

The Introspective Model of Genuine Knowledge in Wang Yangming*

Harvey Lederman

August 12, 2021

Abstract

This paper presents a new interpretation of the great Ming dynasty philosopher Wang Yangming's (1472-1529) celebrated doctrine of the "unity of knowledge and action" (知行合一). Wang held that action was not unified with all knowledge, but only with an elevated form of knowledge, which he sometimes called "genuine knowledge" (真知). I argue for a new interpretation of this notion, according to which genuine knowledge requires freedom from a form of doxastic conflict. I propose that, in Wang's view, a person is free from this form of doxastic conflict if and only if they are acting virtuously.

Keywords: unity of knowledge and action, genuine knowledge, conscience, *liangzhi*, akrasia, Wang Yangming, Neo-Confucianism, Chinese philosophy, moral psychology, moral epistemology

1 Introduction

In 1508, in exile in Longchang in Guiyang, Wang Shouren (王守仁, Yangming 陽明 1472-1529) experienced a "great enlightenment" (大悟), when a voice seemed to call out to

*Thanks to Ernest Brewster, Amin Ebrahimi, Johann Frick, Jane Friedman, Don Garrett, Sabrina Hao, Yingtian He, Grace Helton, Jing Huang, P. J. Ivanhoe, Jessica Moss, Anna Shields, Winnie Sung, Kwong-loi Shun, Justin Tiwald, an editor and three referees, for written comments; to audiences at the Columbia Society for Comparative Philosophy, the University of Pittsburgh, USC, the Bay Area Conference in Chinese Thought, the Association of Chinese Philosophers in America, Nanjing University, McGill, Cosmopolitan Philosophy at Berkeley, and a graduate seminar at Princeton for their questions; to Yingtian He for research assistance; and to Stephen Angle, Jennifer Eichman, Jeremy Goodman, Stephen Menn and Snow Zhang for conversations and correspondence. I wish to say thanks especially to Don Garrett, whose detailed comments on an early draft were an inflection point for the paper; to P. J. Ivanhoe and Justin Tiwald, whose thorough and generous comments on multiple drafts saved me from dozens of errors small and large; to Jane Friedman, without whose encouragement and advice I might not have seen the project through; and to Ernest Brewster, for dozens of conversations on related topics over the years, including those which led me to work on Wang Yangming in the first place.

him in the night. In the following year, Wang distilled this dramatic revelation in the doctrine of the “unity of knowledge and action” (*zhi xing he yi* 知行合一).¹ This doctrine would come to be seen as one of the major achievements of Ming dynasty (1368-1644) thought, and indeed, of the whole tradition now called “Confucian”. It is a central part of the distinctive philosophical outlook which has earned Wang a place on the standard list of the four most important thinkers in this tradition, alongside Confucius, Mencius, and the Song dynasty philosopher Zhu Xi (朱熹, 1130-1200) (“孔孟朱王”).

The unity of knowledge and action is one of the most celebrated doctrines in all of Chinese thought. But it has also attracted trenchant criticism. In letters and recorded conversations, Wang’s interlocutors present him with a series of examples which they take to show that knowledge and action are not, in fact, unified. In his replies to these objections, Wang implicitly qualifies the claim that knowledge and action are unified, stating instead that action is unified only with an elevated form of knowledge, which he sometimes calls “genuine knowledge” (*zhen zhi* 真知). The late Ming materialist Wang Fuzhi (王夫之, 1619-1692) presumably had this move in mind when he complained that Wang Yangming had simply changed the topic: “what [he] calls ‘knowledge’ is not knowledge”.² This now famous complaint can be developed into a pointed criticism. Wang Yangming’s replies to his contemporaries naturally suggest that he sought to defend the unity of knowledge and action by stipulating that “genuine knowledge” be understood as “whatever is unified with action”. But, our critic might observe, while such a stipulation would indeed allow Wang to reject putative counterexamples to his doctrine, it would do so at the expense rendering the doctrine trivial. Far from the epoch-making innovation in moral psychology and the theory of action which it has seemed to so many to be, the “unity of knowledge and action” would be nothing more than a successful marketing trick, a misleading advertisement for a triviality.

I will argue that this criticism is mistaken. I will develop a new interpretation of genuine knowledge, according to which Wang characterizes it as an elevated form of

¹I follow tradition in translating *he yi* 合一 as “unity” in this slogan. The expression can also mean something weaker, more like “correspondence”. Qian Dehong (錢德洪, 1496-1574) gives the standard account of Wang’s “enlightenment” in Wang’s *nian pu*; see Wu *et al.* (2011, 1354-5).

Throughout the paper, I will cite passages from the *Instructions for Practical Living* (hereafter, *IPL*, 傳習錄) by the section number of Chan’s editions (Chan (1963), Chan (1983)), followed by a page number in Wu *et al.* (2011) (indicated by “QJ”). Passages in Wang’s works outside the *IPL* are cited by the *juan* number and page number (e.g. “QJ 6.242”). Where available, I also cite pages in the translations of Ching (1972) and Ivanhoe (2009). Most of Ivanhoe’s translations are reprinted, but with some significant changes, in Tiwald & Van Norden (2014). I recommend that the reader consult the amended versions where possible, but since they do not cover all of the material translated in Ivanhoe (2009), I cite page numbers in the earlier book here.

²其所謂知著非知, Wang (1976, 76), see Chen (1991, Ch. 5.3) for discussion.

knowledge, independently of its relationship to action. In my view, Wang holds that a person has genuine knowledge if and only if they are free from a particular form of doxastic conflict. I will suggest that Wang connects freedom from this form of doxastic conflict to freedom from a particular form of motivational conflict, and connects freedom from this form of motivational conflict, in turn, to virtuous action. The result is that Wang does not simply stipulate the truth of his doctrine. Instead, under the heading of the unity of knowledge and action, he advances striking, substantive claims connecting freedom from doxastic conflict to virtuous action.

My interpretation of genuine knowledge stands in stark contrast to arguably the most prominent interpretation of this notion in the English-language scholarship. On this prominent view, genuine knowledge sometimes consists in part in a person's apprehension of features of their environment.³ In line with this interpretation, the unity of knowledge and action is taken to concern how an ideally virtuous person will respond seamlessly to any situation in which they might find themselves, effortlessly producing an action which is appropriate to what they see around them. On my *introspective model* of genuine knowledge, by contrast, genuine knowledge is a form of knowledge of one's own mind: a person's apprehension of their environment is never any part at all of genuine knowledge.⁴ In line with this interpretation, the unity of knowledge and action does not concern a virtuous person's ability to respond to their environment, but instead a series of proposed connections between doxastic coherence, motivational coherence, and virtuous action.

Section 2 introduces some background and refines the challenge that my interpretation of genuine knowledge will seek to address. Section 3 examines Wang's views about the conscience-like faculty of *liangzhi*. I argue that *liangzhi* invariably acquires knowledge that is importantly related to genuine knowledge, but that this knowledge does not always amount to genuine knowledge. Section 4 discusses a previously unidentified argument in Wang's writings. This argument, which I call the *obscuration argument*,

³For remarks along these lines, see Nivison (1973, 132) (reprinted in Nivison (1996, 243)), Nivison (1973, 134) (reprinted in Nivison (1996, 244)), Cua (1982, 9-14), (Ivanhoe, 2002, 99), Ivanhoe (2009, 113), Ivanhoe (2011, 274), Angle (2005), Angle (2010). The position of Shun (2011) is subtler, though I believe it still implies the claim in the main text. For reasons of space, I have been unable to include thematic discussion of others' interpretations of the unity of knowledge and action in this paper. A companion paper (Lederman (forthcoming)) defends my attribution of this position to the authors listed above and discusses their views in more detail. The two papers are intended to be self-contained, but specialists may wish to read them together.

⁴I use the word "introspection" and its cognates to describe any direct knowledge a person has of their own mind. This technical usage, which is standard in contemporary analytic philosophy, is somewhat broader than the use of the word in ordinary English, where "introspection" is often reserved for an effortful, conscious process of looking inwards. As my discussion in section 3 will make clear, much of the knowledge I describe as "introspective" is crucially acquired automatically and effortlessly.

describes what more is required, beyond *liangzhi*'s knowledge, for a person to have an elevated form of knowledge related to genuine knowledge. Section 5 completes the presentation of my *introspective model* of genuine knowledge, and offers my response to the main challenge of the paper. I show how Wang conceives of genuine knowledge as a form of knowledge, and how it is elevated above ordinary knowledge in a distinctively epistemic or doxastic respect. Section 6 describes how my interpretation of genuine knowledge fits into a new understanding of the unity of knowledge and action as a whole.⁵

2 Knowledge and genuine knowledge

The main aim of this paper is to develop an interpretation of genuine knowledge. To set the stage for this interpretation, this section begins with three preliminaries: first, about the Chinese word I will translate as “knowledge”; second, about the texts I take to be relevant for understanding “genuine knowledge”; and, third, about the objects of genuine knowledge. At the close of the section, and in light of these preliminaries, I present a sharper version of the challenge which the rest of the paper will address.

⁵A word about chronology and the scope of this paper: Wang first advocated the unity of knowledge and action in 1509, a year after his “enlightenment” in Longchang. Wang’s articulation of his views underwent an important shift twelve years later, in 1521, when he first began to emphasize the importance of *liangzhi* (see Qian Dehong’s account of this year in the *nian pu* (QJ 34.141) along with his famous account of the “three turns” (QJ 41.1745-6), cf. e.g. Tu (1976, 10-11), Ching (1976, 41-46)). But Wang continued to endorse the unity of knowledge and action explicitly after 1521, and at least as late as 1526; moreover, there is no evidence that he retracted it before his death in 1529. (The latest explicit mentions I’m aware of are in “Reply to Inquiries from a Friend” (1526) QJ 6.232; Ching (1972, 106-8) and Ivanhoe (2009, 123-7) and in “Letter to Lin Sixun” QJ 8.314. For other post-1521 discussion, see IPL 133 QJ 48, IPL 139 QJ 56 (Chan (1963, 91 n. 1) discusses difficulties with dating this latter); “Second Letter to Lu Yuanjing” (1522) QJ 5.210 Ching (1972, 68-9); “Letter to Zhu Yangbo” (1524) QJ 8.309.)

My aim in this paper will be to explicate the views Wang held about the unity of knowledge and action after 1521. In developing my interpretation, I will freely take conversations and writings prior to 1521 as evidence for Wang’s views after this date. This practice is justified by my belief that, although Wang may not have fully developed his later ideas before 1521, his early views were at least consistent with those he would come to hold later. There is strong evidence that Wang himself understood his philosophical development in this way. First, Wang did not retract or seek to revise a collection of his sayings published in 1518; nor did he intervene when that same collection was republished together with a number of letters in 1524 (for English works on the publication history of the IPL see, Chan (1963, 314), Ivanhoe (2002, Appendix I); for the later publication history of the complete works, see Chu (1988)). Second, Wang is recorded as saying: “From Longchang on [the site of his “enlightenment”], I have not departed from the meaning of the two characters ‘*liangzhi*’. It’s just that I was unable to produce these two characters in speaking to students, and wasted many words describing it. Now, fortunately, this meaning has been made manifest, so that in one expression, one can see clearly the whole substance.” (吾『良知』二字，自龍場已後，便已不出此意，只是點此二字不出，於學者言，費卻多少辭說。今幸見出此意，一語之下，洞見全體 QJ 41.1747) (All translations in the paper are mine, although I have always consulted Chan (1963), Ching (1972) and Ivanhoe (2009) for passages translated in those works.)

First, then, some background on “knowledge”: When it occurs as a free-standing semantic unit, I will translate the character *zhi* 知 by “knowledge” and its cognates, as in “the unity of knowledge and action”.⁶ If Wang had wanted to translate the English “Wei knows that Xin loves Yun”, he would have used the character *zhi*: this character can mean “know”. But the character has a broader semantic range than “know”. It can be used to describe a change of state, and in this usage it is naturally translated as “realize” or “recognize”. *Zhi* can also be used without a complement, and in this usage it can be rendered as “is conscious”, as in the sentence “Wei is conscious again, after months in a coma”. In one passage, for example, Wang is asked whether people in a dreamless sleep still *zhi*; it is natural to understand the interlocutor to be curious about whether such people are in any sense conscious (*IPL* 267, *QJ* 120).

Wang’s theoretical remarks about *zhi* often seem to pick up on the latter two uses of the character *zhi* (“recognize”, “be conscious”) rather than on its use to describe a long-lasting state, making “know” not always a perfect fit for *zhi* in the texts I will discuss. In a battery of important passages, for instance, Wang describes the relationship between *zhi* and what he calls *yi* (意), a term which I will translate as “inclination”, but which is often rendered as “will”, “intention” or “thought”.⁷ In all of these passages, Wang describes inclinations (*yi* 意) as short-lived mental episodes, which “are aroused” or “are moved”; they are not long-lasting states.⁸ In two of them, Wang characterizes inclinations as the “mind when it is aroused” (心之發動), and then describes *zhi* as a property or aspect of such inclinations (*IPL* 174 *QJ* 86-7; *IPL* 201 *QJ* 103).⁹ When Wang

⁶For detailed discussion of the meaning of “know” in classical, pre-Han (before 202 BCE) texts see Harbsmeier (1993). Geaney (2002) is an important treatment of the epistemology of sense perception in that period; Fraser (2011) argues for a particular conception of knowledge among the Mohists and in the *Xunzi* (both also pre-Han) and ties this conception to an alleged relative lack of discussion of skeptical arguments in this period. Angle & Tiwald (2017, Ch. 6) and now Angle (2018) argue that Zhu Xi held there were at least three forms of knowledge, which he ranked in terms of cognitive achievement and practical importance.

