

Wang Yang-Ming and Contemporary Theories of Virtue and Virtue Development
Brad Cokelet
University of Kansas

This document gives the background and motivation for the work-in-progress presentation I will be giving on May 10th at the University of Chicago.

The presentation will be on the neo-Confucian philosophy Wang Yang-Ming (1472-1529) and his optimistic Neo-Confucian model of human self-perfection. It will focus on interpretive debates about how to understand his view and on how to assess it today. In short, Wang argues that all human beings have innate knowledge of good and bad and that to live well and embody true virtue we need only recognize and extend that knowledge into in all our activities and all the corners of our lifeworld. Wang explicitly pitched this optimistic, egalitarian view as an alternative to Zhu Xi's then overwhelmingly influential theory, which pictures people as initially mired in a kind of ignorance that can only be overcome by diligent study and reflection. On Zhu Xi's broadly intellectualist view – which is commonly compared to Aristotle's and Plato's – we must study texts and the things of the world to gain knowledge of the principle or principles that govern the universe. The basic idea is that we need to gain philosophical or metaphysical knowledge via education and long study before we can embody that knowledge in sagely virtue and ease. Wang rejected that potentially elitist picture, offered an egalitarian alternative, and argued that the pursuit of genuine knowledge and action are unified in a way Zhu Xi's followers missed – because they assumed that one should first aim to gain knowledge through long philosophic and textual study and only then aim to embody that knowledge in virtuous activity.

There are natural questions that arise about Wang's conception of innate knowledge and his theory of virtue and its development: Does innate knowledge include just personal knowledge of when one is good or bad, knowledge of good and bad in the world more generally, or even some kind of metaphysical knowledge of the principle or principles that govern the universe and the things in it? What does Wang mean when he says that we need to recognize and extend this innate knowledge and how does his view differ from Chan Buddhist views on which full enlightenment results when an initial sudden insight is gradually extended or embodied in thought and action? Does Wang have plausible grounds to claim that extended innate knowledge as he understands it is necessary and sufficient for sagely virtue of the sort that he and other Confucians (including Zhu Xi) commend? Can we take his views seriously today?

My presentation will focus on some of these questions and current debates about how to interpret Wang. I am interested in these exegetical controversies for their own sake, but my presentation and inquiry are also oriented by questions about whether Wang's views, or some aspects of them, can help to inform *contemporary* research into questions about the nature and development of virtue. To explain that motivation for the project, and the questions I will raise about the contemporary plausibility of Wang's views in the last part of my presentation, I want to share a bit about the larger questions that orient my research as a whole. I would also love to get leads on other thinkers that you take to be relevant to these more general, contemporary questions.

Here are the core questions that drive my work:

- What is human virtue?

- How should we understand human moral virtue?
- Who can develop virtue and moral virtue and how can they do that?
- What is the relation between moral virtue or virtues and human happiness or flourishing?

These are all complicated questions with long histories and my work explores relevant concepts and possible answers from three different angles: contemporary analytic philosophy, interdisciplinary empirical science, and the history of philosophy (broadly construed).

My historical work addresses exegetical questions about how to interpret historical figures, such as Wang, for their own sake, but it is also geared to thinking about how various historical sources can enrich contemporary thinking about the core research questions listed above. My general view, roughly following Gadamer and Bernard Williams, is that by studying Asian philosophy and other philosophically non-dominant traditions we can recognize background presuppositions that contemporary philosophers are making (without perhaps noticing) and identify contemporary alternatives – new options that come into view when we study ancient or “outsider” views that operate with different presuppositions. So, in Wang’s case, I will be discussing how to interpret his view while also trying to think about whether his account of virtue and its development could suggest novel contemporary views that could compete with the western, e.g., Kantian, Aristotelian, Thomist, and Sentimentalist, ones that dominate inquiry in philosophy and psychology.

As I have indicated, I approach questions about virtue and virtue development from both empirical and philosophic angles, and perhaps I should say a bit more to clarify the kinds of questions that I think we can better explore with cross-cultural philosophic differences in mind.

