerns emerged from our interviews and observations. The core concerns are what justice issues in RECs are generally about: risk, effort and power.

We understand risk as the chance of negative consequences for a participant, such as financial loss, property damage or troubles with other people that can happen as a result of entering a REC. We understand effort as the time and work that participants put in to maintaining or improving the REC. We understand power as the ability of participants to influence affairs within the REC.

We observed that whether the justice issues were seen as problematic or not depended on the balance of the impact of these core concerns of the participants. There seemed to be two justice equilibria, in which the state of affairs was considered just. The first equilibrium is a state where participants are exposed to low risk, with low effort required from them and have a low amount of power in governing the REC. The second justice equilibrium is where participants are exposed to a high risk, with a lot of effort required from them and have a high amount of power in the REC. We will now describe how these just equilibria were manifested in our cases, as well as how unjust states of affairs emerged. We will describe three states of affairs which were experienced as just by members, with regard to the internal functioning of the REC. We exclude concerns of justice regarding issues external to the REC, like whether

participants felt their REC was treated justly by the legislation of the country they were in.

Community Heroes: High Risk, High Effort, High Power

Some RECs consisted of very active and risk-taking members. One of our cases, Nautilus in Amsterdam, involved a very high risk for participants. In the planning phase, when the group had been established in order to design, plan and finance the building of the house, each member contributed 16,000 euros, with a risk of losing the money. As time passed and more of the money was spent, the chance of getting the money back decreased. Some inhabitants described the stress that was caused by various disappointments, like the social housing corporation which pulled out of the project or the miscalculation of costs leading to a sudden need to cut 2 million. In the case Nautilus, the heating installation posed a large financial and technical risk. The risk materialized after the construction when many defects were discovered, including a large leak. Having the leak fixed required, according to the inhabitants at the time of the interviews, juridical action against the contractor.

The housing complex and the heating system took a lot of effort to build and still take a lot of effort to maintain. Various mistakes of the contractor in building were found afterwards and took a long time to get fixed. Some inhabitants expressed a desire to carry out fewer house-related duties.

Although participants took great efforts and high risk, what they got in return was a substantial amount of power within the cooperative. In Nautilus we saw the justice equilibrium high risk high effort high power. Nautilus works with a decision making process, designed by a member, based on consensus which does not conform to the legal standard of voting in cooperatives. Every person in the cooperative has veto power on all decisions. This means that one person can block any decision, and if that happens then the majority must propose an acceptable alternative to the disagreeing person. The interviewees viewed this system as being more democratic than voting. One of them mentioned that the reason they had to put an emphasis on democracy was the sensitivity of the financing of the project, especially at the beginning. This narrative portrays the sentiment that if everyone takes a big risk, then they must be given enough say about how things are done. This supports the idea that RECs sometimes compensate for risk with decision-making power.

Some of our case studies were “living lab” projects, which meant that the project was aimed at experimenting with new technologies. In the Svalin community in Denmark and Brainport Smart District in Eindhoven, the innovativeness of

technology and the risk it caused was also met with a movement towards extensive consent-driven procedures. Voting was considered too limiting. BSD and Svalin also had extensive discussions on how the decision-making structure should be.

In Svalin this discussion was prompted by an event. A vote was held on whether the communal dinners should be vegetarian. The vote resulted in a decision for allowing meat, which led to a very active household leaving the community. After this, the community began working with other ways of voting, instead of relying on for and against votes, which were considered too limited. While debating power structures explicitly among members can be considered more empowering than sticking to an uncontested power structure, it is not always practical, especially in large communities. The case studies that did this were all smaller in size.

Sleeping Members: Low Risk, Low Effort, Low Power

Other cases consisted of less active and less risk-taking members. We found that some cases imposed a low risk on participants, required low effort and granted participants low power. This was not felt to be unjust by participants. It is alright to have little influence in something which imposes little risk on you and requires little effort.

HC Wroclaw adopts a minimally democratic but easy strategy because of a general adversity to change among residents and employees. All projects are carefully prepared by the project team to avoid resistance. Any risks in the project and possible resistance are identified. The practice shows that a relatively small number of active and organized members can effectively implement changes and, by building a rhetorical advantage. The vast majority, therefore, is not granted much power in the process, other than a vote.

In HC Wroclaw and EnerGent, which are big communities, the decision-making structure is set and stuck to according to legal default. This default generally consists of a board and the possibility for members to vote on issues. Power is centralized and this structure is not contested. In EnerGent, there is a formal democracy as each member has in principle one vote per issue, and there is a general assembly once a year. However, the agenda is set by a limited group of members who, because they possess specific resources (e.g. technical, juridical knowledge) represent the community as a whole and act on their behalf. In the case of eER, the formal hierarchy also follows fixed statutes: the board is appointed (for three years with the possibility of re-appointment) and recalled by the supervisory board. The supervisory board is elected by the general assembly for three years with the possibility of re-election.

However, the residents in eER, EnerGent and HC Wroclaw did not generally find it problematic that their democratic control was minimal. This is because not much risk was taken by the majority of participants and not much effort was required. Therefore, there was not much at stake for most participants. HC Wroclaw had removed most of the risk and burden to members. HC Wroclaw used a pilot project to eliminate some technology-related risks. A pilot project led to the discovery of a flaw in the attachments of photovoltaics, which would have caused the roof to leak. The project was also financially de-risked due to an external funding partner. The governance structure was traditional and legally supported. Participation in EnerGent is to a large extent 'de-risked'. 60% of investments are covered by the project funds. The project involved only tests under simulated circumstances (e.g. simulated dynamic prices) so there was no substantial risk for the households. The project also targeted low-income participants, used a similar strategy. The low amount of power for participants, in these cases, is compensated by the lack of risk and lack of burden for participants, making the situation acceptable to our interviewees.

The dynamic of power compensating for risk and effort can also be observed at the level of individual decisions on how much to participate in the REC. HC Wroclaw, EnerGent and eER all involved community-hero residents who participated in the board, and a host of sleeping members, who participated less. Those who took functions as board members were often accepting of the fact that other members did not put in as much work as they did. The board members, individually, had more power, in return for their effort. The fact that most members did not participate actively was not seen as a problem in these cases. As one of the block leaders in HC Wroclaw commented - "it is natural for such a sizable organization that most of the persons do not get involved proactively in the community life." eER's board consisted of only three members (out of a total of 230 members). The board members stated that their commitments amount to a lot of work, but they do not perceive this distribution as unfair. They perceive their individual workload as a contribution to Germany's energy transition.

Unjust State of Affairs: High Risk, High Effort, Low Power

When risk, effort and power were not well balanced, participants in our cases experienced unfairness. If one takes a large risk and puts a lot of effort into a REC, but still has a low level of influence, one may, often justifiably, feel wronged.

An example of this occurred within Brainport Smart District. In the project, which is participative and co-creation-focused regarding other areas than energy too, the risk

for the participants was framed as a part of their conscious choice to “pioneer” in “the smartest district in the world.” Many of the participants were comfortable taking risks, as long as they would get to choose how their district would look. Initially, participants believed to be part of a high risk, high effort, high power project. This justice equilibrium only fell apart at the moment that a lack of clarity and information within the BSD project as a whole constrained the future residents’ ability to plan and take action. The multitude of other parties and their indecision made the future residents structurally powerless. At this point, the system became a high risk high effort low power -system, which was perceived as unjust by participants.

Individuals within RECs can also experience an individual lack of due power for the risk one takes and the effort one puts in. A resident of Nautilus expressed sympathy for the idea that those who put in effort should have more influence. He felt that the extensive democratic procedures, which were in place to increase each members power, were in some sense failing the people who were putting in effort, or taking initiative.

“I am used to it being so that the one with the paint bucket decides what color the walls are, that initiatives are encouraged. And that there is a lack of initiatives, so the people who take initiatives get the freedom to do it the way they want. But here that is completely turned around. . . . But what happens now is that people come in with an initiative, and then there are always 2 or 4 people who say: we’re not doing that. And that changes the initiative. . . . And there is a meeting and eventually your nice initiative turns to some grey mash. . . . There is so much resistance against everything.”

This highlights a perspective where equally distributed power is unjust, or at least ineffective for incentivizing initiative-taking, since some members put in more effort than others.

Individuals could end-up in an unjust situation also due to being exposed to a larger risk than other participants. In Brainport Smart District differing levels of commitment among the future residents caused an unequal distribution of risk. A participant stepping out of the project would mean a re-planning for everyone. If one person steps out, there is one household less, which requires spatial and energy-related re-organization, costing time and possibly money for participants. Some of the very committed members perceived the lack of commitment from others as a risk for themselves.

Different people have different needs for becoming empowered. One participant at BSD worried that the wishes of certain people were being ignored, because those people are shyer or less loud in meetings. An interviewee from Nautilus echoed a

similar worry “If two people bring in an opinion very passionately, then they often get what they want.”

Apart from differences in sociability or rhetorical skill, difference in knowledge and know-how were also mentioned as factors that determine the level of empowerment. According to the statutes of eER, any member of the community can be elected supervisory board member or appointed board member. However, in practice, executive roles require expertise in business, administration and energy technology. Because of this, some participants found it difficult to participate in the community in these circumstances. BSD participants expressed similar worries. This goes to show that some members in communities can have needs regarding type of dialogue, use of terminology and atmosphere, which are not always recognized by others. Personal differences, therefore, can sometimes limit the power of individuals within a REC, causing some members to lack due power individually.

5.5 Conclusion

Conclusion for Theory: Energy Justice Theory Should Expand to Include the Interdependency of Tenets

Philosophical theories of justice, like that of Rawls (1971), provide principles of justice as well as a prioritization of principles. Because the energy justice framework does not yet provide such a prioritization, scholars have only been able to point out justice issues in RECs but not determine when a REC is just or unjust all things considered. The importance of a balance between risk, effort and ownership allows us to create a more comprehensive framework.

Risk and effort are matters of distributive justice or how the costs and benefits are distributed. Questions of power could be classified as a matter of procedural justice, or how decisions are made and who gets to influence them. Therefore, the fact that risk and effort can be compensated with power means that distributional justice issues can be compensated with more procedural justice. One tenet of energy justice can compensate for another. Similarly, a lower amount of power, an issue of procedural justice, can be compensated with less risk and less effort. Less risk and effort are a distributional benefit as well as an increase in recognition justice, because it allows those with low incomes or those that are more risk averse to join the REC.

This has an important theoretical implication, which creates a link between the different tenets. It means that the tenets of justice are interdependent. Actions that

procedural justice requires depend on the distributional aspects of the REC. What can be considered a just amount of power depends on how much risk and effort one puts in. And that the actions that are required by distributional justice depend on procedural matters. A just amount of risk and burden is dependent on how much influence one has. Secondly, the negotiation between risk, effort and power can also be seen as a tradeoff between distributional, procedural and recognition justice.

Work on community energy treats the tenets of energy justice as separate components of a justice framework (Frate, Brannstrom, de Morais, & Caldeira-Pires, 2019; Gross, 2007; Liebe, Bartczak, & Meyerhoff, 2017; Mundaca et al., 2018; Walker & Baxter, 2017b). However, as these findings show, there is a need to treat them as interdependent, meaning that no conclusions can be drawn about one, without drawing conclusions on the others.

Conclusion for Practice: REC Governance and Policy can Support the Negotiation of Interdependency of Tenets

Energy communities internally are clearly not immune to justice concerns, despite their empowering potential. The justice equilibria and unjust states that we found among our cases highlight the centrality of risk-bearing and effort in renewable energy communities and how risk-bearing affect the internal politics and ethics of these communities. The negotiability of risk and effort with power means that anyone engaging in REC governance must acknowledge and quantify each core concern and determine whether they are in balance. REC governance is influenced by policy, therefore, our findings are relevant for energy policy too.

