



Virtue and Experimental Philosophy: Understanding Character Excellences Empirically

Matthew Dennis

This is an unpublished conference paper for the 7th Annual Jubilee Centre for Character and Virtues conference at Oriel College, Oxford University, Thursday 3th – Saturday 5th January 2019.

These papers are works in progress and should not be cited without author's prior permission.

Jubilee Centre for Character and Virtues

University of Birmingham, Edgbaston, Birmingham, B15 2TT United Kingdom

T: +44 (0) 121 414 3602 F: +44 (0) 121 414 4875

E: jubileecentre@contacts.bham.ac.uk W: www.jubileecentre.ac.uk



Virtue and Experimental Philosophy: Understanding Character Excellences Empirically

Matthew Dennis

Warwick University, UK

Abstract

Recently virtue theorists have sought to defend the philosophy of virtue against moral psychologists who suggest that empirical data radically undermines the existence of stable character traits. This defense has largely been successful, both insofar as it has rebutted the claim, but also in terms of bolstering the philosophy of virtue itself. My article surveys the extent to which x-phi offers the resources to bring greater philosophical clarity to the question of the cultural relativity of virtue, focusing its potential to contribute to the question of the universality of virtues specifically.

KEYWORDS: Virtue, Experimental Philosophy (x-phi), Moral Psychology, Eudaimonia, Human Flourishing.

1. Introduction

This article interrogates whether recent work in the empirical sciences can enrich character-based practical philosophy, focusing on the question of which character traits we should consider to be virtues. I argue that experimental philosophy (x-phi) can contribute to this question, and that philosophers stand to gain much from greater collaborative work with natural scientists and historians. I also suggest some ways that this could be done programmatically. Using x-phi in a more systematic way has three potential upshots. (1) It can explain how empirically-informed research from the natural and historical sciences does not undercut the constitutive claims of character-based practical philosophy. (2) It shows how empirical research – if properly applied – can bolster the philosophy of virtue by connecting it to other influential accounts of character development in moral and empirical psychology. (3) A more empirically-attuned philosophy of virtue works in tandem with taking a more critical attitude towards our ‘intuitions’; this allows for a number of important real-world social and political benefits. I begin by showing how practical philosophers have been receptive to work in the empirical sciences throughout the history of their discipline, moving on to show why they would benefit from bringing the question of the relationship between philosophical and empirical research into sharper resolution. After this I sketch what a positive programme of x-phi in virtue could look like, focusing on its potential to help answer the question of the universality of virtues specifically.

2. *Virtues: An Unstable Canon of Character Traits*

Perhaps the domain in which moral philosophers have most explicitly acknowledged the pervasive effects of the natural sciences on philosophical enquiry is in anthropology. Aristotle's seven texts comprising the so-called 'Parva Naturalia' deal with anthropological questions related to the conditions of human flourishing, and those that claim that the *Nicomachean Ethics* advocates a straightforwardly universalist conception of flourishing would do well to remember that Aristotle comments favourably on Meno's suggestion that a number of virtues depend on an individual's position in society, as well as specific virtues relating to one's sex, age, and social status (1984 [*Politics*]: 1.13).¹ Determining the *precise* effect of anthropological and natural factors on the good life is notoriously difficult, however, and Aristotle's own investigations into these factors only gives a provisional indication of their likely strength.

In addition to empirically investigating the anthropological factors that affect human flourishing, the *Ethics* begins with Aristotle asking us to start out by considering our intuitions about what human flourishing is. Unfortunately, Aristotle's notorious comments in the *Politics* on the subservient and docile virtues that he views as applying to women, children, and slaves does little to inspire confidence in the idea that philosophers can always trust their intuitions, as these intuitions may be nothing more than the time-bound prejudices of the cultural or society to which the philosopher belongs. We should learn from this that philosophers often import their own prejudices when appealing to intuitions, which becomes especially dangerous when these intuitions are only shared by those of a similarly homogenous demographic. Aristotle is also open to this charge: not only does he proceed using a method that follows the intuitions of his audience, he often starts his discussion of topics in practical philosophy by recounting the views of the 'good and the wise' using his famous endoxic method. This appeal to intuitions of valorised traditional figures is at the root of complaints that he is methodologically biased towards the status quo and that he routinely endorses it uncritically.