⁷*IPL* 6 *QJ* 6; *IPL* 78 *QJ* 27; *IPL* 137 *QJ* 53; *IPL* 174 *QJ* 86-7; *IPL* 201 *QJ* 103.

⁸It is not clear to me that this aspect of Wang’s theoretical view of *yi* corresponds with ordinary usage of the word in the literary Chinese of Wang’s day, but it is clear that it differs from what the English word “inclination” describes. If Wei has an inclination to visit his family over the new year, he may have that inclination over a period of months or longer, and regardless of whether he is considering his plans or not at a given moment. Wang’s theoretical remarks about *yi* suggest that for him, *yi* differ from inclinations in this respect: he seems to take them to be mental episodes, which arise and disappear fairly quickly, rather than longstanding states. The translations of *yi*, as “intention” or “will” also do not capture this aspect of Wang’s usage. Translating the term as “thought” would avoid this problem, but would not adequately capture the conative aspect of *yi* in the uses of this character which will be most important below.

⁹In the other passages, Wang instead describes *zhi* as the “original natural condition” of inclinations (*benti* 本體, *IPL* 6 *QJ* 6), takes inclinations to be “the arousal” of *zhi* (發動 *IPL* 78 *QJ* 27) and says that inclinations are *liangzhi* when “it is stimulated, responsive and moving” (感應而動 *IPL* 137 *QJ* 53). These passages show that he can also use the character *zhi* to describe a capacity for producing the episodes I describe in the main text. For more discussion of this point see n. 41.

says that *zhi* arises as a feature of the short-lived episodes of *yi*, he must be thinking of *zhi* itself as a short-lived episode. These passages thus strongly suggest that when Wang was attending closely to *zhi* in theoretical contexts, he was at least sometimes interested in episodes of recognition, or perhaps something more like episodes of knowledgeably considering or grasping. In spite of this fact, I will continue to translate *zhi* as “know” throughout, although in my explanations, I will sometimes use the expression “episode of knowing” to highlight that Wang is focused on short-lived episodes of the kind just described.

Second, some background on “genuine knowledge”: My term “genuine knowledge” translates an expression composed of two characters, the first of which can be translated as “genuine”, “real”, “true”, or “authentic” (*zhen* 真), and the second of which is the character *zhi* I have just been discussing. In a number of passages, Wang uses this expression to describe an elevated form of knowledge connected to the unity of knowledge and action.¹⁰ But Wang can also use a different expression, *zhizhi* (致知) – which, in the relevant contexts, I will translate as “extended knowledge” – in a very similar way.¹¹ Wang uses “extended knowledge” and its cognates much more broadly than he uses “genuine knowledge”: he can use “extended knowledge” to describe mental states or events that are not relevant to the unity of knowledge and action. But it is natural to suppose that in the passages where Wang does connect extended knowledge to the unity of knowledge and action, he means to describe the same elevated form of knowledge that he describes with “genuine knowledge”. Thus, while I will use “genuine knowledge” throughout as my technical term for this notion, I take uses of “extended knowledge” which are explicitly tied to the unity of knowledge and action

¹⁰*IPL* 5 *QJ* 4; *IPL* 5 *QJ* 5 (知得真); *IPL* 125 *QJ* 42 (where it is used by a disciple, not by Wang himself); *IPL* 133 *QJ* 47-8. (The occurrences of the expression in *IPL* 134 *QJ* 49 and *QJ* 20.829 are not related to the unity of knowledge and action.) In two different texts Wang ties what seems to be the set phrase “Insofar as knowledge is genuine, practical, earnest and substantial, it is action; insofar as action is lucidly aware and precisely discriminating, it is knowledge” (知之真切篤實處，便是行；行之明覺精察處，便是知。) to the unity of knowledge and action (*IPL* 133 *QJ* 48; *QJ* 6.233, 234; Ching (1972, 106-8)). In a few places, Wang also speaks of an advanced stage of ethical training being associated with *liangzhi* itself being “genuine and practical” (真切) (*IPL* 241 *QJ* 114; *QJ* 6.238). In these passages, Wang describes a property of the faculty of *liangzhi*, not of the episodes of knowledge produced by that faculty, but it is natural to think that such a property of the faculty would carry over to the episodes produced by it. In *IPL* 170 *QJ* 83, Wang connects the achievement of this genuineness and practicality to the extension of *liangzhi*, a notion he elsewhere (see next note) ties to the unity of knowledge and action.

¹¹*IPL* 139 *QJ* 56; *IPL* 140, *QJ* 58; *IPL* 321, *QJ* 137; *QJ* 5.211, Ching (1972, pp. 68-9); *QJ* 8.308; and *QJ* 27.1100. A rougher connection is drawn in: *QJ* 6.234, Ching (1972, 106-8). The more common use of *zhizhi* 致知 is as a gerund or a verb phrase, which I translate as “extending knowledge” (sometimes “the extension of knowledge”) or “to extend knowledge” (and variants thereof). In this more common use, the expression describes a process or activity. In the passages I am most interested in, however, the phrase is used to describe the successful result of this process.

to describe genuine knowledge as well.¹²

Third, and finally, some background on the objects of genuine knowledge: The word *zhi*, like the English word “know”, can describe knowledge *prima facie* different kinds of knowledge, depending on the syntactic type of its complement: if it takes a sentential complement, it describes propositional knowledge (“know that”),¹³ if it takes a simple noun-phrase as its complement (“arithmetic”), it describes objectual knowledge (“know arithmetic”); and, finally, if it takes a verb-phrase as its complement (“run”, “say thank you”), it describes knowing how (“know how to run”), or knowing to (“know to say thank you”). Which of these kinds of knowledge does Wang take genuine knowledge to be?

I will now argue that he understands genuine knowledge as objectual knowledge. In perhaps the most famous discussion of the unity of knowledge and action (*IPL* 5, *QJ* 4), Wang uses “know” first with the character for “filial piety” (*xiao* 孝, hereafter “filiality”) and then with the character for “fraternal respect” (*ti* 悌, hereafter “respect”) as its complement. One might think that this fact alone would show that Wang understands genuine knowledge as objectual knowledge. But the characters for “filiality” and “respect” can be used in classical Chinese both as abstract nouns (e.g. “filiality”) and as verb-phrases (e.g. “be filial”). So the grammar here does not in fact determine whether Wang intends to describe objectual knowledge (“know filiality”), knowledge-how (“know how to be filial”), or knowledge-to (“know to be filial”).¹⁴

But even though the grammar does not settle this question on its own, the broader context tells clearly in favor of objectual knowledge. In the passage, Wang presents seven examples, in three groups, to illustrate the way in which genuine knowledge

¹²In some places Wang uses the expression “extending knowledge” generally to describe cultivating any capacity of *liangzhi* (see next section for discussion of this notion). But these capacities are very broad. For instance, in *IPL* 171 (*QJ* 83-4) Wang talks about extending *liangzhi* in connection to anticipating others’ actions (i.e. knowing what they will do in the future) on the basis of their present intentions. The criterion given in the main text does not imply that the “extended knowledge” described in this passage is genuine knowledge, and indeed it is natural to suppose that it is not, in part because it is unclear how it would fit with Wang’s understanding of the unity of knowledge and action itself (for instance, it would not satisfy anything like the principle Unity discussed in section 6).

¹³A common way of expressing propositional knowledge in classical Chinese involves a special construction where a nominalized sentence is the complement of the verb “know”. At least in Wang’s corpus, there does not seem to me a significant semantic difference between this construction and those in which the complement is an un-nominalized sentence. See again Harbsmeier (1993) for discussion of a related issue in pre-Han texts.

¹⁴It is in some sense linguistically possible that Wang means to describe propositional knowledge (“know that one is filial”) but that construal makes little sense in context and I know of no one who has defended such a construal, so I won’t discuss it further here.

For favorable discussion of the knowing-how construal, see Chen (1991, §5.3), Ivanhoe (2000, 71, n. 15) Yu (2014), Yu (2016), Shi (2017). For the knowing-to construal, see Huang (2016b) (translate into English as Huang (2017)) and Huang (2016a).

is connected to action: knowledge of sights (*se* 色) and odors (*chou* 臭); knowledge of filiality and respect; and knowledge of pain (*tong* 痛), cold (*han* 寒) and hunger (*ji* 饑). There are strong arguments that in the first and third groups of examples, Wang describes objectual knowledge. The case is clearest for the examples in the first group (sight and odor). Unlike the characters for “filiality” and “respect”, the characters for “sight” and “odor” must be nouns here: the only grammatical possibility is that the relevant expressions describe objectual knowledge (“see a lovely sight” 見好色, “smell a hateful odor” 聞惡臭 and “know the odor” 知臭). For the third group (pain, cold and hunger) the argument is more involved, but still conclusive. In these cases, a construal as “knowledge-how” or “knowledge-to” is grammatically possible: like the characters for “filiality” and “respect”, the characters for “pain”, “cold” and “hunger” can be used both as nouns and as stand-alone verb-phrases. But in the passage, Wang says that a person can know pain, cold and hunger, only by being pained, cold or hungry, and this idea is hard to make sense of if “pain”, “cold” and “hunger” are taken as verb phrases, regardless of whether “know” is then understood to describe knowledge-how or knowledge-to. For knowledge-how: to the extent that I can understand what it would mean to say that someone knows how to be cold in the relevant sense, surely what is surprising is that infants know how to be cold before they have ever been cold; it seems obviously false to say that in this sense people know how to be cold only by being cold. For knowledge-to: again, to the extent that I can understand what it would mean to say that someone knows to be cold in the relevant sense (as “they know when to feel cold”), surely what is surprising is that infants know (when) to be cold, before they have ever been cold; it seems obviously false to say that they know (when) to be cold only by being cold. By contrast to the odd interpretations which result by taking the characters for pain, cold and hunger as verb-phrases, Wang’s idea is straightforward and intuitive if these characters are taken to be nouns. On this construal, his point is that one does not know cold, pain or hunger – in the sense of being intimately acquainted with these conditions – until one has experienced them.

The objectual construal is thus clearly the correct one for the first and third groups of examples. Since Wang presents these examples to illustrate the character of genuine knowledge, we should prefer an interpretation which gives a uniform sense to all seven of the examples. So, we should take Wang to be describing objectual knowledge of filiality and respect as well. Moreover, given that Wang takes genuine knowledge to be objectual knowledge in this important passage, it is natural to think that he takes genuine knowledge in general, too, to be objectual knowledge.

But what does Wang take the objects of this objectual knowledge to be? In addition

to filiality and respect, Wang also speaks of conscientiousness (*zhong* 忠), humaneness (*ren* 仁), and compassion (*ce yin* 惻隱) in close connection to the unity of knowledge and action.¹⁵ For simplicity, in what follows, I will exclusively discuss filiality as my running example. But I intend my analysis to apply in addition to respect, conscientiousness, humaneness, and compassion, since Wang quite clearly thinks these too can be the objects of genuine knowledge. Beyond these five examples, however, I will not have much to say: I will not take a stand here on whether Wang holds that there are further possible objects of genuine knowledge, or on what these objects might be (though n. 47 contains some further discussion).

As promised at the start of this section, these preliminaries allow us to refine the main challenge of the paper. The challenge will not just be to elucidate a sense in which genuine knowledge is an elevated form of knowledge in general, but in particular to elucidate a sense in which genuine knowledge is an elevated form of objectual knowledge, and in particular, objectual knowledge of filiality, respect, conscientiousness, humaneness or compassion.

3 *Liangzhi* and knowledge of ethical qualities

One of Wang's most important philosophical innovations was his doctrine of the "extension of *liangzhi*". Indeed, later in his life, Wang saw his ideas about the conscience-like faculty of *liangzhi* as encompassing all of his philosophical system (see n. 5).¹⁶ An important component of Wang's understanding of *liangzhi* was the idea that *liangzhi* can acquire and possess ethically relevant knowledge. Given this, and the centrality of *liangzhi* in Wang's thought, one might wonder whether Wang held that the ethically relevant knowledge a person's *liangzhi* acquires just is genuine knowledge. In my view, this idea contains an important seed of truth: my own interpretation will vindicate an intimate connection between genuine knowledge and *liangzhi*. But in this section I will argue that, as it stands, the idea is incorrect: a person's *liangzhi* may have ethically

¹⁵In IPL 139 (QJ 56), Wang describes knowledge of conscientiousness and humaneness. He discusses compassion in IPL 135 (QJ 50-1) cf. IPL 8 (QJ 7). I follow tradition in translating 惻隱 as "compassion" but the term might be better rendered as "being pained by" or "unable to bear"; see e.g. Shun (2018, 90) for discussion.

¹⁶I have already described *liangzhi* as a "faculty", and I will continue to speak of it that way below. It is controversial whether this description, corresponding to what I call the *Faculty Model* of *liangzhi*, is correct. On an alternative interpretation, which I call the *Activity Model*, *liangzhi* is not a faculty but a set of episodes of awareness, emotions, inclinations and perhaps other mental events (Angle & Tiwald (2017, 104) describe it as a "category" of emotions). To simplify the exegesis, I will assume the Faculty Model throughout, but I simply note here that the main work of this paper can be done regardless of which model one adopts.

relevant knowledge, even if the person does not have genuine knowledge. This conclusion – and the background about *liangzhi* that I will develop in arguing for it – will set the stage for the guiding question of the next section: what more must a person have, beyond knowledge acquired by *liangzhi*, to have genuine knowledge?