The choice between the justice equilibria could even be seen at the level of individual members. Members in communities sometimes choose to be ‘community heroes’, taking up high risk high effort high power roles, joining the board or starting or participating in initiatives. Members also sometimes choose a low risk low effort low power role, being one of the less active members in the community.

This dilemma manifests itself for individuals joining a REC and for policy makers. A REC member must decide which role to take and may ask themselves the question: Which is more valuable to me, influence or freedom from burden? And from the community-perspective, one may ask, in which role am I more advantageous for my community, as a community-hero with lots of burden and influence, or as a sleeping member with little of both? Both of these roles have their respective moral demands. The number of community-heroes and sleeping members that an energy community

needs to function depends on the context and the legislation surrounding energy communities.

For the policy maker, the dilemma revolves around questions regarding community heroes and sleeping members: In whose benefit should REC policies be? Should they benefit those who take on burdens and risks, or those who do not, and perhaps cannot, do so? Should REC-related policies stimulate citizens to become community-heroes, or is this unfair to those who cannot? Are benefits for sleeping members perhaps unfair to community-heroes?

A practical example of how the justice equilibria interact with policy: While legally defined default structures served well for some low risk low effort low power communities, they were considered too limiting by the high risk high effort high power communities. This goes to show that policymakers should be aware of the different justice equilibria among energy communities, for policy to be fitting.

In this paper we have shown our case studies to reveal that the tenets of energy justice are in practice negotiable and interrelated. This raises various normative questions about whether this interrelation is ultimately justified according to moral theory, and whether it is justified when taking into account various external factors affecting the energy market. For instance, even if groups of community heroes, or high risk, high effort and high power groups are satisfied with their REC, is it fair that citizens take upon themselves the risks of the energy transition like this? Further normative research is needed on whether such high-risk RECs are justified in the greater scheme of things.

Conclusion for the Relationship Between Collective Agents and Individual Members

This research shows that people in community energy projects experience a need for a balance of risk, effort and power in order to feel satisfied with the energy community. In these cases, justice towards individuals can be realized in different ways. If the group demands little from the individual, then the individual does not need to have much power over the group. However, if the group demands effort and risk-taking, then individuals must be given more power.

The scope of this particular framing of justice in group relations is probably limited to such contractual groups as energy cooperatives. It seems plausible that different types of individual-group relationships, like a person's relationship with their family or friends, might conform to different justice norms. One may be willing to take a lot of risks and put in a lot of effort for their friends without feeling like they need power

inside that particular group. This is why the individual-group agent justice framework which I have put forward here does not generalize to all group agents. It is perhaps limited to energy communities and co-creation projects only, and perhaps also to volunteer organizations. Whether the balance of power, effort and risk, and the negotiation that it indicates between procedural and distributional justice, is an apt way to describe other types of groups would require further research.

6 Do Collective Virtues have to be Scaled-Up Individual Virtues?

So far I have argued for a theory of collective virtues chapter 2, and shown in chapter 3 and 4 how virtues like creativity and responsibility are manifested in the collective form. Additionally, I have argued in chapter 5 that there are some important factors that determine how just a particular group or group membership is for individuals. If we examine the argumentation I have presented in chapters 2, 3 and 4, we see that I have mostly used the individual virtue as a model for the collective virtue.



For instance, in chapter 2 I use Zagzebski's list of properties of individual virtues and show how collective traits can have these properties too. In chapter 3, I break the individual virtue of creativity into 3 components and show that these can apply to collectives.

In fact, most defenses and criticisms of collective virtues so far have focused on proving or disproving that groups can have collective versions of individual virtues. However, one might ask: why should we assume that group virtues are scaled up versions of individual virtues? Is it possible that group virtues are a completely different kind of animal? Does it make sense to demand that they resemble individual virtues? In this chapter I propose a further possibility regarding group virtues, which deviates slightly from the theory I have put forward so far. I propose that group virtues can be something different than individual virtues, and whether this is plausible or not depends on the theory we adopt about how virtues are grounded.

6.1 The Argument in Brief

As I've mentioned in previous chapters, I am not the first to write about collective virtues. In social epistemology, various philosophers have discussed collective epistemic virtues (Lahroodi 2007; Levy and Alfano 2019; Smart 2018a). Some work on collective epistemic vices (the negative opposite of epistemic virtues) also exists

(Broncano-Berrocal and Carter 2019). Collective moral virtues have also been debated in recent years (Beggs 2003; Cordell 2017; Fricker 2010; Konzelmann Ziv 2012; Sandin 2007).

Many existing accounts of collective virtue tend to share a common assumption, that collective virtues must be upscaled versions of individually virtuous psychologies. This assumption makes arguments for collective virtues vulnerable to a particular kind of criticism. Critics of the concept of collective virtue also share the preoccupation with collective virtues needing to resemble individual virtues. Criticisms of collective virtues tend to go along these lines:

- Individual virtues always contain internal component x;
- A collective of people cannot collectively possess component x;
- Therefore, collective virtues cannot exist.

Underlying this argumentation scheme is the assumption that to be called a virtue, the structure of group virtues should resemble individual virtues. But let us ask, why should it? Arguments such as the above only work if one assumes that the good of moral virtue is grounded in the goodness of the psychological components of virtue. Some virtue theorists, notably those endorsing agent-based virtue theories, hold this view (Slote 1995). In my view, the current work on collective virtues makes an implicit commitment to agent-based virtue ethics. This commitment makes it vulnerable to the above critique.

This is why agent-based virtue theories, out of all the types of virtue theories, might be the most difficult to scale up to the collective level. If we are serious about collectivist virtue theory, then perhaps we should consider whether non-agent-based virtue theories offer different pathways to understanding collective virtues. If we were to discard the agent-based virtue theory and adopt exemplarism or eudaimonism about collective virtues, then we may discover a version of collectivism that is less vulnerable to criticism. Exemplarist theories do not require virtue to have a set of any particular components. Rather they ground the goodness of virtues in their contribution to flourishing or admirability. Therefore, collective virtues remain a viable possibility at least for virtue ethicists and responsibilist virtue epistemologists with exemplarist or eudaimonist convictions.

6.2 Collective Virtues in Current Literature

Plato's Republic famously draws a connection between individual flourishing and the flourishing of the state. Harmony in the state mirrors the harmony of the soul. All

virtues must be present for a state or a person to thrive. Plato scholars have shown that this can be interpreted in two ways.

One interpretation is that if all citizens of a state possess all the virtues, then the state will also be virtuous. The other is that the virtues can be distributed. For example, a state can also flourish if only its ruling class is wise, its soldiers courageous and its craftspeople temperate. The first interpretation corresponds with what is now called summativism about collective virtue, meaning that collective virtue is short hand for describing a collection of virtuous individuals. Every member must possess the full spectrum of virtues, for the state to be called fully virtuous.

In the second interpretation, the state is a supervening entity which has mechanisms that are analogous to the components of the individual soul. The ruling class functions like the ratio, the soldiers like the virtue of braveness and the craftspeople like temperance [Mulgan (1968);, 435e 1 - 3 in Plato's *Republic*]. Not all citizens need to be fully virtuous, it is enough that each class operates as it should. This last interpretation is very close to how collective virtues are conceptualized in current literature in virtue ethics and epistemology.

Fricker (2010) argues that institutions can have virtues in two distinct ways. They can arise out of group members' joint commitment to a motive or goal. For this, it is not required that the group members commit to that motive or goal as individuals. For instance, members of an amateur dramatics society may not by themselves be committed to the continuation of their society, but in their role as members, they are. This commitment held by the members is thus irreducible to the individual commitments.

The second way in which institutional virtue can arise is through the practical identities of members. A person, who is a part of a group, can take on a practical identity (wearing a hat) as a manifestation of their membership. An example of this could be a church committee, consisting of people who are judgmental towards poor people in their daily lives. Even such members may display more understanding attitudes in the actions of the church committee, feeling that it is the stance of the church to refrain from judging others for their lack of wealth. Fricker's strategy is to argue that different components of virtue, like motivation, can have an irreducible group analogue. Therefore, given that it is also common to speak about groups as if they have certain virtues or vices, we may as well attribute virtues and vices to groups.

Fricker maintains that her argument works whether one has a skill-based or motivation-based virtue theory. She constructs the skill-based version of her argument, which states that "A collective C has a skill-based virtue V just when the

members of C, qua members of C, commit to achieving the end of V, and they achieve this end reliably” (Byerly and Byerly 2016, p. 41; Fricker 2010). An example of this would be a night-watch group, which jointly act in a way that results in collectively vigilant behavior. A motivational account would state: “A collective C has a motive-based virtue V just when the members of C, qua members of C, commit to achieving the end of V because it is good, and they reliably achieve this end” [Fricker (2010); 242; Byerly and Byerly (2016), 40].

The addition of “because it is good” refers to a joint recognition of the value of V. An example of this would be a diligent and thorough research team where “Its members all jointly commit to the motives of diligence and thoroughness; and the team lives up to those motives by proving reliable, over an appropriate span of time and contexts, in achieving their ends” (Fricker 2010, p. 242). In both cases, the argument rests on the idea of practical identity, or committing to a cause or activity qua being a group member.

Beggs (2003) takes a different strategy in arguing for group virtues. He argues that a group manifesting a certain kind of virtuous behavior, when none of the individuals can be credited with the virtue themselves, is enough to constitute a group virtue. He emphasizes the importance of virtuous behavior in the group being stable and its members being committed enough to the group to have it stay that way. Beggs does not require a group motivation, but states that his thesis merely relies on a certain kind of group agency, where actions can be predicated to the group collectively rather than distributively.

In Beggs’s view, a group can possess a group virtue simply by systematically behaving in a virtuous manner, regardless of the motivations by which the virtuous behavior comes about. Let’s say that a church committee decides to engage in extensive charity work, because all the members are desperate to post about charity work on social media and impress their friends. Even in such a case, the group as a whole is behaving virtuously, even if the individuals are not.

Fricker endorses the practical identity or joint commitments, which are intrinsically collective, as bases of collective institutional virtue. Cordell argues that this is not enough to make a virtue irreducibly collective. Both the joint commitment and the practical identity are irreducible to the personal commitments of the members. But this does not undo the fact that each member still chooses to take that practical identity individually, or commit to the group commitment individually. To quote Cordell (2017): “[Fricker] shows the irreducibility of a group of hat-wearers’ motives to a group of non-hat wearers’ motives, but then takes this to show the

irreducibility of the group motive per se” (48). He means that just because a group of “hat-wearers” may share motives which the same group as non-hat wearers would not share, it does not mean that the motivations in such a group are irreducible to the motives of the members. The same problem goes for practical identities. Hence, Cordell shows that Fricker’s definition of group virtues does not describe anything irreducible, but rather something which is better described as a collection of individual traits, attitudes or commitments.

Cordell (2017) also criticizes another conceptualization of collective virtue, that of Beggs. Beggs argues that a group displaying a certain trait, which is over and above any trait displayed by individuals within the group, can be said to have a collective virtue. The basis of Beggs’s argument is that “virtue” is simply a description of behavior and if that behavior is manifested, then the virtue can be predicated to the group. This, Cordell argues, is too lenient regarding the definition of virtue. If any “good” group trait can be a virtue, then this undermines the purpose of using the terminology of virtues and vices in the first place, rather than just collective good and bad.