Nevertheless, while some may view Aristotle's *modus operandi* as responsible for what has been characterised as his conservative outlook, there is clearly much to be gained by an outward-facing outlook that examines real-world examples, asking what flourishing consists in for this or that cultural group. We need not think of this as putting ourselves at the mercy of the inherited views of this philosopher or that, but rather an effort to systematically enrich our conception of virtues by practically investigating how such character traits feature in the lives of virtuous individuals. In fact, by identifying aggregates of collectively shared intuitions, we could say that Aristotle goes to great effort *not* to import his own idiosyncratic prejudices, but rather tries to tie his intuitions into a broader framework that is collectively shared. Viewing things in this way would suggest that it is not his starting with empirical data about what the virtues are generally considered to be that is the problem. It is rather that he is overly

¹ Historically speaking, many systematic philosophers who are now canonically-recognised have taken an interest in these questions, although sometimes this area of research has only formed a minor and relatively under-researched part of their legacy. Descartes' *Passions of the Soul*, Kant's *Anthropology From a Pragmatic Point of View*, and concomitantly-written lecture series; J. S. Mill's *The Art of Life*.

selective in the analysis of this data: his choice of certain character traits results from his own biases, and not from the inherent merit of the character traits he chooses to examine.

Moving on from worries associated with the endoxic method, another key moment in the history of the empirical grounding of virtue occurs in Alasdair MacIntyre's *Short History of Ethics*. Here MacIntyre gives a persuasive historical account of the diachronic evolution of the character traits we now consider to be virtues, one which tracks the *etymological development* of virtue-terms, as well as the *practical function* specific character excellences had in specific societies. One of the key findings of the first part of this text is that virtues are not only dependant on historically embedded practices, but that character traits that were once understood as virtues have been subject to the effects of diachronic change. Reading MacIntyre can lead to the impression that not only have virtues changed, but that these changes have slowed – perhaps even stagnated and stultified – since modernity in a way that does not do justice to the dynamic and shifting nature of virtuous character traits. This is what makes MacIntyre's work troubling: because many virtue ethicists tend to assume that their conception of virtue is stable, the idea that this has a contingent historical path requires us to mobilise the resources of the empirical sciences (in this case historical sciences) to show how it is enmeshed within many different dimensions of practical life. MacIntyre's work glaringly shows that canonical cardinal virtues *cannot* be viewed as isolated from the historical, social, and material conditions out of which they emerge (MacIntyre 1966, 1981).

Analogously to MacIntyre's retrospective approach to the historical development and change of virtue, we may even say that there is a forward-looking dimension to this problem, one which shows that future developments in technology will impact on *how* we are virtuous, even on the *very virtues themselves*. This is something that I will pass over relatively quickly in this article, but it is worth remembering that we should not restrict MacIntyre's insights to an analysis of the past because his notion of virtue as a situated practice applies equally to how virtue may develop in the future. Many virtue problems of today's world are certainly comparable to those that were dealt with by the ancient virtue tradition, but it would seem odd if they *all* were. It would be strange if technology had not brought important contexts in which old virtues could be or not be realised, or if certain ethical problems involving new technology required entirely new character traits to address them. Shannon Vallor's recent book, *Technology and the Virtues*, does a good job of looking ahead to the virtues we may be expected to enjoy in the Internet age, for example. Perhaps because of how we live in such a connected age, not only are widely differing world views on display like never before, but these world views put extra tension on the view that the virtues are universal in a simpleminded way. The virtues clearly have important cultural, social, economic factors intimately associated with them, so attending to these factors should not be regarded as a threat to the philosophy of virtues, but rather as a potential opportunity to strengthen the philosophy of virtue itself.

