



What Is the “Unity” in the “Unity of Knowledge and Action”?

Harvey Lederman¹

Accepted: 3 September 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

This essay argues for a new interpretation of the notion of “unity” in Yangming’s 王陽明 famous doctrine of the “unity of knowledge and action” (*zhi xing he yi* 知行合一). I distinguish two parts of Wang’s doctrine: one concerning training (*gong fu* 工夫), and one concerning the “original natural condition” of knowledge and action (*ben ti* 本體). I focus on the latter aspect of the doctrine, and argue that Wang holds, roughly, that a person exhibits knowledge in its original natural condition if and only if the person exhibits action in its original natural condition. Moreover, I argue that Wang denies that knowledge in its original natural condition is identical to action in its original natural condition.

Keywords Unity of knowledge and action · WANG Yangming 王陽明 · Moral psychology

1 Introduction

The great Ming 明-dynasty (1368–1644) philosopher WANG Shouren 王守仁 (Yangming 陽明, 1472–1529) saw his philosophical work in part as medicine for the moral maladies of his time. Wang claimed that the orthodoxy of that time—most associated with the Song 宋-dynasty (960–1279) philosopher ZHU Xi 朱熹 (1130–1200)—recommended that students intent on becoming virtuous separate their study into two phases: a first focused on acquiring and perfecting knowledge; and a second focused on applying this perfected knowledge in action. Wang argued that this division of learning into stages had led his contemporaries to become ineffectual. Since they thought they had to wait until they had perfected their knowledge before taking any action, “straight to the end of their lives they do not act.” As a “medicine to treat this

✉ Harvey Lederman
harvey.lederman@princeton.edu

¹ Department of Philosophy, Princeton University, 212 1879 Hall, Princeton, NJ 08544-1006, USA

disease,” he proposed his famous doctrine of “the unity of knowledge and action” (*zhi xing he yi* 知行合一).¹

Wang’s diagnosis of his contemporaries’ disease, and his strategy for treating that disease, are both tolerably clear. But Wang did not only describe the doctrine of the unity of knowledge and action as a form of medicine. He claimed also that “the original natural condition of knowledge and action are fundamentally like this [i.e., unified].” And here it is harder to make out his guiding idea. In describing the “natural condition” of knowledge and action, was he simply adding emphasis to his prescription for his contemporaries’ malady? Or was he making a new, different claim about the *nature* of knowledge and action—about how they in some sense really are? And if he did intend a new claim of this kind, did he endorse the radical idea that knowledge and action are identical—that they are, somehow, one and the very same thing?

These questions are intimately related to broader ones about the overall aims of Wang’s thought. Many philosophers would happily say that the questions they aim to answer, and the answers they give to these questions, have at most an obscure relationship to practical questions about how to make oneself a better citizen, friend, parent, or child. Wang, by contrast, held that one should only consider questions and theories which have some promise to help people with practical matters like these. This strong emphasis on ethical practice has led some to wonder whether Wang ever engaged in anything we might call “theoretical philosophy” at all. Did he intend to make claims that accurately describe the way things are, and which he hoped his hearers would believe (or reject) on the basis of evidence? Or did he instead offer something more like pills than propositions, ideas which might alter a person’s habits of mind through a pathway other than belief, and which should be assessed for their potency, not their truth?

In this essay, I propose and argue for an interpretation of the term “unity” in “the unity of knowledge and action,” with an eye to how this interpretation bears on broader questions about Wang’s aims.² I argue that Wang held, roughly, that a person exhibits the “original natural condition” of knowledge if and only if they exhibit the “original natural condition” of action, and that Wang did not hold that the original natural conditions of knowledge and action are identical. I show how some of Wang’s apparently radical methodological remarks about the therapeutic purpose of his doctrine on inspection support the view that he understood this claim about the original natural condition of knowledge and action to accurately describe knowledge and action. My investigation thus supports a conservative position about Wang’s aims, suggesting that

¹ Throughout the essay, I cite passages from the *Instructions for Practical Living* (*Chuanxi Lu* 傳習錄, hereafter *IPL*) by the section number of Wing-tsit CHAN’s editions (Chan 1963, 1983), followed by a page number in G. Wu, Qian, Dong, and Yao 2011 (indicated by “*QJ*”). Passages in Wang’s works outside the *IPL* are cited by the *juan* 卷 number, a period, and then the page number (e.g., “*QJ* 6.242”); I cite passages from Shu and Zha 2016 using “*QJBB*” and then a page number. Where available, I also cite pages in the translations of Ching 1972. This passage is taken from *IPL* 5, *QJ* 5; see below, [T2]. All translations in the essay are mine, although I have always consulted the translations in Chan 1963 and Ching 1972, as well as the translations in Ivanhoe 2009, and the revised versions of them, in Tiwald and Van Norden 2014, where relevant.