The expression *liangzhi* 良知, which is made up of two characters, “good” *liang* 良 and “knowledge” *zhi* 知, occurs in *Mencius* 7A.15, where it is connected to “good ability” (*liangneng* 良能). In the *Mencius*, these two terms describe people’s putatively innate recognition that they should love their parents and respect their older brothers. Wang’s Song dynasty predecessors often discussed this passage of *Mencius*, but they tended to use the expression *liangzhi* together with *liangneng* and to cite the passage in general discussions of humans’ putatively innate capacity for virtuous action.

Wang split *liangzhi* from *liangneng* and developed a new multifaceted understanding of *liangzhi*, taken on its own. A central aspect of this new understanding was the idea that a person’s *liangzhi* knows that particular mental events are right or wrong, if they are. For instance Wang writes:

- [T1] 意則有是有非，能知得意之是與非者，則謂之良知。
Some inclinations are then right, and others wrong; what is able to know that inclinations are right and wrong is called *liangzhi*. (QJ 6.242; Ching (1972, 114))¹⁷

In [T1], Wang says that *liangzhi* is able to know that inclinations are right and wrong. In other passages, however, he says not only that *liangzhi can* know that mental episodes are right or wrong, but that it does:

- [T2] 爾那一點良知，是爾自家的準則。爾意念著處，他是便知是，非便知非
Your *liangzhi* is your own standard. Insofar as your motivating concerns (*yinian*) are attached, it knows that they are right if they are right, and that they are wrong if they are wrong. (IPL 206, QJ 105)

The expression I have translated as “motivating concerns” (*yinian* 意念) is a compound of the word I translate as “inclination” *yi* together with one which we have not yet encountered, but which I would render “concern” *nian*.¹⁸ For the rest of the paper

¹⁷The words I have translated as “right” and “wrong” (*shi* 是 *fei* 非) can also mean “correct” and “incorrect”. Since Wang clearly thinks of this correctness/incorrectness as ethical or moral correctness/incorrectness, I have opted for “right” and “wrong” here and throughout the paper.

¹⁸This term is often rendered simply “thoughts”. But this translation does not capture the fact that *nian* have more of an affective (and even an action-directed) component than other mental events; for instance, they are more affect-laden than *si* (思, which I translate “thoughts”) which are more often dispassionate. A quotation from Liu Zongzhou (劉宗周, 1578–1645) illustrates this contrast; he writes (criticizing Wang)

I will make the simplifying assumption that Wang held that motivating concerns just are inclinations (and vice versa): the reader should treat these terms as interchangeable. Nothing of substance will turn on this assumption, but it will make the exposition more compact.

[T2] says that *liangzhi* knows that a motivating concern is right or wrong whenever the motivating concern “is attached”. In general, in Wang’s idiom “being attached” would have had a negative connotation, but in this passage, Wang cannot mean that every motivating concern which is attached is thereby wrong or incorrect, since he explicitly says that they can be right or correct. It is natural instead to take Wang’s discussion of motivating concerns’ being “attached” simply to mean something like their being “aroused”. His point is that whenever one has a motivating concern, *liangzhi* knows that it is right or wrong, if it is. Thus, Wang holds that *liangzhi* is not just able to know that motivating concerns are right or wrong, but that it invariably knows that they are right or wrong if they are.

In this passage, Wang speaks of *liangzhi* as itself acquiring or possessing knowledge almost independently of the person to whom it belongs – and will see him describe *liangzhi* in a similar way again below. I will later return to the question of how we might understand this striking aspect of Wang’s way of speaking about *liangzhi*.

In addition to knowing that inclinations are right and wrong, Wang also holds that *liangzhi* knows that inclinations are good (*shan* 善) and bad (*e* 惡), if they are.¹⁹

[T3] 凡意念之發，吾心之良知無有不自知者。其善歟，惟吾心之良知自知之；其不善歟，亦惟吾心之良知自知之。

Whenever a motivating concern arises, your mind’s *liangzhi* automatically knows it. [If it is good] your mind’s *liangzhi* automatically knows that it is good; [if it is bad], your mind’s *liangzhi* also automatically knows that it is

that “A thought which is set in motion by desire is a concern. Thus concerns must be eradicated although thoughts (思) cannot be.” (思而動于欲為念。故念當除而思不可除, in Wu (2007, 遺編 v. 30, 陽明心錄 3); see Chan (1983, 142).) The word *nian* (which will be used as a noun in all the passages I discuss below) should not be understood as a concern in the sense described in the English “to be concerned about” (as in “I am concerned about you”); it fits better with concerns described by “to be concerned with” (“He is primarily concerned with his own reputation”) “to be concerned that” (“I’m concerned that they aren’t here yet”), but fits best with “to be concerned to” (“I’m concerned to get there on time”; “my concern is to ensure all of you get there safely”). Just as I noted earlier that Wang often thinks of *yi* as short-lived episodes (whereas inclinations are typically longer-lasting), so too the reader should bear in mind that Wang often thinks of *nian* as short-lived episodes (whereas concerns are often longer lasting).

¹⁹There is a clear distinction between the qualities expressed by the terms I translate “right” / “wrong” (是非) and those I translate as “good” / “bad” (善惡), but the exact character of this distinction will not be important for my purposes. In particular, I will not take a stand on the relationship of this distinction to the one between right/wrong and good/bad as understood by moral philosophers working in English today.

bad.(QJ 26.1070, Chan (1963, 278), Ivanhoe (2009, 170))²⁰

In many further passages, Wang says that a person's *liangzhi* acquires relevant knowledge, no matter how morally corrupt the person has become.²¹ These passages provide yet more support for the claim that a person's *liangzhi* always knows the ethical qualities of their mental events. Wang does not hold that only the *liangzhi* of ideally virtuous people acquires this knowledge. He says quite clearly that no matter one's state of virtue or vice, *liangzhi* always knows.²²

Wang holds that if a person has a good, right, bad or wrong inclination, their *liangzhi* knows that it is good, right, bad or wrong. Does he also endorse the converse, that if a person's *liangzhi* knows that an inclination is good, right, bad or wrong, then they have the relevant inclination and it is in fact good, right, bad or wrong? The textual evidence here is less direct, but it suggests that he does. First, Wang doesn't explicitly say that if *liangzhi* knows that an inclination is good, right, bad or wrong, it is, but this goes without saying: just as if one knows that grass is green, it must be green, so if one *zhis* that an inclination is good, it must be good. Second, Wang holds that a person's *liangzhi* knows that their inclinations are good, right, bad or wrong on the basis of a kind of bodily experience of the inclination.²³ Since a person cannot have this sort of direct experience of *other people's* inclinations, it is natural to think that he is committed

²⁰See also the second and third sentences of the "Four Sentence Teaching" (四句教) (IPL 315 QJ 133-4), as well as: IPL 162, QJ 76; IPL 259, QJ 118; IPL 318, QJ 135-6; QJ 8.307. Wang makes similar points in his pre-*liangzhi* period, in the 1515 "Preface to the Old Version of the Great Learning" QJ 7.271. Chen Jiuchuan (陳九川) also makes related points in IPL 201 QJ 102.

²¹In "The Preface to the Old Version of the Great Learning" Wang says that the original substance of the mind never fails to know (未嘗不知也 QJ 7.271). (In 1515 when he wrote this work, he did not yet speak of *liangzhi* in the way he later would.) The point is also made in QJ 5.193 (Ching (1972, 49)), and emphatically in IPL 152 (QJ 69), IPL 207 (QJ 105), and QJ 27.1112-1113 (Ching (1972, 121)). He makes related remarks in: IPL 151 QJ 69; IPL 169 QJ 81-2; IPL 289 QJ 126; IPL 320 QJ 136-7; QJ 7.298.

IPL 290 (QJ 126) is at first sight in tension with this evidence. There, Wang says that *liangzhi* can know the qualities of these states, not that it does. Moreover, he says that an immediate consequence of *liangzhi's* awareness of bad emotions is that they will disappear. This passage seems incompatible with the remarks just cited, since if (as Wang says in IPL 207 (QJ 105)) a thief's *liangzhi* knows that they ought not to be a thief, and if this knowledge comes with elimination of the bad emotions, no one would be a thief. A conservative way of reconciling the passages is to see IPL 290 as describing genuine knowledge which, as I will describe below, is naturally understood to arise only if bad mental events are eliminated.

²²In the main text I've only considered knowledge about inclinations and motivating concerns, but Wang makes related points about *liangzhi's* knowledge of thoughts (*si* 思) in IPL 169 (QJ 81-2), and emotions or feelings (*qing* 情) in IPL 290 (QJ 126).

²³"If you do not rely on your own *liangzhi* to genuinely and practically learn by bodily experience, it is like using a scale without markings to weigh what is light and heavy, or using a mirror that hasn't been opened to reflect what is beautiful and ugly..." (若不就自己良知上真切體認，如以無星之稱而權輕重，未開之鏡而照妍媸 IPL 146 QJ 66). In what follows he connects the use of *liangzhi* in this way directly to a quotation from Cheng Yi (程頤, 1033-1107) describing how one is able to distinguish right from wrong (能辨是非).

to the claim that a person's *liangzhi* can acquire knowledge in the relevant way only of their own inclinations.²⁴

Thus, Wang holds that a person has a good, right, bad or wrong inclination if and only if *liangzhi* knows that the inclination is good, right, bad or wrong. For our purposes only *liangzhi*'s knowledge that inclinations are good will be important. In particular the key takeaway from the foregoing discussion will be:

Liangzhi Knows Good If a person has a good inclination, their *liangzhi* knows that it is good. If a person's *liangzhi* knows that an inclination is good, then the person has that inclination and it is good.

This principle will be key to my argument below that not all knowledge which *liangzhi* produces is genuine knowledge. It will also be an important component of my introspective model of genuine knowledge.

In [T3], Wang describes *liangzhi* as acquiring knowledge “automatically” (自). Other passages suggest that he holds that *liangzhi* also acquires knowledge effortlessly.²⁵ Both the idea that people have a faculty which automatically and effortlessly acquires knowledge that certain mental events are good, right, bad or wrong, and the idea that this faculty acquires such knowledge independently of the people to whom it belongs, can seem odd and unfamiliar on first encounter. But reflecting on features of the conscience as it is commonly understood helps to make them less so. A person's conscience is often credited with automatic and effortless reactions which can be a source of ethically relevant knowledge. For instance, someone who has a spontaneous urge to harm someone else, may simultaneously experience revulsion at what they feel an urge to do. In the right circumstances, this revulsion is naturally described as an exercise of the person's conscience; the person might say “although I wanted to do it, my conscience told me not to” or, if the urge was something they initially felt a need to act on (perhaps

²⁴Elsewhere, Wang says that *liangzhi* is primarily focused on features of oneself. He writes that it “has nothing to do with others” (是皆無所與於他人者也 QJ 26.1070, Chan (1963, 278)) and even allows himself to appropriate a quotation from his usual opponent Zhu Xi (Johnston & Ping (2012, 155), Zhu (1983, 7)), describing *liangzhi* as “what others do not know, but I know in private” (人雖不知，而已所獨知 IPL 318 QJ 135). Since others cannot acquire this knowledge about me at all, in particular they cannot acquire it by using their *liangzhi*. And by parallel reasoning it must be that I cannot have this knowledge of *other's* mental lives using *my liangzhi*.

²⁵In the original passage in which the *Mencius* introduces *liangzhi* (7A.15), the text says that *liangzhi* “does not await reflection before it knows” (不待慮而知) and that *liangneng* “does not await learning before it is able to” (不待學而能). In two passages, Wang affirms that these features belong to *liangzhi* (QJ 26.1070; Chan (1963, 278) and QJ 8.311), and, in both of them, goes on to describe *liangzhi* as “the mind which judges right and wrong” or “the mind which approves and disapproves” (是非之心) (see text following ?? for discussion). While the *Mencius* seems to have used the “does not await...” descriptions to indicate that the relevant knowledge and ability were innate, Wang seems to think of them as related to effortlessness, and ties this effortlessness directly to the capacity for judging right and wrong.

stemming from righteous anger) they might say “I almost did it, but my conscience told me it was wrong”. These deliverances of the conscience are automatic, effortless, uncalled-for reactions, and the testimony of the conscience (its “telling”) is a source of knowledge. Moreover, in these examples, the conscience is described as an independent, almost alien force within the person’s mind, in parallel to the way in which Wang can describe *liangzhi* almost anthropomorphically as acquiring and possessing knowledge of its own.²⁶

This discussion of *liangzhi*’s knowledge – and in particular the principle *Liangzhi Knows Good* – finally puts us in a position to give the argument I promised at the start of this section, against the claim that all ethically relevant knowledge acquired by *liangzhi* is genuine knowledge. As we have seen, Wang holds that it is not just the *liangzhi* of virtuous people which knows; a person’s *liangzhi* knows that inclinations are good, right, bad or wrong, even when the person is generally morally corrupt, indeed even when they are in the midst of performing vicious bodily actions. But Wang also makes clear in several places that a key component of the doctrine of the unity of knowledge and action is that anyone who is performing vicious bodily actions does not genuinely know filiality, or any other virtue for that matter (*IPL* 5, *QJ* 4; cf. *IPL* 8, *QJ* 7). Wang is therefore committed to the view that when a person has a good inclination while performing a bad action, their *liangzhi* knows that the inclination is good, but the person does not have genuine knowledge. So, he is committed to the view that not all ethically relevant knowledge acquired by *liangzhi* is genuine knowledge.