Connecting the examples we have examined above, I will now focus on how character-based practical philosophers have recently had an acrimonious relationship with those who have denied the very existence of stable character traits. The debate is well known to philosophers of virtue, even if only as a cautionary tale, and recounting it here only aims to provide an informative analogy to sharpen our thinking about the relationship of practical philosophy with the empirical sciences. Those familiar with the debate will know that virtue ethicists were at first faced with a series of apparently robust and apparently watertight criticisms of the most central concepts of their discipline. Character, and the virtues and vices it is constituted by, was argued to be illusory when subjected to even cursory

investigation by empirical scientists. The key studies in this literature are well known. It began in 2002 with the publication of John Doris's seminal book, *Lack of Character*, the summary of his collected papers on this topic that he had started publishing from the late 1990s. In this text, Doris argues that we would do seriously question our assumption that character traits are stable, that they can be said to belong to us in any real sense, and – most importantly – that they are anything more than responses to environmental cues. Doris' work let loose a broadside of collaborative work by fellow moral psychologists and practical philosophers who were sympathetic to his key findings, and wished to expand upon his work. Such thinkers eventually became known as 'situationists', on account of how they located the most pressing motivational factors in our practical reasoning in the environment we are surrounded by, rather than in character traits we could be said to possess ourselves.

Many situationist studies are memorable because they are comic, and deflate what has sometimes been considered to be the pompous, priggish, or even hubristic dimension of those who stress the overarching importance of character traits in our practical lives. Few will forget the experiment perpetrated on trainee clergymen who were en route to present a sermon on the Good Samaritan. In this study, the trainees were told they only had minutes to spare to get to the lecture. On the way to the lecture hall, however, they were waylaid by a stooge masquerading as a mendicant who began loudly pleading for assistance. The fact that only 10% of the most tardy clergymen stopped to help initially surprised experimenters (Darley & Batson 1973: 100). Nevertheless, when these results were extrapolated and applied to the central contentions of virtue ethicists by Doris and his followers, they were seized on with glee by those who were already hostile to the philosophy of virtue. Indeed, in the first-wave application of this work by figures such as Doris, it seemed as if the edifice of virtue was going to come crashing down altogether.

Although the story of Doris and the situationists might be a cautionary tale when considering if the empirical resources can enrich virtue ethics, it is also a tale with a happy ending. Rather than undermining claims pertaining to the consistency of character, a second-wave of philosophical reception of these kinds of situationist studies, has bolstered the philosophy of virtue. Philosophers have not just been able to explain Doris' findings, but have been able to incorporate them into their own accounts of virtue, stressing the role of environmental factors in virtue formation and maintenance. For many, this defence has been successful, both insofar as it has restricted the scope of situationalist claims, but also insofar as it has shown how character-based practical philosophy can be understood as working in close conjunction with other disciplines (Alfano 2011, 2013, 2014). Much of the mileage that philosophers have got from engaging with the work of Doris has been in the area of how virtues function (specifically whether such character traits are stable, our own, and robust), but deeper questions regarding the identify of virtues themselves have not yet been asked.

Thinking about these three crossovers between empirical studies and the virtue ethics, summarised above, should lead us to think positively about the prospects for a more empirically-informed philosophy of virtue. Indeed, they may even lead us to think more positively about whether the character traits we typically consider to be virtues are as stable as a non-empirically informed theory might lead us to assume. From the studies outlined above we have reason to believe that the profile of the character traits that we consider to be virtues may well be responsive to a number of factors, and that we can only understand the effects of these factors by using an empirical method. Thinking about the history of the virtues, *à la* McIntyre, reveals that they are responsive to historical practices, to such an extent that these practices radically determine which character traits we regard as virtuous and