Byerly and Byerly (2016) argue for collective virtues via a different route. Their argument however also rests on the assumption that the existence of collective virtues depends on the existence of collective analogues of individual psychological phenomena. Byerly and Byerly take virtue to be a disposition towards a certain kind of belief, and a certain kind of action. In seeing virtues as dispositions, the authors claim to take an Aristotelian approach. Instead of looking for collective analogues of skills or motivations, as Fricker does, they construct a theory of collective dispositions and show how this notion has advantages over collective skills or motivations as a basis for collective virtue.

6.3 These Arguments only Make Sense if the Goodness of Virtues is Grounded in the Goodness of Certain Psychological Dispositions

What all of the above accounts of collective virtue have in common, is that they all, in one way or another assume that collective virtues consist of upscaled versions of individual psychological phenomena. Fricker tries to show that collectives can have equivalents of motivation and skill. Cordell (2017) tries to argue that Fricker’s and Beggs’s accounts of collective virtue are not functionally equivalent to individual virtue. Byerly and Byerly (2016) try to show that collectives can have dispositions like the ones that exist in virtuous individuals. This preoccupation with collective virtue resembling individual virtue makes sense in the context of virtue theories that ground

the goodness of virtue in the goodness of the psychological components of virtue. This is why it seems like all of these authors are, perhaps implicitly, committed to agent-based virtue theories.

To clarify this, let us examine what is meant by the grounding of goodness in a virtue. Every moral theory has a foundational concept, or source of intrinsic good. One can ask, for any such intrinsic good, why it is good or what makes it good? One way to ask this question is to ask: in what is its moral goodness grounded? I assume that for one moral fact to be grounded in another means roughly that the former fact is explained by the latter and that the truth of the latter is necessary for the truth of the former. I acknowledge that metaphysical grounding is a contested notion, and that both explainability and necessity as components of grounding relations are also contested.¹⁷ However, for my purposes here defining grounding in this way suffices.

In agent-based virtue theories, the answer to the question “Why be virtuous?” lies in the inherent goodness of the psychological dispositions making up a virtue. For Slote, who endorses an agent-based view of virtue ethics, a virtue is good because it is constituted by a good motivation. The intrinsic goodness of virtue is located in the motivation. The goodness of a good motivation is not grounded in anything else (Slote 1995). Slote’s agent-based virtue ethics has been criticized for its inability to distinguish between doing the wrong thing and doing the right thing for the wrong reasons. Sidgwick presents a counterexample where a prosecutor does his duty by prosecuting the defendant but does so with malicious intentions. In such a case, Slote’s agent-based virtue theory would conclude that the prosecutor acted wrongly. After all, in Slote’s theory motivation is everything and the concept of a right action is derived from the virtuousness of the agent. Slote dismisses this criticism by showing that a prosecutor who, horrified by his own malice, fails to prosecute is failing in terms of having the correct motivation out of a sense of duty, to prosecute (Van Zyl 2005).

If one accepts a view like this, then the most important thing about virtues is the possessor’s motivation. Due to this, the primary method of establishing whether someone possesses a virtue is to check whether they have a virtuous motivation. If an agent-based virtue theorist like Slote wanted to check whether a particular type of agent is capable of possessing a virtue, then the logical place to start is by determining whether the agent is capable of a virtuous motivation. For instance, knowing whether philosophical zombies can have virtues requires determining whether philosophical Zombies can have something like a virtuous motivation, or stable character traits. This

¹⁷ For a more nuanced discussion of grounding, see Trogdon (2013).

will of course go for determining whether a group agent can possess a virtue. It must be determined whether a group agent can have good motivations.

How would one go about determining whether a collective can have good motivations or not? Presumably, exactly in the way Fricker goes about it. The first question one is faced with is: “what grounds a good motivation?” Intuitively, one may answer that a good motivation is grounded in certain cognitive phenomena experienced by an individual mind. Only individuals experience cognitive phenomena, therefore, one may be inclined to think that groups cannot have motivations. However, Fricker (2010) argues, using Gilbert’s plural subject thesis, that virtuous motivations can be shared. Postulating the fundamentality of “we intentions” and collective agency above individual agency are ontological moves that state that the concept of “motivation” is not necessarily grounded in the cognitive phenomena of an individual. Fricker argues that collective virtues can exist, because it is possible for motivations to be collective. In Fricker’s argument it is clear that the existence of collective virtues is seen to depend on the possibility of collective motivations.

Cordell (2017)’s argument attacks the step that Fricker makes between the claim “groups can have motivations” and “practical identities of group members can constitute a group motivation.” Cordell argues that practical identities are better explained as individual commitments of individuals to the group, or individuals insofar as they are group members, not as one supervening group commitment. The lack of a collective and supervening group motivation means that the prospects of collectivism are limited according to Cordell. Cordell also draws the conclusion that upscaling motivations is problematic, and therefore collective virtues probably cannot work. In the making of this inference, Cordell implicitly commits to the view that a good motivation is the most essential property of virtue.

Similarly to Fricker and Cordell, Byerly and Byerly (2016) also commit to an agent-based grounding of virtue in their account of collective virtues. Namely, they define the virtue as a disposition and argue that collectives can have good and bad dispositions. However, not all virtue theories ground the good of virtues in the goodness of psychological dispositions or motivations. Exemplarist and eudaimonist virtue theories have a different grounding, which means that a different kind of argument may be possible for showing that groups can have virtues.

6.4 Collective Virtues can also be Grounded in Eudaimonia or Admirability

The foundational concept of virtue theories is the virtue, an inherently good character trait. Virtue ethicists have tried to ground the good of virtue in a naturalistic manner. Eudaimonism grounds the good of virtue in the human need for eudaimonia. Character virtues, as defined by virtue theories, are dispositions to act in a certain way which is good for others, and necessarily also leads to a flourishing life. Because it is in the interest of every person to flourish, the answer to the question “why be moral?” is provided, because it is good for you. Hence, a connection is made between the “telos” of a person, their happiness and their moral obligations, each stems from the human need for flourishing. Eudaimonists do not necessarily see psychological components of virtue, like good motivation, as inherently good. These things can be instrumental to eudaimonia. Eudaimonia is, in turn, grounded in desirability. Desirability provides an answer to the question “why be moral?”

Other virtue ethicists, notably Zagzebski (2017), adopt an exemplarist view on virtue theories. Exemplarism means that the goodness of virtues is grounded in, not eudaimonia or desirability, but admirability. Zagzebski’s exemplarism is modelled after the theory of direct reference in philosophy of language. The theory of direct reference states that the meaning of a word is determined by pointing at observable phenomena, i.e. to determine the meaning of “water” one should point at water, not try to define water, i.e. “a colourless drinkable liquid”. Zagzebski argues that moral virtues are identified in the same way, by pointing at exemplary people, like Jesus or Socrates. Such people are moral exemplars and they are identified through the emotion of admiration. Hence, in Zagzebski’s account, the exact structure of the virtue is not as important for establishing that it is a virtue. In the same way that the meaning of “water” is not established by describing its structure “H₂O, clear liquid” but by pointing at it.

6.5 Collective Virtues may not have to be Structurally Equivalent to Individual Virtues if they are Grounded in Desirability or Admirability

If we assume that eudaimonism is correct, then what do we make of the question “does trait x count as a virtue?” In order to establish whether a trait is a virtue or not, one would begin by trying to determine whether the trait is likely to, by itself, generally lead the agent to flourish. Of course, in eudaimonist virtue theory, whether one flourishes or not depends crucially on whether one is contributing to the flourishing

of one's community (Macintyre 1986). Because flourishing in connection with the flourishing of others, is the most inherently good aspect of virtue, it becomes most important for a collective virtue also to lead to flourishing in connection to the flourishing of others outside the collective. It can be debated whether in the case of collective virtues this should be the flourishing of individual group members or the flourishing of the group as a whole. Arguably, both should count.

Eudaimonists disagree on what exactly is meant by a virtue's tendency or ability to produce eudaimonia. However, it is generally agreed that the goodness of virtue is grounded in its ability, tendency or likelihood of producing eudaimonia. The significance of virtue as a concept, then, also depends on its reference and connection with eudaimonia. For a eudaimonist, the question of whether virtues can be possessed by collectives becomes the question: "Can collectives display long-term stable collective traits that lead to collective flourishing, in a way that connected to the flourishing of those outside the collective?" An example of a flourishing group might be a military unit with strong internal bonds of trust and friendship, with the collective goal of preserving peace and justice. Another example might be a welcoming and inclusive theatre group, which readily succeeds in its mission to empower people and amplify marginalized voices through the art of performance. Flourishing could be seen as an indication of virtue. No upscaled psychological phenomena are needed.

If one adopts an exemplarist view, then one may assert that whether group virtues exist or not depends on whether or not it makes sense to recognize exemplary groups through the emotion of admiration. If people indeed admire groups, then a theory of group virtues can be constructed using the qualities of groups that are admirable. For an exemplarist, the question about the existence of group virtues is: "Are groups sometimes admired morally in a way that is distinct from admiring the members of the group individually?" Again, the examples of the military unit and the theatre group also serve as examples of groups that trigger the emotion of moral admiration, which serves as proof of group virtue.

Both exemplarist and eudaimonist virtue theories have various conditions, depending on the specific version of the theory. Without going into the specifics of what each type of theory tends to require, I think it is worth exploring how eudaimonist or exemplarist grounding of virtue would affect theories of collective virtues. If flourishing or admiration is more important than the existence of an inherently good motivation, then it becomes possible for a virtue to be multiply realizable. Collective virtue theory would perhaps no longer be affected by the style

of argument that denies collective agents the possibility of having unified motivations, intentions or supervening thoughts.

6.6 A New Direction for Collectivist Virtue Ethics and Epistemology?

It is tempting to believe in the beautiful and elegant idea that what happens in the individual psyche also happens in group dynamics. The idea that the soul mirrors the state reveals a fractal metaphysics, where the same phenomena repeat themselves on different ontological levels. This elegant idea is exactly what lies at the heart of most collective virtue accounts. If we are serious, however, about the potential explanatory benefits of collective virtue theories, it may be worth exploring whether we can construct a theory of group virtues without the problematic assumption that group virtues must contain collective versions of psychological states. Perhaps this can be done by starting from an exemplarist or eudaimonist grounding of virtues.

7 Conclusion

I began this dissertation with the research question: *Do irreducible group level phenomena create problems for, or limit the scope of, virtue ethics as a moral theory?* In chapter 2 I answer this question in the affirmative. Phenomena such as mandevillian morality do indeed create problems for pure virtue ethics. When a bigot causes a group to be more reflective, then the moral goodness of this reflectivity cannot be explained by reference to individual virtues only, since the vice plays a crucial role in constituting it. This goes against a fundamental claim of pure virtue theory, which is that all moral good is grounded in the goodness of virtue.

The first research question lead to the next: *How can virtue ethics respond to collectively valuable phenomena which are not reducible to individual virtue?* I focused on how a pure virtue theory can be adapted so that it can explain the moral goodness of mandevillian morality. I argued that ascribing virtues and vices to collective agents in addition to individual agents can solve this problem. It allows us to say that the bigot is vicious, but the group (including the bigot) is virtuous. This means that even mandevillian morality can be explained with reference to virtue, but in this case, the virtue of the collective agent.

This answer yields a third question: *Does collective virtue theory produce valuable insights when applied to real-life cases?* I argue in chapter 3 and 4 that it offers novel and valuable insights to treat the virtues of creativity and responsibility as collective virtues. Chapter 3 deals with collective creativity, as a virtue of collective agents. Seeing collective creativity as a virtue allows us to explain the consistent creative performance of some groups. There are also other normative reasons for placing inherent value on collective creativity rather than only individual creativity. As I argue in the chapter, discouraging competition and fostering collaboration has proven very productive innovation practices. It also discourages obsession with individual geniuses and a competitive culture that arises from it.