While not all relevant knowledge acquired by *liangzhi* is genuine knowledge, I will propose later that in fact all genuine knowledge is knowledge acquired by *liangzhi*. In the next section, I will examine a passage in which Wang distinguishes between better and worse forms of knowledge that an inclination is good. I will ultimately suggest that the better form of knowledge that Wang identifies here is genuine knowledge.

²⁶There are also important differences between *liangzhi* and the conscience; I am *not* claiming that they are the same thing. For instance, Wang holds that *liangzhi* is responsible for seeing and hearing (*IPL* 168 *QJ* 80 (seeing and hearing are all “functions” 用 of *liangzhi*), cf. *QJ* 6.235, [Ching \(1972, 110\)](#)) and in some moods even suggests that it is responsible for all knowledge (*QJ* 6.243, [Ching \(1972, 115\)](#), though the relevant passage is not translated in her selection; *QJ* 20.871, [Ivanhoe \(2009, 182\)](#)); clearly the conscience does not have such broad powers. Still, some of the core phenomenology which Wang aimed to capture with his theory of *liangzhi* is clearly similar to what is often associated with the conscience. I return to this point briefly below in connection to [\[T5\]](#).

For some early discussion in English of the relationship between the conscience and *liangzhi*, see, e.g. [Graham \(1958, xx\)](#), [Chang \(1955\)](#), [Mou \(1973, 104 n. 3\)](#); cf. [Tang \(1973\)](#) who often uses “conscientious consciousness” to translate *liangzhi*, and [Cheng \(1974\)](#). For some more recent discussion, see [Bol \(2008, 169\)](#), [Kern \(2010, 219-220\)](#), [Huang \(2016b, 60-1\)](#) (in English as: [Huang \(2017, 81-2\)](#)) and [Yu \(2016, 19-21\)](#).

4 The obscuration argument

The passage in which Wang distinguishes these different forms of knowledge comes from *Questions on the Great Learning*, a mature statement of Wang's views on the interpretation of the *Great Learning* (大學), which in turn was one of the most important canonical texts in Wang's tradition. Scholars working in this tradition held that a central part of the *Great Learning* provides an enumeration of four tasks or stages in an individual's personal ethical development. In the passage I will examine here, Wang discusses the relationship between two of these tasks or stages: "extending knowledge" (*zhi zhi* 致知 – an idea which we encountered earlier in the text surrounding n. 11); and "making inclinations wholehearted" (*cheng yi* 誠意).²⁷ In this section and the next I will not say much about what Wang might mean by "wholehearted inclinations"; I'll simply treat this term as a placeholder. I'll return to the notion in section 6.

The passage I will examine in this section in fact contains two ideas that will be important in what follows. First, Wang here distinguishes the knowledge a person's *liangzhi* invariably acquires from an elevated form of knowledge ("extended knowledge"), by describing how the elevated form of knowledge requires freedom from a certain form of doxastic conflict. This idea – that a person has an elevated form of knowledge only if they do not suffer from a form of doxastic conflict – will be key to my account of the sense in which genuine knowledge is an elevated form of knowledge (section 5). Second, Wang here provides a detailed analysis of the relationship between extended knowledge and wholehearted inclinations. Later, in section 6, I will argue that if an account of genuine knowledge can vindicate a close connection between genuine knowledge and wholehearted inclinations, it will help to make progress toward an interpretation of the doctrine of the unity of knowledge and action as a whole. Ideas that Wang develops in the present passage – and which I will spend some time interpreting in this section – will be key to my later account, of how the introspective model itself makes sense of a tight connection between genuine knowledge and wholehearted inclinations (section 6).

- [T4] 故欲正其心者，必就其意念之所發而正之，凡其發一念而善也，好之真如好好色，發一念而惡也，惡之真如惡惡臭，則意無不誠，而心可正矣。然意之所發，有善有惡，不有以明其善惡之分，亦將真妄錯雜，雖欲誠之，不可得而誠矣。故欲誠其意者，必在於致知焉。〇〇
- 凡意念之發，吾心之良知無有不自知者。其善歟，惟吾心之良知自知之；其不善歟，亦惟吾心之良知自知之。是皆無所與於他人者也。故雖小人為不

²⁷The traditional translation of *cheng* 誠 is "sincere". *Cheng yi* is variously rendered as "making the will sincere", "making thoughts sincere" or "making intentions sincere".

善，既已無所不至，然其見君子，則必厭然掩其不善，而著其善者，是亦可以見其良知之有不容於自昧者也。

今欲別善惡以誠其意，惟在致其良知之所知焉爾。何則？意念之發，吾心之良知既知其為善矣，使其不能誠有以好之，而復背而去之，則是以善為惡，而自昧其知善之良知矣。意念之所發，吾之良知既知其為不善矣，使其不能誠有以惡之，而覆蹈而為之，則是以惡為善，而自昧其知惡之良知矣。若是，則雖曰知之，猶不知也，意其可得而誠乎！今於良知之善惡者，無不誠好而誠惡之，則不自欺其良知而意可誠也已。

Therefore if you want to rectify your mind, you must rectify it in regard to the arousal of your motivating concerns. If, whenever a concern arises and it is good, you genuinely love it as you love lovely sights, and whenever a concern arises and it is hateful [bad], you genuinely hate it as you hate hateful [bad] odors, then all of your inclinations will be wholehearted and your mind can be rectified.²⁸ However, some of the inclinations which arise are good and some are bad. If one did not have a means to understand the distinction between good and bad, and wrongly mixed up true (真) and misguided, then although one wanted to make them [viz. one's inclinations] wholehearted, they cannot successfully become wholehearted. Thus making one's inclinations wholehearted must depend on extending one's knowledge of them...

Whenever a motivating concern arises, your mind's *liangzhi* automatically knows it. [If it is good] your mind's *liangzhi* automatically knows that it is good; [if it is bad], your mind's *liangzhi* also automatically knows that it is bad.²⁹ It has nothing to do with other people. Thus, although a petty person has become not good, and there is nothing they will stop at, nevertheless when they meet a virtuous person, they will ashamedly hide the fact that they are not good, and outwardly project that they are good. From this one can see that there are some respects in which their *liangzhi* has not allowed itself to be obscured.

Now, if you want to discriminate good and evil in order to make your inclinations wholehearted, this just depends on extending what your *liangzhi* knows about them and nothing more. Why is this? When a [good] motivating concern arises, the *liangzhi* of your mind already knows that it is good. Suppose you do not wholeheartedly love it but instead turn away from it and diminish it. You would then be taking what is good to be bad and obscuring your *liangzhi* which knows that it is good. When a [bad] motivating concern arises, the *liangzhi* of your mind already knows that it is bad. Suppose you do not wholeheartedly hate it but instead backslide and promote it. You would then be taking what is bad to be good and obscuring your *liangzhi* which knows

²⁸“Love lovely sights” and “hate hateful odors” are quotations from the *Great Learning*; see [T6] below. My translation attempts to simulate the fact that “love” is written with the same character (好) as the adjective “lovely” (although they are pronounced differently) and the verb “hate” is written with the same character (惡) as the adjective “hateful” (although they too are pronounced differently). (Above, I have used “bad” for the character I here translate as “hateful”.)

²⁹This text was also printed above as [T3].

that it is bad. In such cases one says that you know it, but in fact you do not know – how could your inclinations have become wholehearted! [But] now if what *liangzhi* [recognizes as] good or bad is wholeheartedly loved or hated, one's *liangzhi* is not deceived and one's inclinations can be wholehearted. (QJ 26.1070-1, cf. Chan (1963, 277-9))

I propose that Wang aims to argue in this passage that a person has extended knowledge if and only if they have wholehearted inclinations. At the start of the passage Wang says that making inclinations wholehearted “depends on” (在於) extending knowledge, and also that discriminating good and evil in order to make inclinations wholehearted “depends on” (在) extending knowledge. At the close of the passage he says that if *liangzhi* is not deceived (that is, presumably, if the person's knowledge is extended), one's inclinations “can be” (可) wholehearted. These remarks most naturally suggest that he intends to argue only that having extended knowledge is a necessary condition for having wholehearted inclinations. But Wang often repeats (including at the end of this work) his distinctive view that the four aspects of personal ethical training in the *Great Learning* are really all just different ways of looking at the same task: if one of them is brought to completion, they all must be.³⁰ He therefore also holds that having extended knowledge is sufficient for having wholehearted inclinations. And in fact this claim seems to be in his sights in our passage itself. The phrase “can be wholehearted” in the last sentence of the excerpt is most naturally read as “the only obstacles to their being wholehearted will be removed” suggesting that Wang means to claim not only that extended knowledge is required for wholehearted inclinations, but that extended knowledge is sufficient for wholehearted inclinations as well. Even more importantly, as we will see below, the most explicit part of Wang's argument (in the third paragraph of the excerpt) clearly targets the claim that extended knowledge suffices for wholehearted inclinations. Together these facts suggest that Wang's “depends on” expresses a tighter connection than “requires” – perhaps something more like “is constituted by” – and that in the passage as a whole Wang sets out to defend the claim that a person has wholehearted inclinations if and only if their knowledge is extended.³¹

The argument stated explicitly in the third paragraph of the passage, which I will call the *obscuration argument*, is an argument for an intermediate conclusion on the way

³⁰ IPL 137 QJ 54; QJ 26.1069-70, Chan (1963, 277). See Shun (2011, Section IV) and Ching (1976, 82-4) for discussion.

³¹ Since Wang does not discuss the unity of knowledge and action explicitly in this passage, it does not follow from my principle for demarcating cases of “extended knowledge” which describe genuine knowledge that his uses of “extended knowledge” in this passage describe genuine knowledge, and accordingly in what follows I will not assume that they do. See below n. 47 for more discussion.

to this main claim. In this sub-argument Wang seeks to show that if a person does not wholeheartedly love (好) a good (善) motivating concern, then they do not know that it is good, and similarly that if a person does not wholeheartedly hate (惡) a bad (惡) motivating concern, they do not know that it is bad. The argument runs as follows (in the case of a good motivating concern, which I will focus on throughout):

1. If a person does not wholeheartedly love a motivating concern, they turn away from it and diminish it.
2. If a person turns away from and diminishes a motivating concern, they take it to be bad.
3. If a person takes something to be bad, they do not know that it is good.
4. So, if a person does not wholeheartedly love a motivating concern, they do not know that it is good.³²

As stated explicitly above, Wang's argument concerns a person's knowledge, and not whether their knowledge is extended. But in context, it is clear that Wang does not mean to deny that the person knows the relevant claim; he wants only to deny that they have extended knowledge. In the second paragraph of the passage, Wang says that *liangzhi* always knows, and in the passage as a whole, Wang aims to elucidate a contrast between the knowledge that *liangzhi* always has, and extended knowledge. In the second paragraph, Wang says that petty or vicious people recognize their faults and hide them in front of virtuous people. He takes this to show that even vicious people know that they are not good by means of their *liangzhi*. But their knowledge cannot be *extended* knowledge; extended knowledge is something only virtuous people have. Moreover, at the end of the excerpt, Wang admits that in a case where a person fails to wholeheartedly love their good inclinations, we do *say* that the person knows. He here seems to recognize a divergence between his own way of speaking of (extended) knowledge, and the form of knowledge described by ordinary uses of the word "know".³³

³²It may be helpful to think of the second and third premises of Wang's argument by comparing them to the premises of the following argument that cognitivists about desire (that is, those who hold that if one desires that p one believes that it's good that p) cannot allow that people want what they know to be not good:

2. If a person wants that p , they believe that it is good that p .
3. If a person believes that p , they do not know that $\neg p$.
4. If a person wants that p , they do not know that it is not good that p .

³³Elsewhere, too, Wang implicitly concedes that someone may be said to know even though they do not have genuine or extended knowledge (IPL 5, QJ 4; IPL 138, QJ 55).

In light of these observations, the obscuration argument should be understood as focused on whether the person in question has extended knowledge, not whether they have any form of knowledge at all. The third premise of the argument should thus be rephrased as “If a person takes something to be bad, they do not have extended knowledge that it is good”; and the conclusion as “If a person does not wholeheartedly love a good motivating concern, they do not have extended knowledge that it is good.”

The third premise of the argument, therefore, articulates a necessary condition for knowledge that an inclination is good to be extended, namely, that the person not take the inclination to be bad. In what sense might “taking what is good to be bad” (以善為惡) prevent a person from having this elevated form of knowledge that an inclination is good?

Consider a native Chinese speaker who pronounces the Pinyin sound “zh” in the standard way, by putting their tongue against the back of their palette, but who, when they reflect on their practice in pronouncing this sound, comes to believe that the sound is produced by putting their tongue at the front of their mouth against their lower teeth. If there can be any people who both know a claim, and believe the negation of that claim, this person is a good candidate: they know – as evinced by their ability to reliably, intentionally produce the sound correctly – that the sound is produced by putting one’s tongue at the back of one’s palette. But they also believe – as a result of reflecting on their practice – that the sound is not produced in this way. Now compare this person (“the conflicted person”) to someone (“the unconflicted person”) who can produce the sound but who also, on reflection, comes to believe truly that the sound is produced by putting their tongue at the back of their palette. The unconflicted person’s state of mind can be said to be better than the conflicted person’s state of mind in at least one distinctively epistemic or doxastic respect, since the conflicted person suffers from a form of doxastic conflict which the unconflicted person does not suffer from. In a grandiose mood, one might say that, in virtue of this fact, the unconflicted person has an elevated form of knowledge by comparison with the conflicted person.