which we regard as vicious. Thinking about how emergent technology selects for certain character traits and not others, *à la* Vallor, suggests that we should constantly be aware that today's virtues may not be those of tomorrow. Both these approaches realigns the historical emphasis on how virtue must be situated in a scientific context that pays attention to those non-ethical factors that provides the conditions of virtue, or at least markedly impact upon it. Finally, doing this, integrates our practical philosophy with findings in empirical psychology that support these claims. All the above studies have strongly supported the idea that virtues are not stable. Rather virtue shifts in response to non-ethical factors, even seemingly-trivial ones such as the layout of a room, one's lateness for an appointment, or whether one has recently had a positive or negative social encounter.

3. Prospects for an X-Phi of Virtue

The studies summarised above provide a backdrop to make sense of the work of recently-coined 'experimental philosophers' (x-phi), a catch-all term for philosophers who think that empirical data can and should inform philosophical enquiry in a number of specific ways. It will be useful to distinguish these ways to better elucidate my proposal that x-phi has the potential to shed light on our understanding of the scope of the relativity of virtue, although we should bear in mind that the legitimacy of each of these sub-branches of x-phi is currently disputed, not just among those hostile to x-phi, but also – indeed, most fiercely – among experimental philosophers themselves. Understanding the theoretical limits of x-phi will be helpful in clarifying the scope of my claim, however, so following the recent x-phi literature I shall say at the outset that I am only concerned with one branch of x-phi, known as its 'negative programme', in contrast to the 'positive programme', championed by Joshua Knobe amongst others.² Both positive and negative x-phi promote augmenting philosophical theories with the best available empirical data – be this from the social, psychological, historical, or even the natural sciences – but I will focus on negative x-phi due to what I view as its ability to illuminate the question of the relativity of virtue.

X-phi originated outside of practical philosophy (initially in epistemology, but also in philosophy of language and perception), so when it was first applied in moral philosophy it followed various precedents relating to its history of previous application. Many of these precedents characterise what has become known in the literature as negative x-phi. To challenge our assumptions about what our intuitions are about the answer to a specific ethical question, experimental philosophers who pursued the negative branch of their discipline designed experiments to see if their philosophical intuitions on a certain topic were shared by disparate and diverse groups. Many of these experiments yielded remarkable results. Not only did participants from various demographics respond differently to staples

² Alfano situates recent work in x-phi in the context of a longstanding tradition of connecting normative ethical claims with empirical observation in much the same way that I have invoked Aristotle, MacIntyre, and Doris above. See Alfano 2018 for an informative discussion of this, along with discussion and examples of both the positive and negative programmes of x-phi.

in moral philosophy such as trolley problems, but they even showed that intuitions that moral philosophers had relied on for generations had a great degree of variance.³

These findings, experimental philosophers argued, can be to some extent explained by the fact that the philosophical community has a relatively high degree of homogeneity when analysed under even the most rudimentary sociological concepts. Professional philosophers are still lamentably similar in terms of class, ethnicity, gender, and sexuality, and this lack of diversity in the community has been responsible for an artificially high degree of agreement regarding 'our' intuitions on philosophical matters. These intuitions were often shown to wildly diverge from those of a more representative sample of humanity. Indeed, some may say that philosophers have always been an atypical bunch, and decidedly unrepresentative of human beings as a whole. Proponents of the negative programme of x-phi think that empirical data can inform our intuitions on many topics in moral and practical philosophy, intuitions which would instead be formed from our own contingent viewpoint. By honing our intuitions with empirical knowledge on the instability of character, say, or the historical development of virtue, our philosophical theories of these topics are better safeguarded from the time- or culture-bound prejudices of those who create these theories.