- In an empirical vein, I work with psychologists to improve the science of virtue and virtue development. The goal is to orient psychological work so that it tests (i) various speculative accounts of virtue development or cultivation that have been proposed in various philosophic and religious traditions and (ii) various claims about how the pursuit of moral virtue and personal good dovetail or do not. I also co-design empirical studies that measure the behavioral impact of virtue traits and the ways that philosophy classes impact student behavior and think these can be enriched by “outsider” ideas.
- In a philosophic vein, I work on virtue theory and virtue ethics. My main interests are in “radical” virtue ethical proposals to abandon modern moral concepts of moral obligation and right, opposing “competitive” virtue ethical proposals to ground our concepts of moral obligation and right in a virtue-based theory, and in debates about how ancient theories of virtue and personal good differ from modern virtue theories of the sort developed by Kant and contemporary consequentialists (Julia Driver and Tom Hurka). I am currently working on a new form of competitive virtue ethics that is inspired by the neo-Confucian philosopher Dai Zhen and am interested in how engagement with Asian philosophers can expand contemporary philosophic thinking about virtue and its relation to personal good.

To more concretely fill in the connection between Wang and work in empirical psychology, I am sending a book review that I recently published in the *Journal of Moral Education*. The target co-authored book is by a philosopher and two psychologists and they propose an Aristotelian model for the scientific (psychological) study of virtue and its development. In my review I argue that psychologists should be careful to not adopt contentious presuppositions about virtue

and its development without leaving room to also explore other models such as neo-Confucian ones and I mention Wang specifically.

In closing, my current work in progress, which I look forward to presenting next week, is my attempt to get clearer on the scholarly debate about how to interpret Wang and to then think about how engagement with his views can enrich our contemporary thinking both in the context of empirical science and in the context of analytic philosophy. I would also be very interested in hearing about any other thinkers that you take to be relevant to our contemporary thinking about virtue and the good life. Thanks in advance for your interest and taking the time to read this!



Virtue science and productive theoretical neutrality: Review of Wright, J. C., Warren, M., & Snow, N. *Understanding virtue*

Bradford Cokelet 

Department of Philosophy, University of Kansas, Lawrence, KS, USA

ABSTRACT

In this wise and creative book, Wright, Warren, and Snow propose a path-breaking interdisciplinary research program that promises to ground a mature science of moral virtue. Their theoretical framework and ideas for measurement are designed to guide psychologists as they study the individual traits that people have, the ways that traits interact or conflict, and the ways they change over time. While lauding the authors' impressive achievements, I criticize the contentious Aristotelian assumptions they build into their program. I argue that the science of virtue will be better served if researchers restrict themselves to more neutral assumptions and convert philosophically contentious views into competing empirical hypotheses.

KEYWORDS

Virtue; character; moral development; wisdom; methodology

Introduction

In this wise and creative book, Wright et al. (2020) propose a path-breaking interdisciplinary research program that promises to ground a mature science of moral virtue. Their theoretical framework and ideas for measurement are designed to guide psychologists as they study the individual traits that people have, the ways traits interact or conflict, and the ways they change over time. They also propose ways to assess people's overall character and note that their framework could be used to investigate genetic, cultural, educational, and economic forces that impact character development. The book will be of great use and interest to scholars and practitioners in a variety of fields including psychology, philosophy, sociology, and education.

The authors' proposed program fuses a broadly Aristotelian theory of virtue with the whole trait theory in empirical psychology. It includes numerous creative and promising proposals for designing experiments and many of those could be adapted to the study of bad traits ('vices') such as cruelty or intellectual traits such as open-mindedness or dogmatism. My critical discussion in this review will focus on the general theory of virtue that is built into the program, but before I get to that I want to emphasize how impressive and useful the proposals for empirical inquiry are. In a nutshell, the authors show that while there are many good published studies of virtue traits that provide *some* evidence when it comes to questions about moral virtue traits and their development, the empirical study of character needs, and is starting to get, a methodological upgrade. The book's systematic and detailed measurement proposals

provide a vital blueprint for this upgrade and they give us reason to believe that scientists will be able to actually answer many pressing questions about who has good character and how we can facilitate and cultivate it. In effect, they show how scientists can replace armchair speculation about who has good character and why with strong scientific evidence.