Chapter 4 also provides an answer to the question: *Does collective virtue theory produce valuable insights when applied to real-life cases?* The chapter focuses on the collective virtue of responsibility. I describe how a group like an organization can develop this collective virtue by avoiding epistemic, structural and motivational pitfalls. I also argue that seeing collective responsibility as a virtue can also contribute something novel to the debate about collective responsibility. Once we see responsibility as a

virtue we overcome one of the challenges for the notion of collective responsibility, namely its inability to deal with problems of many hands.

Chapters 3 and 4 therefore answer the research question *Does collective virtue theory produce valuable insights when applied to real-life cases?* in the affirmative. The chapters illustrate the usefulness of the notion of collective virtue in different contexts. However, in chapter 5 I add some nuance to this answer. Collective virtue theory, as I have presented it, needs to be supplemented with a description of just conditions for group membership. I have shown how such just conditions can be formulated in the case of renewable energy communities. I propose that a balance between risk, effort and power can provide guidance for justice towards group members.

In chapter 6 I respond to some criticism of collective virtue theory by suggesting a way of grounding collective virtues that does not require the existence of collective mental states. If we adopt an exemplarist grounding of virtue theory, in the style of Zagzebski¹⁸, then we can also ground the goodness of collective virtues in something other than collective equivalents of mental states. This means that collective virtues do not need group-analogues of mental states. Such collective mental states are often vulnerable to criticism. Therefore, doing away with them makes collective virtue theory stronger.

I have proposed that virtue theory should be expanded to become applicable to collective agents. I am not proposing that we abandon individual virtues, but that we supplement virtue theory with a category of virtues for collectives. I do not think that collective virtues are any more important or foundational, but simply that they exist alongside individual virtues. While conceptions of individual virtue and vice play a role in our interpersonal ethics, collective virtues are likely to play a role in public life, organizations and the groups, societies, families and scenes that we are a part of. Others have already argued that collectives can bear responsibility for things (List and Pettit 2011; Smiley 2008). In my work, I have argued that the moral agency of collectives can be understood as moral personhood in a sense similar to moral personhood in the individual psyche.

The findings of this thesis are relevant for all virtue theorists. Some see virtue theories as useful additions to other fundamental moral theories (Nussbaum 1999). Pure virtue theorists see the virtues as the fundamental grounds of moral goodness (Anscombe 1958; Macintyre 1986; p.79 Zagzebski 1996). Whether one is a pure virtue theorist or not, the insights from this thesis are relevant.

¹⁸ See Zagzebski (2017)

7.1 Relevance for Pure Virtue Theorists

Pure virtue theory is a meta-ethical conviction that all moral value and disvalue are attributable to virtues and vices. In other words, pure virtue theorists believe that virtue ethics is a fundamental moral theory which explains all moral value. As I have argued, collective action has features which detach its goodness from the goodness of participating members. I have mentioned a few examples in the chapters.

For example, in the case of mandevillian morality, vice plays a systematic role in good group behavior. This means that the goodness of a groups behavior does not result from the goodness of members only. Similarly, the problem of many hands, which I discussed in a later chapter, also includes the collective acting in a morally disvaluable way, while none of the individuals in the group necessarily act viciously. In these cases, moral virtue on the collective level comes apart from moral virtue on the individual level. This means that pure virtue theorists need to expand the notion of virtue to include traits at the collective level. This may make pure virtue theory more resistant to criticism.

The claims of this dissertation are therefore relevant for pure virtue theory in two ways. First, by pointing out a vulnerability of pure virtue theory. This vulnerability is its inability to account for moral goodness of group behavior which is not reducible to individual virtues. Second, by providing a solution by which pure virtue theory can overcome this challenge, namely embracing collective virtues.

7.2 Relevance for Ethics of Collective Action

Collective virtue ethics allows for normative evaluation of the character and development of groups. Applying this theory to collective action cases, like innovative companies, yields insights about how such collectives should operate.

For instance, reflecting on creativity as a collective virtue reveals the importance and value of practices that enhance that form of creativity. Being a creative group does not necessarily mean that the group should include the most genius and brilliant individuals. Group creativity can arise even if no one in particular is a creative genius. Valuing creativity does not necessarily mean that one allows creative individuals to ideate and lead, but it can also mean enhancing collective ideation. The value of collective creativity is a relevant insight for many organizations, groups, bands or collectives. It also provides reasons to adopt a decentralized, “leading from behind” style of leadership to facilitate collective creativity (Hill 2014).

Seeing collective responsibility as a virtue also provides insights for practice. For instance, as I outlined in chapter 4, it provides a justified way of holding organizations accountable. This is because the collective vice of irresponsibility can be attributed even in cases where the collective agent is too fragmented to fulfill the conditions of moral responsibility. Additionally, seeing collective responsibility as a virtue also provides an incentive to steer the organization towards developing this virtue collectively. This can be done by improving epistemic, structural and motivational features of an organization.

Collective virtue theory can likely be applied to many matters of organizational operation, collective behavior change and collective action. This dissertation is only the beginning.

7.3 Further Research

Collective virtue ethics, as I have presented it here, raises some interesting further research questions. The first set of questions concerns theory development, which will likely contribute to virtue theory and defend pure virtue theory. The second set of questions concerns the potential applicability of collective virtue ethics to various contexts. I will describe each in turn.

As for the first set of questions: I argued for collective virtue ethics as a particular response which virtue ethics can formulate to the challenge of mandevillian morality. However, given that pure virtue theory must expand to collectives, there may be other ways of constructing the theory than the way I have presented. For example, I argued in chapter 6, many proposed collective virtue theories (including the one presented in chapters 1, 2 and 3) take the individual virtue as a model for the collective virtue. For instance, in chapter 3, I construct a model of creativity as a virtue in individuals. Then, I argue that collective versions of the components of this model are possible and can be found in many collectives.

However, there may be another way to conceptualize collective virtues. As I argued in chapter 6, if we adopt an exemplarist view of virtues, we might conceptualize collective virtues differently. An exemplarist approach would start with insights about what groups are admired, and what traits such groups have in common. The viability of this approach to collective virtues needs further research. Exemplarist collective virtues are a promising possibility because it avoids some of the criticisms that can be made about collective virtue accounts that are modelled after individual virtue. Collectives are not always structured in the same way individuals are. This

means that collective virtues which are modelled after individual virtues may be limited in how many groups they are applicable to. Sometimes a group is structured in such a way that a supervenient “group motivation” can be distinguished, as in the case of a single organization with a strong organizational culture. However, often even loosely organized groups are not structured in such a way, as one might say about a community of scientists who work at different institutes.

An investigation into an exemplarist collective virtue theory would probably need to be supported by empirical material about what groups are admired and why. Such admirable groups would also need to be investigated, to understand their supervening traits and mechanisms. After constructing an empirical understanding of what makes groups admirable in the eyes of people, this admirability can be reflected upon philosophically. And the traits of admirable groups can be formulated descriptively and prescriptively without reference to individual virtues as a model.

The second set of questions for further research concerns the applicability of collective virtue theory. Further research should be conducted into how collective virtues can be used as a guide for collective behavior change. This is a promising direction for research because, as I argued in the introduction, many, if not all, of the greatest challenges of humanity are collective action problems. Think of climate change, corrupt governments, aging population and mass surveillance via the internet. All of these problems require mass action and collective behavior change. It seems that our understanding of collective action problems could benefit from collectivistic moral theories. Future research should investigate how collective virtue theories can give guidance in these collective action problems.

Collective virtue theory might for instance give guidance on the development of collective virtues inside organizations that can influence these problems. For instance, multinationals and governments can address climate change by developing the collective virtue of responsibility or collective creativity. It may be interesting to investigate, for instance, what role employees’ identification with company-cultures and roles have for the development of collective virtues. Much work is also needed in defining what exactly collective virtues mean for different types of organizations and collectives. This will in turn determine to what extent theoretical generalizations can be made regarding collective virtues.

Furthermore, the extension of virtues and vices to collective agents may also mean that virtue ethics can give better guidance in large and global questions. Virtue ethics in current discourse is often limited to the small and interpersonal. A more robust

theory of virtue and vice in collective agents may well even serve as a basis for policy decisions.

References

- Adamczyk, I., & Knyszewska, E. (2017). *Housing Economy in the Years 2013–2017*. Statistics Poland. <https://stat.gov.pl/en/topics/municipal-infrastructure/municipal-infrastructure/housing-economy-in-the-years-20132017,9,1.html>
- Adams, C. A., & Bell, S. (2015). Local Energy Generation Projects: Assessing Equity and Risks. *Local Environment*, 20(12), 1473–1488. <https://doi.org/10.1080/13549839.2014.909797>
- Ahmed, S. (2017). *Living a Feminist Life*. Duke University Press.
- Alfano, M., & Skorburg, J. A. (2016). The Embedded and Extended Character Hypotheses. In J. Kiverstein (Ed.), *The routledge handbook of philosophy of the social mind* (1st ed., pp. 465–478). Routledge. <https://doi.org/10.4324/9781315530178>
- Anscombe, G. E. M. (1958). Modern Moral Philosophy. *Philosophy*, 33(124), 1–19. <https://doi.org/10.1017/S0031819100037943>
- Aristotle. (2015). *Ethica Nichomachea*. (C. H. Poortman & Bartel, Eds.) (4th ed.). Uitgeverij DAMON.
- Astola, M. (2021a). Mandevillian Virtues. *Ethical Theory and Moral Practice*, 1–14. <https://doi.org/10.1007/s10677-020-10141-9>
- Astola, M. (2021b). Do Collective Epistemic Virtues Have To Be Scaled-Up Individual Virtues? *Social Epistemology Review and Reply Collective*. <https://social-epistemology.com/2021/07/22/do-collective-epistemic-virtues-have-to-be-scaled-up-individual-virtues-mandi-astola/>
- Astola, M., Bombaerts, G., Spahn, A., & Royackers, L. (2021). Can Creativity Be a Collective Virtue? Insights for the Ethics of Innovation. *Journal of Business Ethics*, 1, 3. <https://doi.org/10.1007/s10551-021-04833-0>
- Axtell, G., & Carter, J. A. (2008). Just the Right Thickness: A Defense of Second-Wave Virtue Epistemology. *Philosophical Papers*, 37(3), 413–434. <https://doi.org/10.1080/05568640809485229>
- Baird, C., & Calvard, T. S. (2019). Epistemic Vices in Organizations: Knowledge, Truth, and Unethical Conduct. *Journal of Business Ethics*, 160(1), 263–276. <https://doi.org/10.1007/s10551-018-3897-z>