Wang’s idea in our passage seems to be related. He considers a structurally parallel contrast between people in two different conditions. In one condition, a person knows that an inclination is good and wholeheartedly loves it; in another, a person knows that an inclination is good, but does not wholeheartedly love it – and, as a result, takes it to be bad. A person who knows that an inclination is good, but also “takes what is good to be bad” suffers from something like the doxastic conflict described above, while the person who wholeheartedly loves their good inclination presumably does

not.³⁴ Wang articulates the contrast between these two conditions by distinguishing two grades of knowledge; his remarks suggest that a person who suffers from this conflict have a worse form of knowledge, by contrast to the better, extended knowledge, the wholehearted person has. But we could paraphrase Wang's idea without speaking of grades of knowledge in this way: the idea is that the state of mind of a person who wholeheartedly loves a good motivating concern is better in a distinctively epistemic or doxastic respect than a person who does not, because the person whose love is not wholehearted suffers from a form of doxastic conflict which the person whose love is wholehearted does not suffer from.

This completes my discussion of the first key idea contained in this passage. This idea – that extended knowledge requires freedom from something like doxastic conflict – will be at the heart of the account of genuine knowledge I present in the next section. There, I will propose that genuine knowledge is an elevated form of knowledge precisely because a person who has genuine knowledge must be free from the kind of doxastic conflict described here.

But before I turn to that account of genuine knowledge, I will first present a second key idea from our passage: how Wang connects extended knowledge and wholehearted inclinations. To better understand this connection, let us begin by reconsidering the first and second premises of the obscuration argument. The first premise states that if one does not wholeheartedly love a motivating concern, one turns away from it and diminishes it.³⁵ Wang's point seems to be that, since the response good inclinations

³⁴The fact that Wang connects "taking what is good to be bad" and "taking what is bad to be good" to the "self-deception" described by the *Great Learning* (see below [T6]) may provide further support for the idea that Wang means to emphasize something like doxastic conflict in his discussion (cf. also IPL 138, QJ 55; IPL 171, QJ 84). Just as a person who suffers from self-deception about the extent of their own accomplishments may be said to know, deep down, that they have not achieved much, in spite of the fact that they outwardly seem to believe they have accomplished a great deal, so too the person described by Wang's gloss on "self-deception" in the *Great Learning*, knows the goodness of a given inclination (via *liangzhi*) but also takes it to be bad.

³⁵In the passage Wang does not make this claim directly; he says "if one does not wholeheartedly love a motivating concern, but instead turns away from it and diminishes it...". On its own this phrase could be read as saying not that wholeheartedly loving an inclination on the one hand and turning away and diminishing it on the other are exhaustive alternatives but only that turning away and diminishing an inclination is a special, extreme form of failing to wholeheartedly love it. This interpretation would seem on an even stronger footing if one reads (as is linguistically possible) the words I have translated as "diminish" (去) and "promote" (為) as "eliminate" and "enact" respectively. But in the context of Wang's argument here, this reading and the associated translation are disfavored. Wang clearly takes his argument to show that *liangzhi's* powers are not fully exercised if one does not wholeheartedly love a good motivating concern. If he thought that turning away and diminishing a good motivating concern was just a special case of failing to wholeheartedly love it, there would be an obvious gap in his argument: he would have failed to show that *liangzhi's* knowledge would also be degraded if the person exhibited different, less extreme ways of failing to wholeheartedly love the relevant inclination.

warrant is wholehearted love, anything short of that response amounts to diminishing their status, and turning away from them.³⁶

The second premise of the obscuration argument connects an affective feature of a person's response to an inclination ("turning away") with the more doxastic "taking". Wang says that if one turns away from and diminishes an inclination, one takes it to be bad. He says nothing in this passage to defend this claim. He may have seen it as simply obvious that there is a doxastic component to the affective reactions he considers: that loving something requires taking it to be good; and that turning away from and diminishing something requires taking it to be bad. Alternatively, Wang's views on this connection may have stemmed from a deeper commitment about the nature of *liangzhi's* response to inclinations. The following passage is suggestive:

[T5] 良知只是個是非之心，是非只是個好惡，只好惡就盡了是非。。。。。
Liangzhi is just the mind which [judges] right and wrong (*shi fei*). [Judging] right and wrong (*shi fei*) is just loving and hating. If you have just loved and hated, then you have exhausted [judging] right and wrong (*shi fei*). (IPL 288, QJ 126)

Here the characters I have elsewhere translated as "right" and "wrong" (*shi fei* 是非) (which are the same as the nouns "rightness" and "wrongness") are used as transitive verbs. I have rendered them "judge right" and "judge wrong" to mark the fact that the same characters are used here, but the expressions might also be translated as "approve" and "disapprove". The passage clearly says that *liangzhi's* approving an inclination or judging it to be right *just is* loving that inclination, but it leaves open a variety of positions about how this "approving" or "judging to be right" is related to what Wang elsewhere describes as *liangzhi's knowing*. An intriguing hypothesis is that Wang takes *liangzhi's* knowledge of the rightness of an inclination to be identical to *liangzhi's* approving the inclination. This idea can seem especially natural when one reflects on phenomenology associated with the conscience: although we can speak of the conscience "telling" us that something is wrong or right, and so conferring knowledge about it, such a "telling" may feel internally more like an affective reaction. If Wang did take the knowing and apt loving of *liangzhi* to be identical, it would be natu-

³⁶In section 6, I'll suggest that a person's inclination is wholehearted if and only if they have no inclination which conflicts with it. In a moment I'll also argue that Wang holds that good inclinations are wholehearted if and only if they are wholeheartedly loved. Given these two ideas, if a person's good inclination is not wholeheartedly loved, it is not wholehearted, and hence the person has an inclination which conflicts with it. On this picture, Wang's point in this first premise might be, in particular, that the tug in a different direction from a conflicting inclination amounts to "turning away from and diminishing" the good inclination.

ral for him to take the mistaken “taking” and inapt “turning away” he describes in [T4] also to be identical. And if the taking and the turning away are the very same thing, then of course the presence of the one requires the presence of the other.

Whatever Wang’s reasons – whether he took it to be simply obvious, or had this more involved justification in mind – it is clear that in this second premise of the argument, Wang connects an affective reaction (“turning away”), with a more doxastic one (“taking”).

Together, these premises bring us to the conclusion of the obscurator argument: that if a person does not wholeheartedly love an inclination, then they do not have extended knowledge that it is good. By contraposition, this conclusion is equivalent to the claim that if a person has extended knowledge that an inclination is good, then they wholeheartedly love it. This claim is very close to establishing a connection between extended knowledge and wholehearted inclinations. To make the final step, we need a final idea. In [T4], Wang repeatedly alludes to one of his favorite quotations from the *Great Learning*:

- [T6] 所謂誠其意者，毋自欺也，如惡惡臭，如好好色
What is called making the inclinations wholehearted is not deceiving oneself.
It is like hating a hateful [bad] odor, like loving a lovely sight.

In [T4], Wang transposes the *Great Learning*’s talk of hating hateful [bad] odors, and loving lovely sights to a person’s metacognitive reactions to their inclinations; he speaks of hating hateful [bad] inclinations and loving good inclinations. He implicitly relies on the claim – which he seems to take to be supported by [T6] – that all of a person’s inclinations are wholehearted if and only if they wholeheartedly love their good inclinations and wholeheartedly hate their bad ones. Given Wang’s commitment to this claim about all inclinations, it is plausible that he also endorses a principle I will call “Wholehearted Love”: if a person has a good inclination and wholeheartedly loves it, then the inclination is wholehearted; and, if a person has a wholehearted good inclination, then they wholeheartedly love it. This principle, together with Liangzhi Knows Good, and the conclusion of the obscurator argument, allows us to derive the desired connection between extended knowledge and wholehearted inclinations, in particular, the claim that: if a person has extended knowledge that an inclination is good, then they have that inclination, and it is wholehearted.³⁷

³⁷Wholehearted Love and the conclusion of the obscurator argument entail that: if a person has a good inclination, and has extended knowledge that it is good, then it is an inclination they have, and it is wholehearted. We can show that the first conjunct of the antecedent of this claim is redundant, and thus derive the claim in the main text, as follows. First, Wang clearly holds that if a person has extended

This conclusion is a natural way of making precise the claim that if a person has extended knowledge, their inclinations are wholehearted. But this is just half of what I earlier described as Wang’s target here. As I mentioned there, Wang’s remarks about his goal in our passage are most readily understood as focused on the converse of our present conclusion, that is, on the idea that if a person’s inclinations are wholehearted, their knowledge is extended. It is admittedly a little surprising that Wang does not give a more explicit argument for this converse, but it is in any case easy to see how the argument Wang does give can be transformed into an argument for it, based on claims he would have accepted. First, it is natural to think that Wang would have accepted something very close to converses of each of the three premises of the obscuration argument, and thus, that he would have accepted something very close to the converse of its conclusion as well. For the first premise: Wang says explicitly that if a person does not wholeheartedly love an inclination, they turn away from it and diminish it. But he clearly also thinks that if a person turns away from and diminishes an inclination, they do not wholeheartedly love it. Similarly, for the second premise: Wang says that if a person turns away from and diminishes an inclination, they take it to be bad. But it is natural to think that he would also hold that taking a good inclination to be bad is a way of turning away from it and diminishing it, and thus that if a person takes a good inclination to be bad, they turn away from it and diminish it. Finally, for the third premise: Wang says that if a person takes what is good to be bad, then they do not have extended knowledge that the inclination is good. But it is natural to think that he would hold that having a mental state which yields this kind of conflicting taking is the only way that a person’s *liangzhi*’s knowledge that an inclination is good could fail to yield extended knowledge, and so, if a person’s *liangzhi* knows that an inclination is good but they do not have extended knowledge that it is good, then they take what is good to be bad.³⁸ If this is right and Wang was committed to these further ideas, then

knowledge that an inclination is good, their *liangzhi* knows that it is good. Second, *Liangzhi Knows Good* states in part that if a person’s *liangzhi* knows that an inclination is good, then it must be an inclination they have. Thus, given our assumptions: if a person has extended knowledge that an inclination is good, it is an inclination they have. This fact together with the claim in the first sentence of this note, gives us the desired result that if a person has extended knowledge that an inclination is good, then it is an inclination they have, and it is wholehearted.

³⁸ Later, it will be useful to have isolated the following claim, which I will (slightly inexactly) refer to as “the third premise of the obscuration argument and its converse”:

A person has extended knowledge that an inclination is good if and only if their *liangzhi* knows that the inclination is good, and they do not take that inclination to be bad.

The right-to-left direction is the claim I have just argued for in the main text. The left-to-right direction slightly strengthens (the contraposition of) premise 3 of the obscuration argument, claiming that extended knowledge requires not only that the person not take the inclination to be bad, but also that their *liangzhi* knows that the inclination is good. But this further claim is clearly something Wang would have endorsed

he would also have been committed to the claim that if a person wholeheartedly loves a good inclination, and their *liangzhi* knows that it is good, then they have extended knowledge that it is good. Together with Wholehearted Love (from the previous paragraph) and Liangzhi Knows Good (from the previous section), this claim implies that if a person has a wholehearted good inclination, then they have extended knowledge that this inclination is good.³⁹

This completes my discussion of the second main idea I promised at the start of this section. We have now seen arguments for a precise connection between extended knowledge and wholehearted inclinations, in the form of the following principle:

Extended/Wholehearted If a person has extended knowledge that an inclination is good, then they have this inclination and it is wholehearted. If a person has a wholehearted good inclination, then they have extended knowledge that this inclination is good.

This principle codifies the idea – which I suggested was Wang’s target in our passage – that a person has extended knowledge if and only if they have wholehearted inclinations. As I mentioned earlier, the principle will play an important role in section 6, when I come to show how my account of genuine knowledge contributes to an interpretation of the unity of knowledge and action as a whole, by vindicating a close connection between genuine knowledge and wholehearted inclinations.

But first, it is time to present the account of genuine knowledge, and answer the main challenge of the paper.

5 Introspective Knowledge

To introduce the account of genuine knowledge, it is worth stepping back to consider what an adequate account of genuine knowledge must achieve. Extrapolating from the connection Wang draws between extended knowledge that an inclination is good and wholehearted inclinations to perform a good action (in Extended / Wholehearted), it is natural to suspect that Wang would endorse a parallel connection between our tar-

(see previous note).

³⁹By Wholehearted Love, we derive the claim that if a person has a wholehearted good inclination, and their *liangzhi* knows that this inclination is good, then they have extended knowledge that it is good. By Liangzhi Knows Good – part of which states that if a person has a good inclination, their *liangzhi* knows that it is good – we can eliminate the second conjunct of the antecedent of this claim, and derive the desired result that if a person has a wholehearted good inclination, then they have extended knowledge that this inclination is good.

get notion of genuine knowledge – genuine knowledge of filiality – and wholehearted inclinations to perform a filial action, that is:

Genuine/Wholehearted A person genuinely knows filiality if and only if they have a wholehearted inclination to perform a filial action.⁴⁰

In section 6 I will argue that this principle does in fact play an important role in Wang's doctrine of the unity of knowledge and action. But for now, I will use it more abstractly, to illustrate what is required for a satisfactory account of genuine knowledge.