Despite its promise, I believe that the proposed program contains contentious theoretical assumptions that will hinder empirical research. My background assumption in reaching that conclusion is that empirical inquiry into virtue traits should be guided by a *productively neutral* virtue theory—a theory that avoids any and all contentious theoretical commitments that can be converted into empirical hypotheses about virtue and tested through empirical inquiry. For example, consider Kohlberg’s infamously contentious theoretical assumption that to develop and express a mature and robust trait of justice, people need to engage in broadly Kantian forms of principled practical reasoning. This is a contentious claim that should not be built into the conceptual scheme or theory guiding empirical inquiry; instead, we can and should stick with more neutral assumptions about the trait of justice and convert Kohlberg’s Kantian presuppositions about the importance of principled reasoning into empirical hypotheses that we can test along with opposing hypotheses that are suggested by various less rationalistic thinkers. Empirical testing of such competing hypotheses could find, for example, that broadly Kantian forms of practical reasoning promote robust justice traits in some subjects but that in other subjects justice traits are best or only developed by empathy or care promoting interventions of the sorts suggested by Mencius, Martin Hoffman, or Carol Gilligan. Generally, the best path forward is to develop productively neutral theories or constructs of traits such as justice and to then conduct experiments to test various contentious and interesting claims about what promotes and hinders their development and expression.

Productively neutral aspects of the Aristotelian theory

As mentioned, the authors meld their impressive blueprint for empirical investigations with a broadly Aristotelian theory of virtue and some parts of the resulting program *are* productively neutral. A prime example is the authors’ helpful and plausible distinction between three functional aspects of virtue traits: patterns of input uptake, characteristic social-cognitive processing of inputs, and outputs in the form of ‘trait relevant responses’. Compassion for example, might be thought to involve dispositions to notice when others are suffering (input), to feel sympathy and a desire to help (social-cognitive processing), and to then act beneficently (output). This three-part model is extremely useful because it gets us to ask conceptual questions about what virtue traits involve and it can guide empirical studies by highlighting various aspects of a virtue for experiments to target and measure. Moreover, the authors distinguish various aspects of the three parts and show how they can be operationalized in experiments in order to measure the dynamic inner structure of virtue traits that are embodied at different times and in different contexts (pp. 34–60, 121–187). The resulting model and the extensions they make to measuring multiple virtues, their interaction, and the ways we can model overall character and longitudinal character change are inspiring and impressive.

Things become contentious when it comes to the authors' claims about practical wisdom and its role in virtue and virtue development, but this will require some unpacking because their views about both practical wisdom and its functional importance for virtue development are quite expansive. After explaining that, I will be in a position to identify some further appealing and productively neutral aspects of their research program and better explain what I think needs to be trimmed.

First, the authors take practical wisdom to involve both developed forms of intelligent responsiveness to situations *and* a set of more strongly rational tendencies that involve things like reflection, deliberation, and reason-based judgment (Snow et al., 2021, pp. 22–30). In the former vein, they mention (1) situational comprehension or apt perception and (2) sympathy informed awareness of how it is appropriate to respond to perceived situations and people. In the latter vein, they mention (3) theoretical and practical problem-solving abilities, (4) sound deliberative reasoning about how to achieve chosen ends, (5) wise reflective choice of ends to pursue, and (6) wise reflection on one's values, one's roles and relationships, the kind of person one is and wants to be, and the narrative structure or unity of one's life as a whole. There is nothing inherently problematic about using 'practical wisdom' to refer to all these abilities or achievements, but to avoid confusion I will use 'non-deliberative practical wisdom' to refer to items one and two above and 'deliberative practical wisdom' to refer to items three through six. I should also note that my six aspect list above doesn't exactly match the five-part scheme they adopt from Aristotle; this is just an expository choice aimed at highlighting the distinction between the parts of practical wisdom that have more to do with reasoning processes and those that do not (but that may involve cognitive construals or broadly intelligent evaluative sensitivities).