- Baregheh, A., Rowley, J., & Sambrook, S. (2009). Towards a Multidisciplinary Definition of Innovation. *Management Decision*, 47(8), 1323–1339. <https://doi.org/10.1108/00251740910984578>
- Bassett-Jones, N. (2005). The Paradox of Diversity Management, Creativity and Innovation. *Creativity and Innovation Management*, 14(2), 169–175. <https://doi.org/10.1111/j.1467-8691.00337.x>
- Beggs, D. (2003). The Idea of Group Moral Virtue. *Journal of Social Philosophy*, 34(3), 457–474. <https://doi.org/10.1111/1467-9833.00194>
- Belbin, R. M. (2012). *Team Roles at Work, Second Edition* (pp. 1–153). Taylor; Francis. <https://doi.org/10.4324/9780080963242>
- Beuse, E., & Organisationen for Vedvarende Energi. (2000). *Vedvarende Energi I Danmark: En Krønike Om 25 Opvækstår 1975–2000*. OVE. https://www.saxo.com/dk/vedvarende-energi-i-danmark_ejvin-beuse_haefdet_9788787660853
- Bjarne, K. (2012). *Project Management: Theory and Practice* (p. 587). Nyt Teknisk Forlag.
- Björnsson, G. (2020). Collective Responsibility and Collective Obligations Without Collective Moral Agents. In *The routledge handbook of collective responsibility*. Routledge. <https://philpapers.org/rec/BJRCRA>
- Boden, M. A. (2003). *The Creative Mind: Myths and Mechanisms*. Taylor & Francis.
- Bommel, N. van, & Höffken, J. I. (2021). Energy Justice Within, Between and Beyond European Community Energy Initiatives: A Review. *Energy Research & Social Science*, 79, 102157. <https://doi.org/10.1016/j.erss.2021.102157>
- Bradley, B. (2005). Virtue Consequentialism. *Utilitas*, 17(03), 282–298. <https://doi.org/10.1017/S0953820805001652>
- Bratman, M. E. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford University Press. <https://philpapers.org/rec/BRASAA-5>
- Brenkert, G. G. (2009). Innovation, Rule Breaking and the Ethics of Entrepreneurship. *Journal of Business Venturing*, 24(5), 448–464. <https://doi.org/10.1016/j.jbusvent.2008.04.004>
- Broncano-Berrocal, F., & Carter, J. A. (2019). The Philosophy of Group Polarization: A Collective Vice Theory. In *Routledge studies in eistemology*. Routledge. <https://eprints.gla.ac.uk/205418/>
- Byerly, T. R., & Byerly, M. (2016). Collective Virtue. *Journal of Value Inquiry*, 50(1), 33–50. <https://doi.org/10.1007/s10790-015-9484-y>

- Carter, J. A. (2016). Robust Virtue Epistemology As Anti-Luck Epistemology: A New Solution. *Pacific Philosophical Quarterly*, 97(1), 140–155. <https://doi.org/10.1111/papq.12040>
- Carter, J. A. (2017). Virtuous Insightfulness. *Episteme*, 14(4), 539–554. <https://doi.org/10.1017/cpi.2016.37>
- Catmull, E. (2008). How Pixar Fosters Collective Creativity. *Harvard Business Review*. <https://hbr.org/2008/09/how-pixar-fosters-collective-creativity>
- Catney, P., MacGregor, S., Dobson, A., Hall, S. M., Royston, S., Robinson, Z., et al. (2014). Big Society, Little Justice? Community Renewable Energy and the Politics of Localism. *Local Environment*, 19(7), 715–730. <https://doi.org/10.1080/13549839.2013.792044>
- Chadwick, R. (2021). On the Politics of Discomfort. *Feminist Theory*, 22(4), 556–574. <https://doi.org/10.1177/1464700120987379>
- Code, L. (1984). Toward a “Responsibilist” Epistemology. *Philosophy and Phenomenological Research*, 45(1), 29. <https://doi.org/10.2307/2107325>
- Cordell, S. (2017). Group Virtues: No Great Leap Forward With Collectivism. *Res Publica*, 23(1), 43–59. <https://doi.org/10.1007/s11158-015-9317-7>
- Creswell, J. W. (2006). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches*. SAGE Publications.
- Davis, M. (2020). A Whistle Not Blown: VW, Diesels, and Engineers. In A. E. Abbas (Ed.), *Next generation ethics* (pp. 217–229). Cambridge University Press.
- Demetriou, D. (2013). The Virtues of Honorable Business Executives. In M. W. Austin (Ed.), *Virtues in action: New essays in applied virtue ethics* (pp. 23–38). Palgrave Macmillan. <https://doi.org/10.1057/9781137280299>
- Driver, J. (1989). The Virtues of Ignorance. *The Journal of Philosophy*, 86(7), 373. <https://doi.org/10.2307/2027146>
- European Commission. (2020). Open and Participative Innovation | Digital Single Market. <https://ec.europa.eu/digital-single-market/en/open-and-participative-innovation>. Accessed 9 January 2020
- Ewing, A. M. (1932). Turning Accidents Into Profit Through Careful Observation. *Journal of Chemistry Education*, 16, 67–79. <https://pubs.acs.org/sharingguidelines>
- Fehr, E., & Fischbacher, U. (2004). Third-Party Punishment and Social Norms. *Evolution and Human Behavior*, 25(2), 63–87. [https://doi.org/10.1016/S1090-5138\(04\)00005-4](https://doi.org/10.1016/S1090-5138(04)00005-4)

- Finley-Brook, M., & Thomas, C. (2011). Renewable Energy and Human Rights Violations: Illustrative Cases From Indigenous Territories in Panama. *Annals of the Association of American Geographers*, 101(4), 863–872. <https://doi.org/10.1080/00045608.2011.568873>
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press. <https://philpapers.org/rec/FISRAC-3>
- Foot, P. (1978). *Virtues and Vices and Other Essays in Moral Philosophy*. University of California Press.
- Forsman, A., De Moor, T., Weeren, R. van, Bravo, G., Ghorbani, A., Ale Ebrahim Dehkordi, M., & Farjam, M. (2020). Eco-Evolutionary Perspectives on Emergence, Dispersion and Dissolution of Historical Dutch Commons. *PLOS ONE*, 15(7), e0236471. <https://doi.org/10.1371/journal.pone.0236471>
- Fricker, M. (2010). 10. Can There Be Institutional Virtues? In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (3rd ed., pp. 235–252). Oxford University Press. <https://philpapers.org/rec/FRICT>
- Gilbert, M. (1987). Modelling Collective Belief. *Synthese*, 73(1), 185–204. <https://doi.org/10.1007/BF00485446>
- Gilbert, M. (2004). Collective Epistemology. *Episteme*, 1(2), 95–107. <https://doi.org/10.3366/epi.2004.1.2.95>
- Gilbert, M. P. (2002). Collective Wrongdoing: Moral and Legal Responses. *Social Theory and Practice*, 28(1), 167–187. <https://philpapers.org/rec/GILCWM>
- Giubilini, A., & Levy, N. (2018). What in the World Is Collective Responsibility? *Dialectica*, 72(2), 191–217. <https://doi.org/10.1111/1746-8361.12228>
- Heffron, R. J., McCauley, D., & Sovacool, B. K. (2015). Resolving Society’s Energy Trilemma Through the Energy Justice Metric. *Energy Policy*, 87, 168–176. <https://doi.org/10.1016/j.enpol.2015.08.033>
- Heldeweg, M. A., & Séverine Saintier. (2020). Renewable Energy Communities As “Socio-Legal Institutions”: A Normative Frame for Energy Decentralization? *Renewable and Sustainable Energy Reviews*, 119, 109518. <https://doi.org/10.1016/j.rser.2019.109518>
- Hicks, D. J., & Stapleford, T. A. (2016). The Virtues of Scientific Practice: Macintyre, Virtue Ethics, and the Historiography of Science. *ISIS*, 107(3), 449–472. <https://doi.org/10.1086/688346>

- Hill, L. A. (2014). *Collective Genius: The Art and Practice of Leading Innovation*. Harvard Business Review Press.
- Holstenkamp, L., & Stier, B. (2018). Dezentrale Energieversorgung in Deutschland in Historischer Perspektive. In *Handbuch energiewende und partizipation* (pp. 387–401). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-09416-4_23
- Horst, D. van der. (2014). Climate Policy and the Siting of Renewable Energy Projects; Towards Common but Differentiated Responsibility at the Community Level. *People, Place and Policy Online*, 8(3), 222–234. <https://doi.org/10.3351/ppp.0008.0003.0007>
- Huebner, B. (2016). The Group Mind: In Commonsense Psychology. In Justin Sytsma & Wesley Buckwalter (Ed.), *Blackwell companion to experimental philosophy* (pp. 292–305). John Wiley & Sons, Ltd.
- Hursthouse, R., & Pettigrove, G. (2018). Virtue Ethics. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy (winter 2018 edition)*. <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=ethics-virtue>
- Irvin, S. (2005). Appropriation and Authorship in Contemporary Art. *The British Journal of Aesthetics*, 45(2), 123–137. <https://doi.org/10.1093/aesthj/avi015>
- Isaacson, W. (2013). *Steve Jobs*. Little Brown Book Group.
- Jenkins, K., McCauley, D., Heffron, R., Stephan, H., & Rehner, R. (2016). Energy Justice: A Conceptual Review. *Energy Research & Social Science*, 11, 174–182. <https://doi.org/10.1016/J.ERSS.2015.10.004>
- Karnøe, P., & Jensen, J. S. (2016). Struggles in Denmark's Transition Towards a Low Carbon Future: Shifts in the Energy Technology Assemblage. Routledge. <https://vbn.aau.dk/en/publications/struggles-in-denmarks-transition-towards-a-low-carbon-future-shif>
- Kieran, M. (2014). Creativity as a Virtue of Character. In *The philosophy of creativity* (pp. 125–144). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199836963.003.0007>
- Kieran, M. (2018). Creativity as an Epistemic Virtue. In H. Battaly (Ed.), *Routledge handbook of virtue epistemology* (pp. 167–177). Routledge.
- Konzelmann Ziv, A. (2012). Institutional Virtue: How Consensus Matters. *Philosophical Studies*, 161(1), 87–96. <https://doi.org/10.1007/s11098-012-9933-4>

- Lacey-Barnacle, M. (2020). Proximities of Energy Justice: Contesting Community Energy and Austerity in England. *Energy Research and Social Science*, 69, 101713. <https://doi.org/10.1016/j.erss.2020.101713>
- Lahroodi, R. (2007). Collective Epistemic Virtues. *Social Epistemology*, 21(3), 281–297. <https://doi.org/10.1080/02691720701674122>
- Lepoutre, M. (2020). Democratic Group Cognition. *Philosophy & Public Affairs*, 48(1), 40–78. <https://doi.org/10.1111/papa.12157>
- Levine, M. P., & Cox, D. (2016). Academic Virtues: Site Specific and Under Threat. *The Journal of Value Inquiry*, 50(4), 753–767. <https://doi.org/10.1007/s10790-016-9579-0>
- Levy, N., & Alfano, M. (2019). Knowledge From Vice: Deeply Social Epistemology. *Mind*. <https://doi.org/10.1093/mind/fzz017>
- Lewis, D. (1986). *On the Plurality of Worlds*. Blackwell Publishing.
- List, C., & Pettit, P. (2011). *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford University Press.
- Macdonald, J. E., & Beck-Dudley, C. L. (1994). Are Deontology and Teleology Mutually Exclusive? *Journal of Business Ethics*, 13(8), 615–623. <https://doi.org/10.1007/BF00871809>
- Macintyre, A. (1986). *After Virtue* (2nd ed.). University of Notre Dame Press. <https://philpapers.org/rec/MACAVN>
- Mandeville, B. (1714). *The Fable of the Bees: Or, Private Vices, Publick Benefits* (1962nd ed.). Capricorn Books. <https://www.worldcat.org/title/fable-of-the-bees-or-private-vices-publick-benefits/oclc/1421916>
- Mauser, W., Klepper, G., Rice, M., Schmalzbauer, B. S., Hackmann, H., Leemans, R., & Moore, H. (2013, September). Transdisciplinary Global Change Research: The Co-Creation of Knowledge for Sustainability. Elsevier. <https://doi.org/10.1016/j.cosust.2013.07.001>
- Maxwell, L., Beltrán, C., Threadcraft, S., White, S. K., Leonard, M., & Honig, B. (2019). The “Agonistic Turn”: Political Theory and the Displacement of Politics in New Contexts. *Contemporary Political Theory*, 18(4), 640–672. <https://doi.org/10.1057/s41296-019-00346-8>
- McCabe, A., Pojani, D., & Broese van Groenou, A. (2018). Social Housing and Renewable Energy: Community Energy in a Supporting Role. *Energy Research & Social Science*, 38, 110–113. <https://doi.org/10.1016/j.ERSS.2018.02.005>