Furthermore, while experimental philosophers have been good at enlisting previous empirical studies into philosophical debates, they have also speculated on the value of programmes that would direct, for example, psychological, social, or cultural research towards a specific philosophical question. Prima facie the question of whether virtues are local universal seems especially amenable to this approach. Although no experimental philosopher could claim we can *answer* this question with recourse to empirical data, those sympathetic to x-phi would regard the empirical and philosophical dimensions to this question as mutually informative, even if only by fine-tuning our intuitions regarding the extent to which virtues are *resistant* to cultural fluctuation.⁴ Even if an empirical approach does not claim to discover new virtues, by understanding what character traits *non-philosophers* think of as virtues can galvanise *philosophers* to think more deeply about the problems they approach, or even precipitate a

³ Cultural differences affecting psychological studies have been tackled since the 1950s. They are usefully summarised in a well-known 2010 article, 'The Weirdest People in the World?' by the empirical psychologists Henrich, Heine, Norenzayan. Playfully using the acronym WEIRD to stand for psychological participants that are drawn from 'Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies' (2010: 61), these psychologists argue that many of the most influential psychological studies are simply not replicable when tested with individuals who are not WEIRD.

⁴ An example of a study that is explicitly geared towards a shared and symbiotic research focus between philosophy and the empirical sciences is the recently-funded John Templeton Foundation project, *The Geography of Philosophy* (2018–21). In this project, the philosophers Edouard Machery and Stephen Stich have joined anthropologist Clark Barrett to investigate how a range of anthropological factors affect our responses to the most pressing questions in practical philosophy. This project will build on Stich's early work in how such anthropological factors affect our reasoning and responses to epistemological questions, along with that of Shaun Nicols and Jonathan Weinberg, both of whom were early pioneers in this field.

robust response that defends those character traits that have traditionally been considered to be universal virtues.

So, are there aspects of the *methodology* of x-phi – its *negative modus operandi* – that could complement the methods currently employed by practical philosophers. Also, given that the use of empirical data has already benefitted character-based practical philosophy (in the ways I have outlined above), how could these tools be applied to the question of the relativity of virtue specifically? From my above discussion, we can construct a theoretical platform to make a number of key claims regarding these questions. These can be summarised as follows:

1. An experimentally-informed philosophy of virtue (v-x-phi) does not just establish that there are different conceptions of virtue operating; it is capable of identifying aggregates of virtuous character traits which are regarded as virtues by more than one group.
2. V-x-phi can provide a targeted analysis of the factors affecting the perceived cultural relativity of virtues. Not only can the effect of each of these factors be closely examined, but such a careful targeting can reveal how these factors may interact in the formation of virtuous character traits.
3. V-x-phi can and should build on previous studies that investigate empirically-measurable factors that affect virtue. While these studies should not be considered as v-x-phi themselves, they offer incisive analyses of many different factors (anthropological, historical, technological) that deeply effect our conception of virtues.
4. V-x-phi can participate in a critical reappraisal of what philosophers refer to as their ‘intuitions’. While this is welcome in all areas of philosophy, there is reason to think that it is especially important in practical philosophy because intuitions about admirable character traits must be closely tied to the values of person who intuits them.
5. V-x-phi may lead to a number of desirable real-world social and political benefits. Using an empirically-informed research method can prevent us overlooking the value of character traits that apply traditionally marginalised groups. It is not implausible to think that there are character traits that are closely connected to the very sociological categories that these groups define themselves by, so understanding the virtues these groups admire is especially important.

There are longstanding and sophisticated reasons why moral philosophers have traditionally cardinalised certain character traits as virtues, but we stand to learn much from an empirical analysis of the character traits that are valorised as virtues by a wide range of groups. This is not to propose that philosophers need to don white coats, and begin experimenting or collecting data on far-flung virtues, because there are already copious existing resources to be used. The work of those in the natural sciences has been most closely explored in other kinds of x-phi, but as well as developing we should also consider those who have contributed to contemporary discussion of virtue with their own empirically-informed method.