Next, we need to consider the four functions that the authors take practical wisdom to serve in the manifestation and development of virtue traits: 'guiding the action of specific virtues,' 'regulating interactions among multiple virtues,' 'regulating virtuous emotions,' and 'guiding reflection on one's life as a whole' (p. 30). I believe that the first three of these are productively neutral. Anyone interested in studying good character and promoting virtue can agree that to develop robustly good character traits many if not all developing humans initially need to improve at expressing their traits, accurately hit the 'targets' of the relevant virtues, integrating or regulating their dispositions in cases where different virtues seem to point in different directions, and harboring and expressing virtuous emotions in the most appropriate fashion. For example, when children are still developing a robust trait of justice, they may need guidance in order to express their concern for justice in a productive and compassionate way or they may need to learn from experience in order to develop the courage to act on their sound sense of justice.

With these parts of the authors' views on the table, I can more fully identify the productively neutral part of their theory of virtue. In short, the authors hold that practical wisdom is vital for virtue expression and development because it provides needed action guidance, multiple virtue regulation, emotional regulation, and reflection on one's values and life as a whole, but I think they should pare this down to the claim that *non-deliberative* practical wisdom can help provide action guidance, multiple virtue regulation, and emotional regulation. Of course, in some situations we might doubt that our virtues need guidance or regulation, but when we do, it is plausible to assume that agents can develop and rely on situational comprehension or apt perception and sympathy

informed awareness of how it is appropriate to respond to perceived situations and people, in order to achieve action guidance, multiple virtue regulation, and emotional regulation. So, a productively neutral virtue theory can assume that non-deliberative practical wisdom plays those roles.

Contentious aspects of the Aristotelian theory

Things get more problematic when it comes to the more robustly rational and deliberative aspects of practical wisdom and the authors' apparent assumption that to develop overall virtue, practical wisdom needs to guide 'reflection on one's life as a whole'. These Aristotelian claims are philosophically contentious in ways that are analogous to the broadly Kantian assumptions that Kohlberg built into his research program, so I suggest that the authors should delete them from the theory of virtue that guides inquiry and convert them into empirical hypotheses. Of course, even if this is right, some educators, parents, and others aiming to cultivate virtue may still rightly promote the reflective and deliberative aspects of practical wisdom in hopes of catalyzing moral character development. Given our limited current evidence about what will and will not promote development, there is nothing wrong with relying on the personal experience, tradition, and other non-scientific sources for guidance in at least some practical domains. My point is that when it comes to science, we should aim for productively neutral research programs that can help us identify and test competing hypotheses about what promotes virtue effectively and what does not. We should value a productively neutral research program because it will give us more knowledge but also because it promises to give educators, parents, and so forth robust evidence-based guidance when it comes to questions about how to promote compassion, justice, honesty, and overall good character.

I will illustrate my worries about the authors' program by noting three contentious Aristotelian assumptions about virtue that they build into their theory. The first is the assumption that deliberative practical wisdom is the best mechanism to help agents when it comes to the needs for action guidance, multiple virtue regulation, and emotional regulation. Anti-rationalist philosophers and psychologists as diverse as David Hume, Michael Slote, Julia Driver, and Jonathan Haidt cast doubt on such Aristotelian claims about practical wisdom; they doubt that rational deliberation and reflection can shape our emotions and actions in the ways that Aristotelians hold and they tend to think that being virtuous is within the reach of uneducated or poorly educated people who lack the kinds of rationality and conceptual sophistication that seem to be built into Aristotelian wisdom. In addition, many of these thinkers propose alternative mechanisms to serve the needs for action guidance, multiple virtue regulation, and emotional regulation. They point, for example, to extended forms of sympathy or empathy and to various complicated forms of perspective taking that need not involve deliberative, rational excellence or conceptually sophisticated articulation.

Of course, the mere fact that there is philosophic debate on this score does not show that the authors' assumptions about deliberative practical wisdom are wrong, but it highlights why it will be unproductive to adopt a program of empirical research that assumes as a matter of background theory that one side is right and another is wrong. It will be more scientifically productive to trim the contentious Aristotelian assumptions

about deliberative practical wisdom from the theory of virtue that guides inquiry and to convert them into a set of empirical hypotheses that can be tested along with a set of opposing 'anti-rationalist' alternatives.