- Mccauley, D., Heffron, R. J., & Jenkins, S. K. (2013). Advancing Energy Justice: The Triumvirate of Tenets. *International Energy Law Review*, 32(3), 107–110. <http://dspace.stir.ac.uk/handle/1893/18349>
- Mill, J. (1956). *On Liberty*. Bobbs-Merrill.
- Miller, C. A., Iles, A., & Jones, C. F. (2013). The Social Dimensions of Energy Transitions. *Science as Culture*, 22(2), 135–148. <https://doi.org/10.1080/09505431.2013.786989>
- Miller, S. (2020). Joint Abilities, Joint Know-How and Collective Knowledge. *Social Epistemology*, 34(3), 197–212. <https://doi.org/10.1080/02691728.2019.1677799>
- Morton, A. (2013). *Bounded Thinking: Intellectual Virtues for Limited Agents* (pp. 1–192). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199658534.001.0001>
- Morton, A. (2014). Shared Knowledge From Individual Vice: The Role of Unworthy Epistemic Emotions. *Philosophical Inquiries*, II(1), 163–172.
- Mouffe, C. (1999). Deliberative Democracy or Agonistic Pluralism? *Social Research: An International Quarterly*, 66(3). <https://philpapers.org/rec/MOUDDO>
- Mulgan, R. G. (1968). Individual and Collective Virtues in the Republic. *Phronesis: A journal for Ancient Philosophy*, 13(1-2), 84–87. <https://doi.org/10.1163/156852868X00056>
- Mundaca, L., Busch, H., & Schwer, S. (2018). “Successful” Low-Carbon Energy Transitions at the Community Level? An Energy Justice Perspective. *Applied Energy*, 218, 292–303. <https://doi.org/10.1016/j.apenergy.2018.02.146>
- Nihlen Fahlquist, J. (2015). Responsibility as a Virtue and the Problem of Many Hands. In I. van de Poel, L. Royakkers, & S. D. Zwart (Eds.), *Moral responsibility and the problem of many hands* (1st ed.). Routledge.
- Nussbaum, M. C. (1999). Virtue Ethics: A Misleading Category? *Journal of Ethics*, 3(3), 163–201. <https://doi.org/10.1023/A:1009877217694>
- Nuyen, A. T. (2007). Confucian Ethics as Role-Based Ethics. *International Philosophical Quarterly*, 47(3), 315–328. <https://doi.org/10.5840/ipq200747324>
- Nyborg, S., & Røpke, I. (2015). Heat Pumps in Denmark—From Ugly Duckling to White Swan. *Energy Research & Social Science*, 9, 166–177. <https://doi.org/10.1016/j.erss.2015.08.021>
- Palermos, O., & Pritchard, D. (2013). Extended Knowledge and Social Epistemology. *Social Epistemology Review and Reply Collective*, 2(8), 105–120.

<https://social-epistemology.com/2013/07/24/extended-knowledge-and-social-epistemology-orestis-palermos-and-duncan-pritchard/>

- Palermos, Spyridon Orestis. (2016). Spreading the Credit: Virtue Reliabilism and Weak Epistemic Anti-Individualism. *Erkenntnis*, 81(2), 305–334. <https://doi.org/10.1007/s10670-015-9741-2>
- Palermos, Spyridon Orestis. (2020). Epistemic Collaborations: Distributed Cognition and Virtue Reliabilism. *Erkenntnis*, 1–20. <https://doi.org/10.1007/s10670-020-00258-9>
- Palermos, S. Orestis, & Tollefsen, D. P. (2018). *Group Know-How* (Vol. 1). Oxford University Press. <https://doi.org/10.1093/oso/9780198801764.003.0007>
- Paulus, P. (2003). *Group Creativity Innovation Through Collaboration*. Oxford University Press.
- Peter, F. (2013). The Procedural Epistemic Value of Deliberation. *Synthese*, 190(7), 1253–1266. <https://doi.org/10.1007/s11229-012-0119-6>
- Plato. (2010). *Republic*. Crane Library at the University of British Columbia. https://www.worldcat.org/title/republic/oclc/1011725329&referer=brief_results
- Poel, I. van de, Royackers, L., & Zwart, S. D. (2015). Moral Responsibility and the Problem of Many Hands. <https://philpapers.org/rec/VANMRA-8>
- Pritchard, D. (2018). Extended Virtue Epistemology. *Inquiry*, 61(5-6), 632–647. <https://doi.org/10.1080/0020174X.2017.1355842>
- Pritchard, D., & Palermos, S. O. (2016). The Distribution of Epistemic Agency. In P. J. Reider (Ed.), *Social epistemology and epistemic agency: De-centralizing epistemic agency*. Rowman & Littlefield.
- Producent, U., Energia, N., & Energia, N. (2012). Czy Spółdzielnie Energetyczne Przyjmą Się W Polsce? <https://www.cire.pl/artykuly/serwis-informacyjny-cire-24/186031-czy-w-polsce-przyjma-sie-spoldzielnie-energetyczne>
- Purcell, M. (2006). Urban Democracy and the Local Trap. *Urban Studies*, 43(11). <https://doi.org/10.1080/00420980600897826>
- Rasch, E. D. D., & Köhne, M. (2017). Practices and Imaginations of Energy Justice in Transition. A Case Study of the Noordoostpolder, the Netherlands. *Energy Policy*, 107, 607–614. <https://doi.org/10.1016/j.enpol.2017.03.037>
- Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.

- Sand, M. (2018). The Virtues and Vices of Innovators. *Philosophy of Management*, 17(1), 79–95. <https://doi.org/10.1007/s40926-017-0055-0>
- Sanders, E. B.-N., & Stappers, P. J. (2008). Co-Creation and the New Landscapes of Design. *CoDesign*, 4(1), 5–18. <https://doi.org/10.1080/15710880701875068>
- Sanders, E. B.-N., & Stappers, P. Jan. (2012). *Convivial Design Toolbox: Generative Research for the Front End of Design* (p. 310). BIS. https://books.google.de/books/about/Convivial_Toolbox.html?id=a8miuAAACAAJ&redir_esc=y
- Sandin, P. (2007). Collective Military Virtues. *Journal of Military Ethics*, 6(4), 303–314. <https://doi.org/10.1080/15027570701755505>
- Sandler, R. (2009). *Character and Environment: A Virtue-Oriented Approach to Environmental Ethics*. Cambridge University Press. <https://philpapers.org/rec/SANCAE-7>
- Sandler, R. L. (2013). Environmental Virtue Ethics. In *International encyclopedia of ethics*. Blackwell Publishing Ltd. <https://doi.org/10.1002/9781444367072.wbiec090>
- Sawyer, R. K. (Robert. K. (2012). *Explaining Creativity: The Science of Human Innovation* (p. 586). Oxford University Press. <https://www.worldcat.org/title/explaining-creativity-the-science-of-human-innovation/oclc/785779526>
- Scott, D. N., & Smith, A. A. (2018). “Sacrifice Zones” in the Green Energy Economy: Toward an Environmental Justice Framework. *McGill Law Journal*, 62(3), 861–898. <https://doi.org/10.7202/1042776ar>
- Searle, J. (1990). Collective Intentions and Actions. In Philip R. Cohen Jerry Morgan & Martha Pollack (Ed.), *Intentions in communication* (pp. 401–415). MIT Press.
- Slote, M. (1995, September). Agent-Based Virtue Ethics. Blackwell Publishing Ltd. <https://doi.org/10.1111/j.1475-4975.1995.tb00306.x>
- Smart, P. R. (2018b). Mandevillian Intelligence: From Individual Vice to Collective Virtue. In J. A. Carter, A. Clark, J. Kallestrup, O. S. Palermos, & Duncan. Pritchard (Eds.), *Socially extended epistemology* (pp. 253–274). Oxford University Press.
- Smart, P. R. (2018a). Mandevillian Intelligence. *Synthese*, 195(9), 4169–4200. <https://doi.org/10.1007/s11229-017-1414-z>

- Smiley, M. (2008). Collective Responsibility. In *Stanford encyclopedia of philosophy*.
<https://philpapers.org/rec/SMICR>
- Smith, N. V. (1982). *Mutual Knowledge*. Academic Press.
<https://catalogue.nla.gov.au/Record/174199>
- Sovacool, B. K., & Dworkin, M. H. (2015). Energy Justice: Conceptual Insights and Practical Applications. *Applied Energy*, 142, 435–444.
<https://doi.org/10.1016/j.apenergy.2015.01.002>
- Sovacool, B. K., Heffron, R. J., McCauley, D., & Goldthau, A. (2016). Energy Decisions Reframed As Justice and Ethical Concerns. *Nature Energy*, 1(5), 16024. <https://doi.org/10.1038/nenergy.2016.24>
- Sovacool, B., & Dworkin, M. (2014). *Global Energy Justice: Problems, Principles, and Practices*. Cambridge University Press.
- Stamou, L. (2002). Plato and Aristotle on Music and Music Education: Lessons From Ancient Greece. *International Journal of Music Education*, 39(1), 3–16.
<https://doi.org/10.1177/025576140203900102>
- Steadman, J. M. (1964). The “Inharmonious Blacksmith”: Spenser and the Pythagoras Legend. *PMLA/Publications of the Modern Language Association of America*, 79(5), 664–665. <https://doi.org/10.2307/461152>
- Steen, M. (2013). Virtues in Participatory Design: Cooperation, Curiosity, Creativity, Empowerment and Reflexivity. *Science and Engineering Ethics*, 19(3), 945–962. <https://doi.org/10.1007/s11948-012-9380-9>
- Steiber, A., & Alänge, S. (2013). A Corporate System for Continuous Innovation: The Case of Google Inc. *European Journal of Innovation Management*, 16(2), 243–264. <https://doi.org/10.1108/14601061311324566>
- Stephens, J. C. (2019). Energy Democracy: Redistributing Power to the People Through Renewable Transformation. *Environment: Science and Policy for Sustainable Development*, 61(2), 4–13.
<https://doi.org/10.1080/00139157.2019.1564212>
- Stephens, J. C., Burke, M. J., Gibian, B., Jordi, E., & Watts, R. (2018). Operationalizing Energy Democracy: Challenges and Opportunities in Vermont’s Renewable Energy Transformation. *Frontiers in Communication*, 3. <https://doi.org/10.3389/fcomm.2018.00043>
- Summeren, L. F. M. van, Wiczorek, A. J., Bombaerts, G. J. T., & Verbong, G. P. J. (2020). Community Energy Meets Smart Grids: Reviewing Goals, Structure, and Roles in Virtual Power Plants in Ireland, Belgium and the Netherlands. *Energy Research and Social Science*, 63, 101415.
<https://doi.org/10.1016/j.erss.2019.101415>