4. Conclusion

From my discussion above, I hope to have shown that there are good reasons to think that practical philosophy would be benefitted by the kind of targeted empirical enquiry that has recently been called as x-phi. Much of this work is not new, but rather x-phi should be situated as part of a long-standing tradition of engaging with the empirical sciences, starting in ancient philosophy and continuing in the work of more recent thinkers, such as MacIntyre and Doris. Furthermore, the claim is not restricted to philosophers who are interested in how virtue has changed historically (MacIntyre) or how it is affected by environmental cues (Doris). Even philosophers of technology have shown that we can – indeed should – inform our discussion and analysis of virtue with respect to developments in emergent technologies (Vallor).

Describing these disparate projects as x-phi would be distort and conflate their differing and overriding philosophical concerns, however, as well as misrepresenting their potential contribution to understanding the cultural relativity of virtue. Rather these projects are best described as providing a platform upon which experimental philosophers can build a more sustained and informative conversation between philosophy and the empirical sciences. Not only have these various projects shown that it is possible to integrate empirical research with a philosophical enquiry into the virtues, they have also shown why it is valuable to do this. Put simply, virtues are context-dependant character traits that are directly constitutive of human flourishing, so understanding the context out of which they arise is essential for understanding the philosophy of virtue itself. If we neglect this contextual dimension – whether it be historical (MacIntyre), situational (Doris), or technological (Vallor) – we will not only fail to understand virtue completely, but we will also show how it must respond to the specific challenges that each of these factors present. By taking the idea that empirical research must play a constitutive role in any viable theory of virtue, then we make room for the possibilities that x-phi offers, not just in terms of developing those aspects of philosophical theories that complete them, but also in terms of selectively choosing whatever aspects of these theories that need strengthening.

References

- Annas, J. (2005). 'Comments on John Doris' *Lack of Character.*' *Philosophy and Phenomenological Research*. 71 (3): 636–42.
- Alfano, M. (2012). 'Expanding the Situationist Challenge to Responsibilist Virtue Epistemology', *Philosophical Quarterly*. 62 (247): 223–49.
- Alfano, M. (2013). *Character as Moral Fiction*. Cambridge: Cambridge University Press.
- Alfano, M. (2014, revised 2018). 'Experimental Moral Philosophy', *Stanford Encyclopedia of Philosophy*. Stanford: Stanford University Press. Accessed 11th December 2018.
- Aristotle (1984). *The Complete Works of Aristotle: Revised Oxford Translation*. Jonathan Barnes (ed.), Princeton: Princeton University Press
- Badhwar, N. (2014). 'Objectivity and Subjectivity in Theories of Well-Being'. *Philosophy and Public Policy Quarterly*. 32 (1): 23–28.
- Darley, J., & Batson, C., (1973). "'From Jerusalem to Jericho": A study of Situational and Dispositional Variables in Helping Behavior'. *Journal of Personality and Social Psychology*, 27: 100–08.
- Doris, J. (1998). 'Persons, Situations, and Virtue Ethics'. *Nous*, 32 (4): 504–40.
- Doris, J. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Henrich, J., Heine, S., Norenzayan, A. (2010). 'The Weirdest People in the World?', *Behavioural and Brain Sciences*. 33 (2):61–83.
- Knobe, J. (2003). 'Intentional action and side effects in ordinary language". *Analysis*, 63 (3): 190–94.
- Knobe, J. and Nicols, S. (2008) (eds.). 'An Experimental Philosophy Manifesto'. *Experimental Philosophy*, Oxford: Oxford University Press.
- MacIntyre, A. (1998). *A Short History of Ethics: A History of Moral Philosophy from the Homeric Age to the Twentieth Century*. South Bend, IN: University of Notre Dame Press.
- Tiberius, V. (2006). 'Well-Being: Psychological Research for Philosophers'. *Philosophy Compass*, 1 (5): 493–505.
- Tiberius, V. (2013). 'Philosophical Methods in Happiness Research'. D. Boniwell & E. Ayers (eds.), *The Oxford Handbook of Happiness*. Oxford University Press.