For an example of how this would play out in detail, consider the authors' useful idea that we could present subjects with vignettes in which the demands of various virtues are in apparent conflict. They seem to assume that practical wisdom is the best way for subjects to regulate or adjudicate the apparent interaction or tension in such cases; they say 'practical wisdom must help her navigate the situation' (p. 152) and to assess people's degree of wisdom, and hence virtue, they suggest that experimenters track how people reflectively think about the situations as new nuances are introduced. The assumption seems to be that people who are 'able to successfully articulate how and why they would respond to [changes in virtue-relevant stimuli] and why they would respond to them in the way they've reported' (p. 153) are more likely to navigate actual situations of virtue conflict better. (Cf., Snow et al., 2021, pp. 71–2). But anti-rationalists would surely disagree and it is not something that we can assume as a conceptual matter before we do experiments; we need a neutral theory that enables us to identify and test competing hypotheses.

Next, consider the authors' suggestion that overall virtue and virtue development can be best promoted by reflection on one's values, one's roles, virtues one has or aspires to have, and the narrative structure of one's life as a whole (for example, pp. 27–29, 47–48, 278–291). Once again, I think that comparison of competing philosophic theories of virtue gives us reason to reduce how much is built into the theory of virtue and to increase our stock of competing empirical hypotheses. Consider, for example, the lively and complex debate among Chinese neo-Confucian philosophers when it comes to virtue and virtue development (see, Angle & Tiwald, 2017, especially chapters six and seven). Very roughly, some philosophers such as Zhu Xi thought virtue could be best developed by intellectual study of the moral exemplars illustrated in various canonical texts, by comparative rational reflection on one's virtues or vices, and by deliberate attempts to then instantiate ideals derived from exemplary models. This approach seems to fit hand in glove with the Aristotelian model of virtue and virtue development that our authors recommend. But this approach was strongly rejected by philosophers such as Wang Yangming, who thought it was a mistake to divide virtue development into a two-aspect process involving the development of intellectual rational knowledge and then the internalization or application of that knowledge in action. He held that human beings have an innate ability to know right from wrong and that the task of virtue cultivation should focus on eliminating things like selfish desires or self-centered psychological tendencies that block our natural tendency to act on our knowledge of what is morally best. So, he would presumably reject our authors' assumption that practical wisdom as they understand it is needed, or would be useful, for virtue development and suggest an alternative. And when it comes to the actual practice of virtue cultivation he would suggest that educators, parents, and others interested in virtue cultivation should adopt a set of strategies that will promote awareness of and expression of innate moral knowledge, and that they should drop any existing attempts to promote Aristotelian practical wisdom of the sort that involves reflection on exemplars and self-reflection on an ideal that one hopes to instantiate through reflective self-assessment and deliberative planning over time. Here again, my thought is that we should convert contentious aspects of the Aristotelian model into empirical hypotheses, look to the philosophic traditions to develop alternative hypotheses, and to then pursue virtue science based

on a neutral model to determine which view is correct (keeping in mind that it could be both or that the answers could be different for different subject populations). As things stand now, we lack the relevant scientific evidence and need to build neutral theories in order to gather it.

Finally, consider the authors' apparent Aristotelian assumption that wise evaluative self-reflection in a broad sense—for example, reflecting on the narrative unity of one's life as a whole, on one's values, and on the roles and relations one inhabits—will be conducive to the development of overall virtue or good character. Before stating my doubts on this score, it is worth mentioning a useful and wise section of the book that discusses philosophic debates about the unity of virtue and the kind of overall virtue that one must exhibit to have good character (pp. 213–245). The authors discuss a variety of views that they take to be overly strong on these topics, and introduce a moderate view that strikes me as very sensible, namely that 'ideally a virtuous person would possess all the virtues necessary to flourish and live a good life within her own cultural/historical contingencies' (p. 228). This view plausibly allows, for example, that people with overall good character may have only a subset of the maximal set of possible virtues, depending on the contingent contours of their lifeworld. In addition, the authors have also recently and plausibly emphasized that less than virtuous people can possess or develop the 'critical thinking capacities' that will enable them to reason in the ways that are characteristic of deliberative practical wisdom as they understand it (Snow et al., 2021, p. 76).