Genuine/Wholehearted on its own is not an adequate account of genuine knowledge. The principle states conditions under which a person has genuine knowledge, but it does not say what genuine knowledge is. It thus leaves open the possibility that Wang held that the event of genuinely knowing filiality just is the event of acting filially – if he thought that a person acts filially if and only if they have a wholehearted inclination to perform a filial action. But this is precisely the kind of commitment that would make the unity of knowledge and action a triviality disguised by a misleading definition. If Wang stipulated that genuinely knowing filiality is acting filially, he would have simply stipulated that knowledge and action are unified. He would have given no explanation of why actions should count as knowledge (as that term is ordinarily understood), never mind an elevated form of knowledge. So an account of genuine knowledge must go beyond Genuine/Wholehearted. It must not just say when people have genuine knowledge, but also what genuine knowledge is.

In section 3 I considered the idea that all ethically relevant knowledge possessed by *liangzhi* is genuine knowledge. I argued that this idea is mistaken: genuine knowledge requires virtuous conduct, but even the *liangzhi* of vicious people has ethically relevant knowledge. The conclusion of that section – that not all ethically relevant knowledge of *liangzhi* is genuine knowledge – however, left open the possibility that a subclass of *liangzhi*'s knowledge could be genuine knowledge. And in the previous section, we saw that Wang himself distinguishes between two grades of knowledge that an inclination is good. He says that a person has a better form of knowledge (extended knowledge) if the person's *liangzhi* not only knows that a good inclination is good (as it does whenever a person has a good inclination), but if in addition the person does not take the relevant inclination to be bad. Wang does not discuss the unity of knowledge and action in [T4], but it is plausible that the elevated form of knowledge he describes

⁴⁰Below, when I display principles like Genuine/Wholehearted which discuss filiality explicitly, I intend also to endorse the obvious variants of them for the other virtues mentioned at the end of section 2, and when I use names for principles like "Genuine/Wholehearted" I will sometimes mean the family of these principles, not just the one which describes filiality.

there is closely related to genuine knowledge. In many places Wang says explicitly that *liangzhi* must be extended to achieve an elevated state of knowledge connected to the unity of knowledge and action, that is, genuine knowledge (see n. 11). In one passage, he connects his idea that *liangzhi* is “the mind which judges right and wrong” – a description which calls to mind *liangzhi*’s capacity to know that inclinations are good, bad, right and wrong – directly to the unity of knowledge and action (QJ 5.211 Ching (1972, 68-9)). These passages provide circumstantial evidence that genuine knowledge does not only coincide with the extended knowledge described in [T4] (so that a person experiences an episode of genuine knowledge if and only if they have extended knowledge that an inclination is good) but, moreover, that the two are identical. More exactly, and building on the characterization of extended knowledge from the previous section, the passages suggest:

Introspective Knowledge To be an episode of a person’s genuinely knowing filiality is to be an episode of their *liangzhi*’s knowing that an inclination – which is in fact an inclination to perform a filial action – is good and for the person not to take the inclination to be bad.⁴¹

This principle is the last piece of my introspective model of genuine knowledge. This model consists of four claims: Liangzhi Knows Good; the third premise of the obscuration argument and its converse (see n. 38); Extended/Wholehearted, and Introspective Knowledge. In the next section, it will become clear why the model includes all four of these claims, and not Introspective Knowledge on its own. But in the remainder of this section, I will focus on Introspective Knowledge. We can already see that this principle avoids the problem with Genuine/Wholehearted that I mentioned a moment ago: it does not merely characterize the conditions under which a person has genuine knowledge, but identifies episodes of genuine knowledge with particular mental events, episodes of *liangzhi*’s knowing that an inclination is good.

⁴¹I use the expression “to be...is to be...” to express a symmetric relation akin to identity. I assume that such identities among properties entail that the properties are necessarily coextensive. So, for instance, this principle entails that necessarily every episode of genuinely knowing filiality is an episode of their *liangzhi*’s knowing that an inclination is good.

As mentioned in n. 9, Wang sometimes uses “knowledge” not for episodes of knowledge, but for a capacity or disposition to experience such episodes in the appropriate circumstances. Introspective Knowledge is intended to be neutral on a parallel question about genuine knowledge: the principle does not say how genuine knowledge is related to episodes of genuine knowledge, so it does not take a stand on whether genuine knowledge is itself episodic or is instead a disposition to experience episodes of genuine knowledge. If forced to guess I would say that Wang had not settled on univocal senses for “genuine knowledge” and related technical terms: he could use them to describe episodes of knowing, but also to describe a disposition to experience such episodes in the appropriate circumstances.

Introspective Knowledge allows us at last to answer the main challenge of the paper. The challenge can be thought of as having two parts: first, to show that genuine knowledge of filiality is recognizably a form of knowledge – and in particular, of objectual knowledge; and, second, to elucidate the sense in which genuine knowledge is an elevated form of that knowledge, and not simply “whatever knowledge happens to be unified with action”. I will take these parts of the challenge one at a time.

According to Introspective Knowledge, episodes of genuine knowledge are clearly a form of knowledge: they are episodes of *liangzhi*'s knowing that an inclination is good. But can this knowledge reasonably be thought of as objectual knowledge of filiality?

As a prelude to my answer, a few remarks about filiality are in order. Wang fairly clearly thinks that a person's actions are filial, if they are, because the person has a “mind which is filial to their parents” (孝親之心).⁴² Since the filiality of a person's mind explains the filiality of their actions (and isn't itself explained by the fact that some further entity is filial), there is a sense in which a person's mind is the primary bearer of the quality of filiality. Wang does not describe inclinations themselves as filial, but elsewhere he identifies “the mind which...” with an inclination, and it is natural to think that he takes the mind which is filial to be a filial inclination (albeit, most plausibly, a wholehearted one).⁴³ Accordingly, I will assume that what I have been calling “inclinations to perform filial actions” to this point just are such filial inclinations.

Given Wang's commitment to the claim that the mind is, in the sense just described, the primary bearer of filiality, it is natural to think that he would hold that the most direct cognitive relation one could have to the property of filiality – which is naturally described as “knowledge of filiality” – would be a form of introspective knowledge, acquired by *liangzhi*. In particular, one might think that he would hold that knowledge of filiality just is *liangzhi*'s knowing that an inclination is filial. But Wang never says explicitly that *liangzhi* knows that inclinations are filial, or, for that matter, that it knows

⁴²See IPL 133 (QJ 48) where he also makes the same point about the relationship between actions which are conscientious toward one's ruler and “the mind which is conscientious toward the ruler” (忠君之心). Cf. IPL 3 QJ 2-3; IPL 135 QJ 50-1.

⁴³In IPL 132 QJ 7, Wang says that “the mind which desires food is an inclination” (欲食之心即是意), and that “the mind which desires to travel is an inclination” (欲行之心即是意).

Why then would he not describe them as filial? Here is one hypothesis. Wang holds that it is misleading to describe such inclinations as filial when they are not wholehearted, since that description might suggest that people with filial inclinations already count as filial, whereas only a person with wholehearted inclinations does. However, if a person does have a wholehearted inclination to perform a filial action, they count as having “a mind which is filial”, and this is a better description since it is less likely to lead to confusion.

that they are respectful, conscientious, humane or compassionate.⁴⁴ What should we make of his silence on this point? I will consider two different responses, each of which seems to me plausible, though I slightly favor the second.

The first response is to dismiss Wang's silence, and to hold that *liangzhi* does know that inclinations are filial, if they are. There is certainly no in principle reason that Wang could not have said that, in addition to knowing that inclinations are good or bad, right or wrong, *liangzhi* also knows that they are filial (or respectful, conscientious, humane or compassionate). As we have seen (n. 26), Wang accords *liangzhi* extremely broad powers. The claim that *liangzhi* also knows that inclinations are filial would be a minor and natural addition to the list of the abilities he attributes to it. So, according to this response, in addition to Liangzhi Knows Good, Wang endorses "Liangzhi Knows Filiality" (the result of replacing "good" with "filial" in Liangzhi Knows Good). Moreover, he endorses a slight variant of Introspective Knowledge, where again, "good" is replaced with "filial". And this latter principle yields a straightforward sense in which genuine knowledge is objectual knowledge of filiality: *liangzhi's* knowledge that an inclination is filial is the most direct cognitive relation a person can have to the property of filiality.

This first response seems to me reasonable, and I would consider my defense of the introspective model a success if readers are persuaded by it. But I myself am not completely convinced. I argued in section 2 that Wang centrally takes genuine knowledge to be genuine knowledge of filiality, respect, conscientiousness, humaneness and compassion. If his explanation of the sense in which genuine knowledge is objectual knowledge of filiality and these other qualities depends on the claim that *liangzhi* itself can know that inclinations are filial, respectful, conscientious, humane or compassionate, it would be odd that Wang never explicitly makes this claim.

This concern leads me to prefer a second response. Even supposing that *liangzhi* itself can only know that inclinations are good or right, and not that they are filial or respectful, *liangzhi's* knowledge that an inclination (which is in fact filial) is good would still be the most direct sensitivity a person can have to the filiality of the inclination. So, even if what *liangzhi* knows is not that the inclination is filial, but only that it is good, it would still be natural to call this direct sensitivity to filiality "knowledge of filiality". On this view, *liangzhi's* knowledge that an inclination (which is in fact filial) is good is knowledge of filiality; its knowledge that an inclination (which is in fact respectful) is

⁴⁴In IPL 8 (QJ 7), Wang does say that the mind (and in context quite clearly *liangzhi*) "knows filiality", "knows respect" and "knows compassion". But it is not clear in that passage whether he has in mind knowledge of qualities of mental events, or something more general.

good is knowledge of respect, and so on. The difference between knowing filiality and knowing respect lies not in the content of *liangzhi*'s knowledge – *liangzhi* only knows that the inclination in question is good – but rather in whether the inclination itself is in fact filial or respectful. Although *liangzhi* does not know that inclinations are filial (as opposed to respectful, or conscientious, for instance) its knowledge that a filial inclination is good is still the most direct sensitivity a person can have to the filiality of an inclination. And, given that filiality is in the sense described earlier primarily instantiated in the mind, this is the most direct sensitivity a person can have to filiality itself.

In either case – whether on the basis of the first or the second response to Wang's silence about whether *liangzhi* itself knows that inclinations are filial – the ideas behind Introspective Knowledge allow us to make sense of the idea that genuine knowledge is objectual knowledge of filiality, respect, conscientiousness, humaneness and compassion. Since I myself slightly prefer the second response, I'll focus on it in the main text from now on – although, as I have said, I am open to the view being developed instead on the basis of the first.

I now turn to the second part of our challenge: in what sense is genuine knowledge of filiality an elevated form of knowledge of filiality? In the previous section we saw that Wang draws a distinction between, on the one hand, a person whose *liangzhi* knows that an inclination is good and who takes that inclination to be bad, and, on the other, a person whose *liangzhi* knows that an inclination is good, but does not take it to be bad. I suggested that the first person suffers from a form of doxastic conflict which degrades their knowledge; it prevents them from having extended knowledge that the relevant inclination is good.

This paradigm applies to gradations in a person's propositional knowledge that an inclination is good. But Introspective Knowledge allows us to extend the paradigm to objectual knowledge of filiality. I suggested a moment ago that knowledge of filiality in general (whether or not it is genuine) should be identified with *liangzhi*'s propositional knowledge that an inclination (which is in fact filial) is good: this knowledge is the most direct sensitivity a person can have to the property of filiality as it is instantiated in the mind. But then, given that episodes of knowing filiality are identical to episodes of propositional knowledge, if the propositional knowledge is degraded by a relevant form of doxastic conflict, it is natural to think that the knowledge of filiality will be degraded as well. If a person knows filiality by knowing that an inclination (which is in fact filial) is good, but their knowledge that the inclination is good is not of the highest form, then their knowledge of filiality is also not of the highest form. So, gen-

uine knowledge of filiality, which requires freedom from this form of doxastic conflict, is elevated above the more ordinary form of knowledge of filiality, in a distinctively doxastic or epistemic respect.⁴⁵

In closing this section, I want to bring out what is at stake in Introspective Knowledge by developing and responding to an important objection to it. In perhaps the most famous discussion of the unity of knowledge and action (*IPL* 5, *QJ* 4), a student asks Wang whether the unity of knowledge and action is threatened by the fact that many people know that they ought to be filial but fail to act filially. In his response, Wang says that such people do not have genuine knowledge, and goes on to quote [T6] to illustrate his idea: “The *Great Learning* points to genuine knowledge and action for people to see. It says they are ‘like loving lovely sights and hating hateful odors’” (故《大學》指個真知行與人看，說『如好好色，如惡惡臭』。 *IPL*, 5, *QJ*, 4). This passage has led many interpreters to believe that genuine knowledge has a close relationship to perception of objects in one’s environment. Indeed, as I said in the introduction, on what is arguably the most prominent view of genuine knowledge in scholarship written in English, episodes of genuine knowledge of filiality can at least sometimes have episodes of perceiving features of the environment as parts.

The introspective model is incompatible with this prominent view. A person’s perceiving features of their environment is a different event from their *liangzhi*’s knowledge that an inclination is good. It is also not *part* of the event of their *liangzhi*’s knowing that an inclination is good.