- Swanton, C. (2016). A Virtue Ethical Theory of Role Ethics. *The Journal of Value Inquiry*, 50(4), 687–702. <https://doi.org/10.1007/s10790-016-9582-5>
- Swierstra, T. (2016). Introduction to the Ethics of New and Emerging Science and Technology. In R. Nakatsu, M. Rauterberg, & P. Ciancarini (Eds.), *Handbook of digital games and entertainment technologies* (pp. 1–25). Springer Verlag. <https://doi.org/10.1007/978-981-4560-50-4>
- The Energy Collective. (2021). Context and Motivations. <https://the-energy-collective-project.com/context/>. Accessed 18 January 2021
- Tollefsen, D. (2004). Collective Epistemic Agency. *Southwest Philosophy Review*, 20(1), 55–66. <https://doi.org/10.5840/swphilreview20042015>
- Tollefsen, D. P. (2003). Participant Reactive Attitudes and Collective Responsibility. *Philosophical Explorations*, 6(3), 218–234. <https://doi.org/10.1080/10002003098538751>
- Trogdon, K. (2013). Grounding: Necessary or Contingent? *Pacific Philosophical Quarterly*, 94(4), 465–485. <https://doi.org/10.1111/papq.12009>
- Välikangas, L., Hoegl, M., & Gibbert, M. (2009). Why Learning From Failure Isn't Easy (and What To Do About It): Innovation Trauma at Sun Microsystems. *European Management Journal*, 27(4), 225–233. <https://doi.org/10.1016/j.emj.2008.12.001>
- Vallor, S. (2016). *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (p. 309). Oxford University Press.
- Van den Hoven, J., Lokhorst, G.-J., & Van de Poel, I. (2012). Engineering and the Problem of Moral Overload. *Science and Engineering Ethics*, 18(1), 143–155. <https://doi.org/10.1007/s11948-011-9277-z>
- Van Veelen, B. (2018). Negotiating Energy Democracy in Practice: Governance Processes in Community Energy Projects. *Environmental Politics*, 27(4), 644–665. <https://doi.org/10.1080/09644016.2018.1427824>
- Van Zyl, L. (2005). In Defence of Agent-Based Virtue Ethics. *Philosophical Papers*, 34(2), 273–288. <https://doi.org/10.1080/05568640509485159>
- Walker, C., & Baxter, J. (2017). “It’s Easy To Throw Rocks at a Corporation”: Wind Energy Development and Distributive Justice in Canada. *Journal of Environmental Policy & Planning*, 19(6), 754–768. <https://doi.org/10.1080/1523908X.2016.1267614>
- Walker, R. L., & Ivanhoe, P. J. (2006). Working Virtue: Virtue Ethics and Contemporary Moral Problems. <https://philpapers.org/rec/LWAWVV-2>

- Walzer, M. (1973). Political Action: The Problem of Dirty Hands. *Philosophy & public affairs*, 2(2), 160–180. <https://doi.org/10.2307/2265139>
- Wayman, A. (1982). The Human Body As Microcosm in India, Greek Cosmology, and Sixteenth-Century Europe. In *History of religions* (Vol. 22, pp. 172–190). University of Chicago Press. <https://doi.org/10.1086/462918>
- Werre, M. (2003). Implementing Corporate Responsibility – The Chiquita Case. In *Journal of business ethics* (Vol. 44, pp. 247–260). Kluwer Academic Publishers. <https://doi.org/10.1023/a:1023316303587>
- Wikipedia. (2021). List of Pixar Awards and Nominations (Feature Films). [https://en.wikipedia.org/wiki/List_of_Pixar_awards_and_nominations_\(feature_films\)](https://en.wikipedia.org/wiki/List_of_Pixar_awards_and_nominations_(feature_films)). Accessed 13 January 2021
- Williams, G. (2008). Responsibility as a Virtue. *Ethical Theory and Moral Practice*, 11(4), 455–470. <https://doi.org/10.1007/s10677-008-9109-7>
- Woerther, F. (2008). Music and the Education of the Soul in Plato and Aristotle: Homoeopathy and the Formation of Character. *Classical Quarterly*, 58(1), 89–103. <https://doi.org/10.1017/S0009838808000074>
- Zagzebski, L. T. (1996). *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge* (p. 365). Cambridge University Press. https://books.google.nl/books/about/Virtues_of_the_Mind.html?id=w-YC4foOSBwC&redir_esc=y
- Zagzebski, L. T. (2017). *Exemplarist Moral Theory* (p. 274). Oxford University Press. <https://www.worldcat.org/title/exemplarist-moral-theory/oclc/985422340>

Summary

Virtue ethics has been guiding moral practice since ancient times. Even today, virtue ethics continues to give helpful answers to many moral questions in contemporary life. Various scholars have applied virtue ethics to contemporary challenges like questions around emerging technologies or environmental degradation. The work of such scholars goes to show that many of the challenges in our lives today can be fruitfully conceptualized with traditional virtue ethics.

Virtue ethics is a future-looking and development-focused moral theory. For these reasons it is seen as a very practical theory which excels at giving people concrete guidance in life. The guidance of virtue ethics also focuses on long-term character development, rather than individual actions. Since many of the problems humanity now faces, like climate change, require long-term change rather than one-off decisions, one could argue that virtue ethics is particularly suitable as a guiding moral theory for our times.

However, the strength of virtue ethics, that it focuses on character development, may also be its caveat. The focus on character makes it difficult to apply the theory to collective behavior. Virtue ethics seems to struggle with prescribing conduct for collectives of people. Therefore, virtue ethics falls short in analyzing collective action cases.

One way in which virtue ethics might be able to give guidance for collective behavior is if it could make judgements about collective agents. This dissertation is an exploration into that possibility. I propose, based on earlier work, that collective agents can, much like individuals, have character traits, which can be inherently good or bad. This dissertation justifies and refines a theory of collective virtues and vices.

Chapter 2 begins by arguing that collective phenomena can create problems for virtue theory. An example of this is that sometimes vicious traits of individuals can be a systematic part of something inherently valuable on the group-level. A bigot in a group can cause a lot of discussion and make the group more reflective overall. I call this the challenge of *mandevillian morality* for pure virtue theory.

Is this challenge detrimental to pure virtue theory? This depends on whether pure virtue theory can formulate a response to it. I formulate a possible pure virtue theoretical response in the latter part of chapter 2. I put forward a type of virtue theory with three features that allow it do account for the value of mandevillian morality. This theory has three distinctive features: a distinction between individual and group

virtues, a distinction between motivational and teleological virtues, and an acknowledgement of the normativity of “vicious” roles in groups.

Having established collective virtue theory as one of the possible responses of pure virtue theory to mandevillian morality, the subsequent chapters focus on showing that collective virtue theory can yield valuable insights when applied to real life cases. I do this for two collective action cases.

Chapter 3 focuses on the collective action context of innovating organizations and the virtue of collective creativity. I explore how the virtue of creativity often takes collective forms and show that all elements of a broadly Aristotelian definition of creativity as a virtue can be found on a supervening level in creative collectives (Astola et al. 2021).

Creativity only produces morally justifiable behavior if it is coupled with responsibility. In chapter 4, I show that responsibility can also be a collective virtue. Furthermore, seeing collective responsibility as a virtue solves problems in the notion of collective responsibility. This is because seeing collective responsibility as a virtue can deal better with attributing responsibility in cases of “problems of many hands,” which are morally problematic situations where a collective can be held responsible for something but no individual in the group can.

Having illustrated how collective virtues can make sense of collective action cases, I then turn to some limitations of collective virtue theory in this area. Chapter 5 explains how collective virtue functions as an ideal, but how some minimal conditions often need to be met in order to make group membership just for the members of the group. I use the example of volunteer run energy communities to show that in order for a group to function justly, all members must experience a balance of effort, risk and power. Collective virtue theory therefore can provide useful insight in many contexts, but we need more individual-focused reflection aside it, as is clear from the case of community energy projects.

After presenting a possible response for pure virtue theory, to the challenge of group-level phenomena, I reflect on whether there are more ways to argue for the response I have presented. Chapter 6 is a reflection on the current literature on collective virtues, including the claims presented in this dissertation, which tends to see collective virtues as entities that resemble individual virtues. It is often assumed that collective virtues require collective mental states, like individual virtues. I call this assumption into question and propose that with an exemplarist grounding, collective virtues need not rely on the existence of collective mental states.

In chapter 7, the concluding chapter, I summarize my findings and how they answer the research questions. I also present two directions for future research that are worth investigating: the possibility of an exemplarist theory of collective virtues and the application of collective virtue ethics to various societal collective action problems.

Samenvatting

Deugdethiek is al sinds de oudheid een leidraad voor de morele praktijk. Deugdethiek blijft nuttige antwoorden geven op veel morele vragen in het hedendaagse leven. Verschillende wetenschappers hebben deugdethiek toegepast op de uitdagingen van nu, zoals vragen over opkomende technologieën of aantasting van het milieu. Dergelijke onderzoek laat zien dat veel moderne morele problematiek nog steeds duidelijk geconceptualiseerd kan worden met deugdethiek.

Deugdethiek is een toekomstgerichte en ontwikkelingsgerichte morele theorie. Om deze redenen wordt het gezien als een zeer praktische theorie die uitblinkt in het concreet begeleiden van mensen in het leven. De begeleiding van deugdethiek richt zich op karakterontwikkeling op de lange termijn, in plaats van op individuele acties. In toenemende mate kunnen de problemen waar de mensheid mee geconfronteerd wordt niet meer opgelost worden met eenmalige beslissingen maar eisen deze veranderingen over langere termijn, neem klimaatverandering bijvoorbeeld. Daarom zou men kunnen stellen dat deugdethiek bijzonder geschikt is als morele theorie voor onze tijd.

De kracht van deugdethiek, dat het zich richt op karakterontwikkeling, kan echter ook zijn grootste zwakte zijn. De focus op karakter maakt het moeilijk om de theorie toe te passen op collectief gedrag. Deugdethiek schiet daarom tekort bij het analyseren van collectieve actie.

Een manier waarop deugdethiek een leidraad zou kunnen zijn voor collectief gedrag, is als het oordelen zou kunnen vellen over het gedrag van groepen. Dit proefschrift is een verkenning van die mogelijkheid. Ik stel op basis van eerder werk voor dat groepen, net als individuen, karaktereigenschappen kunnen hebben, die inherent goed of slecht kunnen zijn. Dit proefschrift rechtvaardigt en verfijnt een theorie van collectieve deugden en ondeugden.

Hoofdstuk 2 begint met de stelling dat collectieve verschijnselen problemen kunnen veroorzaken voor de deugdethiek. Een voorbeeld hiervan is dat slechte eigenschappen van individuen soms een systematisch onderdeel kunnen zijn van iets dat inherent waardevol is op groepsniveau. Een onverdraagzame in een groep kan veel discussie veroorzaken en de groep in het algemeen meer laten reflecteren. Ik noem dit de uitdaging van de *mandevillian moraliteit* voor deugdethiek.

Is deze uitdaging problematisch voor deugdethiek? Dit hangt ervan af of deugdethiek er een antwoord op kan geven. In het laatste deel van hoofdstuk 2

formuleer ik een mogelijk deugdtheoretisch antwoord op het probleem. Ik pleit voor een nieuwe vorm van deugdethiek met drie kenmerken die het mogelijk maakt om de waarde van de mandevillian moraliteit te verklaren. Deze theorie heeft drie onderscheidende kenmerken: een onderscheid tussen individuele en groepsdeugden, een onderscheid tussen motivatiegerelateerde en teleologische deugden en een erkenning van de normativiteit van 'vicieuze' rollen in groepen.

Nadat de theorie van collectieve deugden is vastgesteld als één van de mogelijke reacties van de deugdethiek op mandevillian moraliteit, richten de volgende hoofdstukken zich op het aantonen dat de theorie van collectieve deugden waardevolle inzichten kan opleveren wanneer toegepast op gevallen uit het echte leven. Dit doe ik voor twee collectieve actiezaken.