I remain skeptical however, about the authors' theoretical presupposition that the broad evaluative self-reflection that they take to be part of practical wisdom will be effective and is perhaps a necessary means to achieving overall good character, and their apparent assumption that there is a conceptual connection between having overall good moral character and living a good or flourishing life. Of course, there are various philosophers and psychologists east and west who agree with those views, but there are others who defend contrary views, and in this case the bone of contention is not *just* about whether deliberative or reflective rationality will be an effective, or the best, mechanism for the promotion of overall virtue. While it is true that philosophers such as Wang Yangming (and many psychoanalysts) would doubt that broad Aristotelian evaluative self-reflection is an efficacious mechanism for overall virtue development and they would propose alternative mechanisms, there are also philosophers, such as Nietzsche, Bernard Williams and Susan Wolf, and psychologists, such as Freud, who raise doubts about whether overall good moral character is in fact always conducive to living a good and flourishing life or to being a good parent, friend, or human being. As in other cases, these contentious views should be converted to empirical hypotheses and not taken as given in empirical inquiry.

In addition, awareness of the philosophic debate about the relation between morality and the good life should lead us to doubt the empirical hypothesis that broad evaluative self-reflection will reliably promote overall good character. Presumably many non-philosophers will, like Nietzsche, Williams, and Wolf, conclude their self-reflections with the view that becoming more overall virtuous will impede their flourishing or their ability to have a life with the kind of narrative unity that they want. So even if their episodes of rational self-examination do shape their evolving character, they will not promote an increase in overall moral virtue. For example, it is easy to imagine some successful people in various professions thinking that they would have failed more if they had been more just or compassionate and concluding that they are better off being decent but not overall virtuous. Or, for slightly

different reasons, it is likely that some parents and friends will conclude that they are fulfilling their relational roles well precisely because they are less than ideally virtuous. In Wolf's (1997) terms, these people may conclude that becoming more virtuous would threaten their ability to characteristically do well in the projects and relationships that give meaning, and narrative unity, to their lives.

Summing up, I have been arguing that the authors' path-breaking program for virtue science is very promising but that they should trim back the Aristotelian theory of virtue that is built into their research program and replace it with a more productively neutral alternative. In some places in the book, I think the authors may tacitly recognize this point as they sometimes talk about some of their Aristotelian assumptions as if they are open to empirical validation if not falsification, so I should end by emphasizing that my call to self-consciously adopt a productively neutral virtue theory is something they could take on board while keeping almost all of the book's rich and creative content. It would mainly involve more clearly distinguishing between the philosophical views that can be rightly built into the theory that guides inquiry and those that should be converted into empirical hypotheses (and contrasted with ones inspired by opposing philosophic theories). The authors have already laid the groundwork for this sort of approach with their creative and impressive interdisciplinary reflections on Aristotelian virtue theory and the ways a mature science of virtue could test the predictions it suggests. This work will no doubt inspire and orient future empirical researchers and it also provides a model that researchers with different philosophic leanings can fruitfully emulate.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Bradford Cokelet works on analytic moral philosophy, the empirical study of virtue and philosophy instruction, and Chinese philosophy. He has published papers in *Ethics*, *Oxford Studies in Normative Ethics*, *Cognition*, *Perspectives on Psychological Science*, *the European Journal for the Philosophy of Religion*, and *the Journal of Confucian Philosophy and Culture*.

ORCID

Bradford Cokelet  <http://orcid.org/0000-0001-8497-0669>

References

- Angle, S., & Tiwald, J. (2017). *Neo-confucianism: A philosophical introduction*. Polity Press.
- Snow, N., Wright, J. C., & Warren, M. T. (2021). *Phronesis and whole-trait theory: An integration*. In M. S. Vaccarezza & M. De Caro (Eds.), *Practical wisdom: Philosophical and psychological perspectives* (pp. 70–95). Routledge.
- Wolf, S. (1997). Meaning and morality. *Proceedings of the Aristotelian Society*, 97(3), 299–315. <https://doi.org/10.1111/1467-9264.00018>
- Wright, J., Warren, M., & Snow, N. (2020). *Understanding virtue: Theory and measurement*. Oxford University Press.