This passage – and the usual, prominent interpretation of it – might seem to give us reason to reject the introspective model. But, as I will now argue, it is in fact the usual interpretation of the passage, and not the introspective model, which should be rejected. In this passage, Wang does not say unambiguously that the examples from the *Great Learning* are instances of genuine knowledge. The examples are introduced by the word *ru* 如, which *can* mean “for example”, but can also equally naturally mean “like”, introducing an analogy or simile. Those who use this passage to argue that perception

⁴⁵Those attracted to the first response above can endorse a similar line of thought. The claim that an inclination is filial obviously entails that the inclination is good, and there is a clear sense in which if believing one claim conflicts with believing an obvious entailment of another, then believing the one claim is directly in conflict with believing the other. Since taking an inclination to be bad is in conflict with knowing that the inclination is good, there is thus a sense in which it is also in conflict with knowing that the inclination is filial. According to the first response, knowledge of filiality would be identified with *liangzhi*’s knowledge that an inclination is filial. So, anything which degrades this propositional knowledge that the inclination is filial would naturally be understood to degrade the objectual knowledge as well.

of the environment can be part of genuine knowledge understand the word *ru* as “for example” here. But there is strong independent evidence against this interpretation. In the following quotation, Wang responds to a correspondent who has argued that the passage from the *Great Learning*, [T6], describes how one should respond when confronted with a beautiful sight:

[T7] 人於尋常好惡，或亦有不真切處，惟是好好色，惡惡臭，則皆是發於真心。大學是就人人好惡真切易見處，指示人以好善惡惡之誠當如是耳，亦只是形容一「誠」字。今若又於好色字上生如許意見，却未免有執指為月之病。

Some of peoples’s ordinary loves and hates are not genuine and practical (真切). But loving lovely sights, and hating hateful [bad] odors all arise from the genuine mind (真心)...The *Great Learning* relies on the fact that it is easy to recognize the genuineness and practicality of love and hate that everyone has, to teach people what the wholeheartedness with which one loves goodness and hates badness [hatefulness] should be like. It just describes the one character “wholehearted”. From the way in which you now again put forward so many opinions about the words “beautiful sights”, it seems you have not avoided the mistake of “taking the finger to be the moon”. (QJ 5.218, cf. Ching (1972, 91))

“Taking the finger to be the moon” is a saying which describes someone who confuses the pointing finger with the object pointed at. According to Wang, his correspondent mistakenly focuses on a suggestive example (loving lovely sights), instead of focusing on the idea the example is used to illustrate (the wholeheartedness with which one ought to love what is good). In this sense, the correspondent takes the finger to be the moon.⁴⁶ Loving lovely sights, says Wang, is not an example of wholehearted inclinations. Instead, the *Great Learning* uses the example to illustrate what wholeheartedness is like. It is natural to think that Wang would have said the same about genuine knowledge: the examples of seeing a lovely sight or hating a hateful odor are not themselves examples of genuine knowledge; they merely illustrate the connection between knowledge and action. Interpretations which ascribe to Wang the view that perception is a component of genuine knowledge, by reading “*ru*” 如 as “for example” and taking seeing a lovely sight or hating a hateful odor to be examples of genuine knowledge, ignore Wang’s own advice. They mistake the finger Wang used, for the moon.⁴⁷

⁴⁶In a number of other passages, Wang also ties the passage from the *Great Learning* to a person’s wholehearted love of the good (IPL 229 QJ 110-11, IPL 318 QJ 135-6, and also in [T4] above; in QJ 7.276-8 Wang speaks of *liangzhi* as loving filiality and other virtues.

⁴⁷Can there be non-ethical objects of genuine knowledge on the present view? The interpretation here does not force a stance on this much-discussed question (see e.g. Cua (1982), Frisina (1989) (now Frisina

6 The unity of knowledge and action

I have now shown how the introspective model of genuine knowledge answers the challenge with which I began. On this model, Wang characterizes genuine knowledge independently of its relationship to action. He holds that genuine knowledge requires freedom from a form of doxastic conflict, and thus is elevated in a distinctively doxastic or epistemic respect.

But my discussion to this point has not yet touched on an important question: can the introspective model help us to understand how genuine knowledge could be “unified” with virtuous action? Since Wang introduces the notion of genuine knowledge in the course of responding to apparent counterexamples to the doctrine of the unity of knowledge and action, an adequate interpretation of genuine knowledge must explicate the role genuine knowledge plays in this doctrine, and in particular how genuine knowledge is connected to virtuous action. In this concluding section I will first show how the introspective model meets this important demand, before stepping back to briefly consider two further consequences of the model for our understanding of Wang Yangming.

At the beginning of section 4 I noted that “making inclinations wholehearted” and “extending knowledge” are two of the four tasks of personal ethical development described in the *Great Learning*. Essentially everyone in Wang’s tradition agreed that a person would be wholly virtuous if and only if they had successfully completed all four of these tasks. But, as I said there (n. 30), Wang himself held a further, distinc-

(2002, Ch. 4), Yang (2009), and Zheng (2018)). The general picture described by the introspective model is incompatible with the opening examples from *IPL* 5 (*QJ* 4) (along with the examples of knowing soup, knowing clothes and knowing the road one will travel on from *IPL* 132 (*QJ* 46-7)) being examples of, rather than analogues for, genuine knowledge. But that does not mean it forces us to the position that there are no non-ethical examples of genuine knowledge. For instance, later in *IPL* 5, Wang shifts his examples from colors and sights to hunger, cold and pain. Those examples are all plausibly states which are perceived by an “inner sense”. The letter of Introspective Knowledge, which describes knowledge that an inclination is good, does not apply to these cases, but the spirit of the introspective model is compatible with taking them to be examples of genuine knowledge, and perhaps they should be.

Within the ethical domain, I am uncertain whether Wang thinks there is genuine knowledge of bad ethical qualities. It is clear that he thinks *liangzhi* knows that inclinations are bad or wrong just as much as it knows that they are good or right. But allowing genuine knowledge of badness would allow that the variant of Genuine/Wholehearted could fail in either direction if “filiality” were replaced with “wrongness”: the conditions under which a person would have genuine knowledge of a bad quality are naturally taken to be conditions under which the person in fact was removing a bad inclination; it would certainly not be a case of having a wholehearted inclination to be preform a bad action. Since (as I will suggest in the next section) failures of Genuine/Wholehearted conflict with the general picture Wang seems to be developing, that is a point against taking there to be genuine knowledge of bad ethical qualities. But, there are also points in favor of taking there to be genuine knowledge of bad qualities. For instance, if one holds that the “extended knowledge” in [T4] is genuine knowledge, then Wang would be committed to there being genuine knowledge of bad ethical qualities.

tive view about the relationship among these four tasks: he held that a person would have completed one of them if and only if the person had completed them all. As a consequence, Wang held that having wholehearted inclinations is not just necessary for a person to be wholly virtuous (as essentially all who took the *Great Learning* to be a canonical text would have agreed), but also sufficient; he held that a person is fully virtuous if and only if their inclinations are wholehearted. Wang unquestionably held that there is an important connection between a person's being fully virtuous in this sense and their performing virtuous actions. So, he held that there is also an important connection between having wholehearted inclinations (and thus being wholly virtuous) and performing virtuous actions. Given this fact, if an account of genuine knowledge could make sense of an intimate connection between genuine knowledge and wholehearted inclinations, it would thereby make sense of a corresponding connection between genuine knowledge and virtuous action.

A broad array of reasonable interpretations of what Wang meant by the "unity of knowledge and action" (*zhi xing he yi* 知行合一) will see such a connection between genuine knowledge and virtuous action as sufficient to capture a key idea behind this slogan. In the slogan, the expression *he yi* 合一, which I have followed tradition in translating as "unity", could mean "identity", but it could also mean something much weaker, more like "correspondence", "co-relatedness" or "co-extensiveness". Those who understand "unity" to mean "identity" here, and who accordingly hold that Wang meant to say that knowledge and action are identical, will hold that a connection between genuine knowledge and virtuous action of the kind described in the previous paragraph is too weak to capture the core of Wang's doctrine. But on a wide array of interpretations (which in my view are much more plausible), "unity" is taken instead to mean "correspondence", "co-relatedness" or "co-extensiveness". On these views, the unity of knowledge and action centers on exactly the kind of weaker (but nonetheless intimate) connection between genuine knowledge and virtuous action sketched above. In what follows, I will focus on how my interpretation of genuine knowledge fits with this second kind of interpretation of the unity of knowledge and action.

At the start of section 5, I introduced the principle Genuine/Wholehearted as a natural extension of ideas in [T4] about the connection between extended knowledge and wholehearted inclinations. There, I used this principle simply to illustrate what an adequate account of genuine knowledge must achieve. But this principle articulates an intimate connection between genuine knowledge and wholehearted inclinations of exactly the kind that would help us to make progress in understanding the unity of

knowledge and action.⁴⁸ It is therefore of great significance that, against natural background assumptions, the introspective model – Liangzhi Knows Good (section 3), the third premise of the obscurity argument and its converse (n. 38), Extended/Wholehearted (section 4), and Introspective Knowledge (section 5) – in fact entails Genuine/Wholehearted.⁴⁹ It is for this reason that these claims are usefully taken together as the introspective model of genuine knowledge. For, given that the introspective model entails Genuine/Wholehearted, a broad array of interpretations of “unity” can agree that the introspective model provides an important step toward a full understanding of the unity of knowledge and action.

This point applies to many views on which “unity” in Wang’s slogan expresses something weaker than identity. But it is worth working an example of a particular such interpretation, to make the idea more concrete. The following principle articulates

⁴⁸It is worth noting that, in the key discussion of the unity of knowledge and action and genuine knowledge in *IPL* 5 (*QJ* 4), Wang appeals to the *Great Learning*’s discussion of wholehearted inclinations in [T6] (“loving a lovely sight, hating a hateful odor”) to illustrate what genuine knowledge is like. This fact provides further support for the idea that something like Genuine/Wholehearted is important to understanding the unity of knowledge and action itself.

⁴⁹The assumptions are: (i) a person genuinely knows filiality if and only if something is an episode of their genuinely knowing filiality; (ii) a person’s *liangzhi* knows that an inclination is good if and only if something is an episode of their *liangzhi*’s knowing that it is good; and (iii) a wholehearted inclination to perform a filial action is a wholehearted good inclination. I will understand the third premise of the obscurity argument and its converse (see n. 38) as: a person has extended knowledge that an inclination is good if and only if their *liangzhi* knows that the inclination is good, and they do not take the inclination to be bad. (Liangzhi Knows Good will not be used in the following derivation, but it features essentially in the argument for Extended/Wholehearted.)

Suppose a person genuinely knows filiality. Then by (i) something is an episode of their genuinely knowing filiality. Given that identical properties are coextensive (see n. 41), Introspective Knowledge implies that the episode of their genuinely knowing filiality is an episode of their *liangzhi*’s knowing that an inclination – which is in fact an inclination to perform a filial action – is good, and that the person does not take that inclination to be bad. By the first conjunct of this claim and assumption (ii), the person’s *liangzhi* knows that this inclination is good. By the second conjunct, they do not take the inclination to be bad. So, by (*), they have extended knowledge that the inclination is good. Extended/Wholehearted says in part that if a person has extended knowledge that an inclination is good, then they have this inclination, and it is wholehearted. So the person has a wholehearted inclination to perform a filial action.

Suppose a person has a wholehearted inclination to perform a filial action. By (iii) this inclination is a wholehearted good inclination. Extended/Wholehearted says in part that if a person has a wholehearted good inclination, they have extended knowledge that this inclination is good. By (*), their *liangzhi* knows that this inclination is good, and they do not take the inclination to be bad. By (ii) something is an episode of their *liangzhi*’s knowing that the inclination is good, and they do not take the inclination to be bad. By Introspective Knowledge (and the coextensiveness assumption in n. 41), this episode of *liangzhi*’s knowledge that an inclination – which is in fact an inclination to perform a filial action – is good, is an episode of genuine knowledge of filiality. By (i) they genuinely know filiality.

Note that this argument, along with the discussion which will follow in the main text, does assume that genuine knowledge is itself episodic (in the form of (i) and (ii)), and not a disposition to experience episodes of knowledge (see above n. 41). The argument and the claims below can easily be reformulated to accommodate a dispositional conception of genuine knowledge, but I will not present those reformulations myself.

a close connection between genuine knowledge of filiality and filial action:

Unity A person genuinely knows filiality if and only if they are acting filially.⁵⁰

Let us suppose that this principle captures a core part of what Wang means by “the unity of knowledge and action”. Suppose, moreover, that Wang endorses:

Inclination Action A person is acting filially if and only if they have a wholehearted inclination to perform a filial action.

Inclination Action and Genuine / Wholehearted together entail Unity. Moreover, against the backdrop of the introspective model, they provide a rich account of the unity of knowledge and action, as I will now explain.

Officially, in this paper I have not committed myself on how the notion of wholehearted inclinations should be understood; in section 4 I treated the notion as a black box. But on my favored interpretation, Wang holds that a given inclination of a person is wholehearted if and only if the person has no other inclinations which conflict with it. So, for instance, a person’s inclination to cool their parents in summer is wholehearted if and only if they have no other inclination which conflicts with it. If the person has an inclination to go hiking in the mountains on a given day, leaving their parents sweltering behind them, then even if in the end they stay to cool their parents, the inclination to cool their parents is not wholehearted. On this interpretation, to have a wholehearted inclination is to be free from motivational conflict of a particular form, and Inclination Action says that being free from this form of motivational conflict with respect to a filial inclination is necessary and sufficient for acting filially.