Hoofdstuk 3 richt zich op de context van collectieve actie van innoverende organisaties en de deugd van collectieve creativiteit. Ik onderzoek hoe de deugd van creativiteit vaak collectieve vormen aanneemt en laat zien dat alle elementen van een breed Aristotelische definitie van creativiteit als een deugd op groepsniveau te vinden zijn in creatieve collectieven.

Creativiteit levert alleen moreel verantwoord gedrag op als het gepaard gaat met verantwoordelijkheid. In hoofdstuk 4 laat ik zien dat verantwoordelijkheid ook een collectieve deugd kan zijn. Bovendien lost het zien van collectieve verantwoordelijkheid als een deugd problemen op in de notie van collectieve verantwoordelijkheid. Bij 'problemen van vele handen,' is het toekennen van verantwoordelijkheid lastig. Wanneer we collectieve verantwoordelijkheid echter als deugd zien kunnen we hier veel beter mee omgaan.

Nadat ik heb geïllustreerd hoe collectieve deugden zin kunnen geven aan gevallen van collectieve actie, ga ik in op enkele beperkingen van de theorie van collectieve deugden op dit gebied. Hoofdstuk 5 legt uit hoe collectieve deugd functioneert als een ideaal, maar niet alle belangrijke morele kwesties dekt. Om groepslidmaatschap rechtvaardig te maken voor de leden van de groep, moeten de condities van lidmaatschap aan bepaalde eisen voldoen. Ik gebruik het voorbeeld van door vrijwilligers geleide energiegemeenschappen om te laten zien dat om een groep rechtvaardig te laten functioneren, alle leden een evenwicht moeten ervaren tussen inspanning, risico en macht. Collectieve deugdethiek kan daarom in veel contexten nuttig inzicht verschaffen, maar daarnaast hebben we meer individueel gerichte reflectie nodig, zoals blijkt uit het geval van gemeenschapsenergieprojecten.

Nadat ik een mogelijk deugdtheoretisch antwoord heb gegeven op de uitdaging van fenomenen op groepsniveau, ga ik in op de vraag of er meer manieren zijn om

een collectieve deugdentheorie te construeren. Hoofdstuk 6 is een reflectie op de huidige literatuur over collectieve deugden, inclusief de beweringen die in dit proefschrift worden gepresenteerd. Ik beargumenteer dat de standaard aanpak op collectieve deugden de collectieve deugd ziet als een entiteit die qua structuur lijkt op een individuele deugd. Vaak wordt er aangenomen dat collectieve deugden collectieve mentale toestanden vereisen, net als dat individuele deugden mentale toestanden vereisen. Ik trek deze veronderstelling in twijfel en stel voor dat met een exemplaristische gronding van de deugdethiek, het bestaan van collectieve mentale toestanden niet nodig is voor collectieve deugden.

In het afsluitende hoofdstuk, hoofdstuk 7 vat ik mijn bevindingen samen en hoe ze de onderzoeksvragen beantwoorden. Ik presenteer ook twee richtingen voor toekomstig onderzoek die het onderzoeken waard zijn: de mogelijkheid van een exemplaristische theorie van collectieve deugden en de toepassing van collectieve deugdethiek op verschillende maatschappelijke problemen van collectieve actie.

About the Author

Mandi Astola holds a bachelor's degree in philosophy (BA) and a master's degree in history and philosophy of science (MSc), from Utrecht University. She graduated in the summer of 2018 with a master's thesis titled "Virtue Epistemology and the Ethics of Profiling in Electronic Coaches." In the spring of 2018, Mandi began her PhD project at Eindhoven University of Technology, as a part of the Horizon 2020 project SCALINGS.

Mandi moved to the Netherlands from Finland with her family when she was a child and has lived mostly in Eindhoven ever since. She enjoys good music, especially electronic music, clubbing, drawing, painting, fencing, gardening, caring for her pets, being in the spotlight and drinking good beer. Her greatest passion is philosophy.

List of Publications

Academic Publications

1. de Hoop, Evelien. van Oers, Laura. Becker, Sören. Macrorie, Rachel. Späth, Philipp. Astola, Mandi. Boon, Wouter. “Smart as a global vision? exploring smart in local district development projects” (2019) *Architecture and Culture*, 7, Issue 3: Smartness.
2. Astola, Mandi. “Mandevillian Virtues” (2021) *Ethical Theory and Moral Practice*, 24: 19–32.
3. Astola, Mandi. Bombaerts, Gunter. Spahn, Andreas. Royakkers, Lambèr. “Can creativity be a collective virtue? Insights for the ethics of innovation” (2021) *The Journal of Business Ethics*.
4. Astola, Mandi. “Do Collective Epistemic Virtues have to be Scaled-Up Individual Virtues?” (2021) *Social Epistemology Review and Reply Collective* 10 (7): 40-47.
5. Astola, M. “Collective Responsibility Should be Seen as a Virtue” Accepted for publication in an edited volume titled *Virtues and Values in a Changing World*. Cambridge University Press. 2022.
6. Astola, Mandi. “The Epistemic Innocence of Irrational Beliefs, Lisa Bortolotti, 2020” (2021) **Book review** for *Ethical Theory and Moral Practice* 24 (3): 879-881.
7. Astola, Mandi. “Next Generation Ethics: Engineering a better society, Edited by Ali e. Abbas” **Book review** accepted for publication in *Prometheus: Critical Studies in Innovation*, 2022.

Professional and Popular Publications

1. Astola, Mandi. “Gaat onze zorgzaamheid het coronavirus overleven?” (2020) *Bij Nader Inzien*. <https://bijnaderinzien.com/2020/04/30/gaat-onze-zorgzaamheid-het-coronavirus-overleven/>. English: “Will our virtues of care survive the coronavirus?”
2. Astola, Mandi. “Moet innovatie democratisch zijn?” (2020) *Bij Nader Inzien*. <https://bijnaderinzien.com/2020/11/02/moet-innovatie-democratisch-zijn/>. English title: “Should innovation be democratic?”

3. Astola, Mandi “Fundamental disagreements about online tools” (2020) for website of SEFI (European Society for Engineering Education). <https://www.sefi.be/2020/11/18/fundamental-disagreements-about-online-tools/>.
4. Astola, Mandi. “Teaching ethics to engineers” (2020) Guest posts on the The Philosopher’s Cocoon. <https://philosopherscocoon.typepad.com/blog/2020/02/teaching-ethics-to-engineers.html#more>.
5. Astola, Mandi. “Een existentialistische blik op eco-paniek: Recensie van Henk Oosterlings ‘Verzet in ecopanische tijden’” (2021) *Book review* for *Bij Nader Inzien*. <https://bijnaderinzien.com/2021/04/29/een-existentialistische-blik-op-eco-paniek/>. English: An existentialist view of eco-panic: A review of Henk Oosterling’s ‘Resistance in Eco-panicked Times’
6. Astola, Mandi. “Should philosophers lead interdisciplinary projects?” (2021) Guest post on The Philosopher’s Cocoon. <https://philosopherscocoon.typepad.com/blog/2021/06/should-philosophers-lead-interdisciplinary-projects-guest-post-by-mandi-astola.html#more>.
7. Astola, Mandi. “When unintelligence makes a group smarter” (2021) Guest piece on *Imperfect Cognitions* (edited by Kengo Miyazono and Lisa Bortolotti) <http://imperfectcognitions.blogspot.com/2021/12/when-unintelligence-makes-group-smarter.html>.

Simon Stevin Series in Ethics of Technology
Delft University of Technology, Eindhoven University of Technology,
Univeristy of Twente & Wageningen University
Editors: Philip Brey, Anthonie Meijers, Sabine Roeser and
Marcel Verweij

Books and Dissertations

Volume 1: Lotte Asveld, *'Respect for Autonomy and Technology Risks'*, 2008

Volume 2: Mechteld-Hanna Derksen, *'Engineering Flesh, Towards Professional Responsibility for 'Lived Bodies' in Tissue Engineering'*, 2008

Volume 3: Govert Valkenburg, *'Politics by All Means. An Enquiry into Technological Liberalism'*, 2009

Volume 4: Noëmi Manders-Huits, *'Designing for Moral Identity in Information Technology'*, 2010

Volume 5: Behnam Taebi, *'Nuclear Power and Justice between Generations. A Moral Analysis of Fuel Cycles'*, 2010

Volume 6: Daan Schuurbiers, *'Social Responsibility in Research Practice. Engaging Applied Scientists with the Socio-Ethical Context of their Work'*, 2010

Volume 7: Neelke Doorn, *'Moral Responsibility in R&D Networks. A Procedural Approach to Distributing Responsibilities'*, 2011

Volume 8: Ilse Oosterlaken, *'Taking a Capability Approach to Technology and Its Design. A Philosophical Exploration'*, 2013

Volume 9: Christine van Burken, *'Moral Decision Making in Network Enabled Operations'*, 2014

Volume 10: Faridun F. Sattarov, *'Technology and Power in a Globalising World, A Political Philosophical Analysis'*, 2015

Volume 11: Gwendolyn Bax, *'Safety in large-scale Socio-technological systems. Insights gained from a series of military system studies'*, 2016

Volume 12: Zoë Houda Robaey, *'Seeding Moral Responsibility in Ownership. How to Deal with Uncertain Risks of GMOs'*, 2016

Volume 13: Shannon Lydia Spruit, *'Managing the uncertain risks of nanoparticles. Aligning responsibility and relationships'*, 2017

Volume 14: Jan Peter Bergen, *Reflections on the Reversibility of Nuclear Energy Technologies*, 2017

Volume 15: Jilles Smids, *Persuasive Technology, Allocation of Control, and Mobility: An Ethical Analysis*, 2018

Volume 16: Taylor William Stone, *Designing for Darkness: Urban Nighttime Lighting and Environmental Values*, 2019

Volume 17: Cornelis Antonie Zweistra, *Closing the Empathy Gap: Technology, Ethics, and the Other*, 2019

Volume 18: Ching Hung, *Design for Green: Ethics and Politics for Behavior-Steering Technology*, 2019

Volume 19: Marjolein Lanzing, *The Transparent Self: a Normative Investigation of Changing Selves and Relationships in the Age of the Quantified Self*, 2019

Volume 20: Koen Bruynseels, *Responsible Innovation in Data-Driven Biotechnology*, 2021

Simon Stevin (1548-1620)

‘Wonder en is gheen Wonder’

This series in the philosophy and ethics of technology is named after the Dutch / Flemish natural philosopher, scientist and engineer Simon Stevin. He was an extraordinary versatile person. He published, among other things, on arithmetic, accounting, geometry, mechanics, hydrostatics, astronomy, theory of measurement, civil engineering, the theory of music, and civil citizenship. He wrote the very first treatise on logic in Dutch, which he considered to be a superior language for scientific purposes. The relation between theory and practice is a main topic in his work. In addition to his theoretical publications, he held a large number of patents, and was actively involved as an engineer in the building of windmills, harbours, and fortifications for the Dutch prince Maurits. He is famous for having constructed large sailing carriages.

Little is known about his personal life. He was probably born in 1548 in Bruges (Flanders) and went to Leiden in 1581, where he took up his studies at the university two years later. His work was published between 1581 and 1617. He was an early defender of the Copernican worldview, which did not make him popular in religious circles. He died in 1620, but the exact date and the place of his burial are unknown. Philosophically he was a pragmatic rationalist for whom every phenomenon, however mysterious, ultimately had a scientific explanation. Hence his dictum ‘Wonder is no Wonder’, which he used on the cover of several of his own books.