We can now see more vividly how the introspective model opens the way to a new understanding of the unity of knowledge and action as a whole. On the introspective model, to have genuine knowledge is in part to be free from a certain form of doxastic conflict. [T4] elucidates a connection between being free from this form of doxastic conflict on the one hand, and having wholehearted inclinations on the other. If having wholehearted inclinations is understood – as I have just suggested it can be – as freedom from motivational conflict, this argument forges a connection between freedom from a form of doxastic conflict on one side and freedom from a form of motivational conflict on another. The unity of knowledge and action can then be seen as consisting of two central claims: first, that a person is free from a relevant form of doxastic conflict if and only if they are free from a relevant form of motivational conflict, and, second, that

⁵⁰I argue for attributing Unity to Wang in [Lederman \(2021\)](#).

a person is free from this form of motivational conflict if and only if they are acting virtuously (as stated in Inclination Action).

I myself am attracted to this view of the unity of knowledge and action. But it is well beyond the scope of this paper to argue for Unity, Inclination Action, or my preferred understanding of Wang's view of wholehearted inclinations. Each of these ideas will be very controversial. For the purposes of our discussion here, I have simply wanted to show more concretely how, since the introspective model implies Genuine/Wholehearted, it opens the way to a new understanding of the unity of knowledge and action.

In closing, I want to highlight two broader consequences of the introspective model for our understanding of Wang Yangming. First, if my interpretation is correct, Wang's views on the relationship of knowledge and action differed strikingly from those of his Song dynasty predecessors as he understood them. According to Wang, Cheng Yi (程頤, 1033-1107) and Zhu Xi had held that ethical knowledge facilitates virtuous action in part through its role in deliberation. On this Cheng-Zhu orthodoxy as Wang interpreted it, knowledge comes first and virtuous action later (*IPL* 133, *QJ* 48). By contrast to this knowledge-first position, according to the introspective model, Wang advanced a position according to which knowledge, in a certain sense, comes last. On this interpretation, Wang held that genuine knowledge is not an ingredient in a reasoned process of deliberation; it is an automatic recognition of the goodness of one's own mental events. Genuine knowledge is closely connected to virtuous action: plausibly it is both necessary and sufficient for acting virtuously. But genuine knowledge is in an important sense a consequence of being in a state of mind which produces virtuous action; it does not drive or even facilitate that action.

Second, say that a person is *akratic* if and only if they know that an action is not among the best available to them but they voluntarily do it nevertheless. Wang's doctrine of the unity of knowledge and action is sometimes claimed to amount to the denial of the possibility of *akrasia*. But the introspective model casts doubt on this idea.

To see this, let us suppose that Wang is committed to Inclination Action. If he is, then given the introspective model (on which he is committed to Genuine/Wholehearted), he is committed also to Unity, not just for filiality but for other virtues as well. So he holds that if a person has genuine knowledge of a good ethical quality, they are acting in a way that exhibits that quality: it is not possible to act badly, and at the same have genuine knowledge of a good ethical quality. This claim is indeed in the vicinity of a denial of the possibility of *akrasia*. But it is not, strictly speaking, a denial of this possibility.

Wang does not characterize the relevant knowledge as knowledge of a proposition (i.e. that an action is not among the best available to a person). Moreover, he emphasizes genuine knowledge of good qualities, whereas whether someone is *akratic* depends on their knowing that an action is *not* among the best actions available to them. Finally, Wang is focused on genuine knowledge, whereas whether a person is *akratic* depends on what they just plain know.

In fact, it is natural to think that Wang would have held that *akrasia* is absolutely pervasive, and that any person who voluntarily performs a bad action suffers from *akrasia*. Whenever a person voluntarily acts badly, plausibly Wang would say that they act on the basis of a bad inclination. Whenever a person has a bad inclination, Wang holds that their *liangzhi* knows that it is bad. So, Wang seems committed to the claim that if a person voluntarily acts badly, their *liangzhi* knows that their inclination (and presumably the action they are performing on the basis of it) is bad. Wang does not speak about “best available actions”, but he also does not consider cases like ethical dilemmas which are typically used to argue that the best actions available to a person need not be good. So perhaps he would have held that an action is not among the best available to a person if and only if it is bad. If he did, it would then be natural for him to think that a person who knows that an action is bad, knows that the action is not among the best available to them. So he would be committed to the claim that anyone who voluntarily performs a bad action suffers from *akrasia*.

These remarks help to highlight an important theme. Wang holds that the conscience-like faculty of *liangzhi* always knows that inclinations are good, bad, right or wrong if they are. With *liangzhi*'s knowledge in the background, Wang can reduce substantive evaluation of people's mental states to an evaluation of the coherence of their mental states. On the view suggested in the previous paragraph, Wang holds that any person who voluntarily acts badly exhibits *akrasia*, that is, that the substantive badness of their action coincides with a form of incoherence. Similarly, on the introspective model itself, a person fails to act virtuously (or at least, fails to have wholehearted inclinations) if and only if they suffer from a form of doxastic incoherence.

References

- Angle, Stephen. 2018. Buddhism and Zhu Xi's Epistemology of Discernment. *Pages 156–192 of: Makeham, John (ed), The Buddhist Roots of Zhu Xi's Philosophical Thought.* Oxford University Press.
- Angle, Stephen C. 2005. Sagely ease and moral perception. *Dao*, 5(1), 31–55.
- Angle, Stephen C. 2010. Wang Yangming as a Virtue Ethicist. *Pages 315–335 of: Dao Companion to Neo-Confucian Philosophy.* Springer.
- Angle, Stephen C, & Tiwald, Justin. 2017. *Neo-Confucianism: A Philosophical Introduction.* Polity.

- Bol, Peter Kees. 2008. *Neo-Confucianism in history*. Cambridge, MA: Harvard University Asia Center.
- Chan, Wing-tsit. 1963. *Instructions for Practical Living and other Neo-Confucian Writings by Wang Yang-ming*. Columbia University Press.
- Chan, Wing-tsit (陳榮捷). 1983. 王陽明傳習錄詳註集評 *Wang Yangming's Instructions for Practical Living with detailed notes and collected discussions*. 臺灣學生書局.
- Chang, Carsun. 1955. Wang Yang-ming's Philosophy. *Philosophy East and West*, 5(1), 3–18.
- Chen, Lai (陈来). 1991. 有无之境：王阳明哲学的精神 *The Boundary between Being and Non-Being: The Spirit of Wang Yangming's Philosophy*. 北京: 人民出版社.
- Cheng, Chung-ying. 1974. Conscience, mind and individual in Chinese philosophy. *Journal of Chinese Philosophy*, 2(1), 3–40.
- Ching, Julia. 1972. *The philosophical letters of Wang Yang-ming*. Canberra: Australian National University Press.
- Ching, Julia. 1976. *To Acquire Wisdom: The Way of Wang Yang-ming*. Columbia University Press.
- Chu, Hung-lam. 1988. The Debate Over Recognition of Wang Yang-ming. *Harvard Journal of Asiatic Studies*, 48(1), 47–70.
- Cua, Antonio S. 1982. *The Unity of Knowledge and Action: A Study in Wang Yang-ming's Moral Psychology*. University Press of Hawaii Honolulu.
- Fraser, Chris. 2011. Knowledge and error in early Chinese thought. *Dao*, 10(2), 127–148.
- Frisina, Warren G. 1989. Are Knowledge and Action Really One Thing?: A Study of Wang Yang-ming's Doctrine of Mind. *Philosophy East and West*, 39(4), 419–447.
- Frisina, Warren G. 2002. *The Unity of Knowledge and Action: Toward a Nonrepresentational Theory of Knowledge*. State University of New York Press.
- Geaney, Jane. 2002. *On the epistemology of the senses in early Chinese thought*. Vol. 19. University of Hawaii Press.
- Graham, Angus Charles. 1958. *Two Chinese Philosophers*. London: Lund Humphries.
- Harbsmeier, Christoph. 1993. Conceptions of knowledge in ancient China. *Pages 11–30 of: Lenk, Hans, & Paul, Gregor (eds), Epistemological issues in classical Chinese philosophy*. State University of New York Press Albany, NY.
- Huang, Yong (黄勇). 2016a. 再论动力之知: 回应郁振华教授“Knowing To” Revisited: A Reply to Professor Yu Zhenhua. *学术月刊*.
- Huang, Yong (黄勇). 2016b. 王阳明的良知概念: 命题性知识, 能力之知, 抑或动力之知? *Knowing-that, Knowing-how, or Knowing-to: Wang Yangming's Conception of Moral Knowledge*. *学术月刊*, 49–66.
- Huang, Yong (黄勇). 2017. Knowing-That, Knowing-How, or Knowing-To? *Journal of Philosophical Research*, 42, 65–94.
- Ivanhoe, Philip J. 2000. *Confucian moral self cultivation*. Hackett Publishing.
- Ivanhoe, Philip J. 2002. *Ethics in the Confucian tradition: The thought of Mengzi and Wang Yangming*. Hackett Publishing.
- Ivanhoe, Philip J. 2009. *Readings from the Lu-Wang school of neo-confucianism*. Hackett Publishing.
- Ivanhoe, Philip J. 2011. McDowell, Wang Yangming, and Mengzi's contributions to understanding moral perception. *Dao*, 10(3), 273.
- Johnston, Ian, & Ping, Wang. 2012. *Daxue and Zhongyong: Bilingual Edition*. Chinese University Press.
- Kern, Iso (耿宁). 2010. *Das Wichtigste im Leben: Wang Yangming (1427–1529) und seine Nachfolger über die "Verwirklichung des ursprünglichen Wissens" (The most important thing in life: Wang Yangming [1472–1529] and his successors on the "Realization of Original Knowledge")*. Schwabe Verlag. Translated as 人生第一等事——王阳明及其后学论“致良知” by 倪梁康, published by 商务印书馆, 北京 2016; citations are to the translation.
- Lederman, Harvey. 2021 (July). *What is the "unity" in the "unity of knowledge and action"?* Unpublished MS.
- Lederman, Harvey. forthcoming. Conceptions of Genuine Knowledge in Wang Yangming. *Oxford Studies*

in Epistemology.

- Mou, Zongsan. 1973. The Immediate Successor of Wang Yang-ming: Wang Lung-hsi and his Theory of ssu-wu. *Philosophy East and West*, 23(1/2), 103–120.
- Nivison, David S. 1973. Moral Decision in Wang Yang-ming: The Problem of Chinese “Existentialism”. *Philosophy East and West*, 23(1/2), 121–137.
- Nivison, David S. 1996. *The Ways of Confucianism: Investigations in Chinese philosophy*. Open Court Publishing.
- Shi, Weimin. 2017. The Quest for Ethical Truth: Wang Yangming on the Unity of Knowing and Acting. *Comparative Philosophy*, 8(2), 7.
- Shun, Kwong-loi. 2011. Wang Yang-ming on Self-Cultivation in the *Daxue*. *Journal of Chinese Philosophy*, 38(s1), 96–113.
- Shun, Kwong-loi. 2018. On the Idea of ‘No Self’. *Proceedings and Addresses of the American Philosophical Associations*, 92, 78–107.
- Tang, Chun-i. 1973. The Criticisms of Wang Yang-ming’s teachings as raised by his contemporaries. *Journal of Chinese Philosophy*, 23(1/2).
- Tiwald, J., & Van Norden, B.W. 2014. *Readings in Later Chinese Philosophy*. Hackett Publishing Company, Incorporated.
- Tu, Wei-ming. 1976. *Neo-Confucian Thought in Action: Wang Yang-Ming’s Youth*. Berkeley: University of California Press.
- Wang, Fuzhi (王夫之). 1976. 尚書引義 *The Extended Meaning of the Book of Documents*. 中華書局.
- Wu, Guang (吳光) (ed). 2007. *Collected Works of Liu Zongzhou 劉宗周全集*. 浙江: 浙江古籍出版社.
- Wu, Guang (吳光), Qian, Ming (錢明), Dong, Ping (董平), & Yao, Yanfu (姚延福) (eds). 2011. *Collected Works of Wang Yangming 王陽明全集*. 上海: 上海古籍出版社.
- Yang, Xiaomei. 2009. How to Make Sense of the Claim “True Knowledge is What Constitutes Action”: A New Interpretation of Wang Yangming’s Doctrine of Unity of Knowledge and Action. *Dao: A Journal of Comparative Philosophy*, 8(2), 173–188.
- Yu, Zhenhua (郁振华). 2014. 论道德—形上学的能力之知—基于赖尔与王阳明的探讨 On Knowing How in Morality and Metaphysics: An Exploration of Gilbert Ryle and Wang Yangming. *中国社会科学*, 12.
- Yu, Zhenhua (郁振华). 2016. 再论道德的能力之知——评黄勇教授的良知诠释 Moral Knowing How Revisited: A Critical Examination of Professor Huang Yong’s Interpretation of Wang Yangming’s Notion of “Liangzhi”. *学术月刊*.
- Zheng, Zongyi (Cheng Chong-yi, 郑宗义). 2018. 再论王阳明的知行合一 A New Interpretation of Wang Yangming’s Doctrine of the Unity of Knowing and Acting. *学术月刊*, 5–19.
- Zhu, Xi (朱熹) (ed). 1983. *Collected Commentaries on the Four Books (四書章句集註)*. 北京: 中華書局.