² The essay does not aim to offer a comprehensive interpretation of “the unity of knowledge and action” as a whole, but only of how we should understand the notion of “unity” in this doctrine. In developing my interpretation of this notion, I will to a great extent treat “knowledge” and “action” as placeholders, abstracting from substantive issues about how Wang understood these terms. I provide more substantive interpretations of the kind of knowledge operative in the doctrine in Lederman 2022, forthcoming.

Wang was indeed engaged in something naturally understood as “theoretical philosophy,” at least some of the time.³ Section 2 lays out evidence that Wang’s doctrine falls into two parts, one about ethical training (*gong fu* 工夫) and one about the “original natural condition” (*ben ti* 本體) of knowledge and action. Section 3, the core of the essay, focuses on the original natural condition of knowledge and action, and argues that Wang endorses a principle I call “Unity.” Section 4 presents a principle I call “Identity,” considers the merits of an interpretation which attributes it to Wang, and argues that Wang does not endorse it. Section 5 concludes, returning to broader issues about Wang’s aims.

2 Training and the Original Natural Condition

Wang’s striking slogan “the unity of knowledge and action” inspired his contemporaries to wonder whether he really meant what he seemed to say. In an important letter to Gu Lin 顧璘 (Dongqiao 東橋, 1476–1545), Wang quotes Gu as follows:⁴

Your letter says: “Genuine knowledge is just what is employed in performing an action. Without action, it is not sufficient to be called “knowledge”.’ If you set forth this teaching as an urgent measure (*chi jin li jiao* 喫緊立教) for those who are learning, to help them in their personal conduct, that is acceptable. But if you genuinely mean that acting just is knowing, I’m afraid that [students] will exclusively focus on seeking their original mind and leave behind the *li* of things, so that there must be some places where they will be stuck and can’t get through. How then could this be the sages’ established method of the joint advancement of knowledge and action?” (*IPL* 133, *QJ* 47)

Here Gu presents a sharp distinction between “an urgent measure for those who are learning” and “genuinely” asserting a doctrine. Gu’s description of an “urgent measure” calls to mind the Buddhist notion of an “expedient means” (Ch. *fangbian* 方便; Sk. *upaya*), broadly the idea that a teacher may sometimes assert falsehoods in order to help their students make progress (see, e.g., Watson 1993: 56–60). Gu expresses the concern that, if Wang’s doctrine is not to be understood merely as a

³ A word about how I will handle the secondary literature in this essay: many scholarly discussions of the unity of knowledge and action (Ching 1976: 66–68; Cua 1982; Ivanhoe 2002: 78–80, 99–100; Shun 2011: §II; Wu 2011: ch. 5; Yu 2014; Angle and Tiwald 2017: 127–131; Shi 2017; Cheng 2018; Van Norden 2019; Zheng 2019) include claims (and some arguments for those claims) which clearly *bear* on the question of the sense in which knowledge and action are unified, but they do not consider this question directly in its own right. Since it would require a great amount of space to settle what these authors’ views on my central question are, for the sake of space and tractability, I have largely (though not exclusively) focused on engaging in detail with authors who do discuss the question more thematically, for example Lao 1984–86/2019, Frisina 1989, L. Chen 1991, Lee 1994, L.-S. Chen 2015, Huang 2017.

⁴ Throughout this essay I will assume without argument that Wang’s views on the unity of knowledge and action remained consistent from 1509, when he first proposed the unity of knowledge and action, until his death in 1529. I will thus freely draw on works like this letter, which was written considerably later than 1509 (at least after 1524). Those skeptical of this assumption can see the essay as arguing for conditional claims about how we should understand Wang’s views, *if* they were consistent across this period.

false but possibly useful urgent measure, it will lead to a one-sided focus on inner development, to the exclusion of important external matters, indicated here by the expression “the *li* of things” (*wu li* 物理).⁵

In his reply, Wang accepts something like Gu’s distinction between modes of teaching (urgent vs. genuine), but rejects Gu’s implicit assumption that the two options exclude each other:

[T1] [a] [a1] 知之真切篤實處，即是行；行之明覺精察處，即是知。[a2] 知行工夫本不可離。[a3] 只爲後世學者分作兩截用功，失却知行本體，故有合一並進之說。[a4] 「真知即所以爲行，不行不足謂之知」……[a5] 此雖喫緊救弊而發，然知行之體本來如是，非以己意抑揚其間，姑爲是說以苟一時之效者也。

[a] [a1] The genuine, practical, earnest and effective aspects of knowledge are [a matter of] action. The lucid, perceptive, focused, discriminating aspects of action are [a matter of] knowledge.⁶ [a2] The training (*gong fu* 工夫) of knowledge and action at root cannot be separated. [a3] It is only because later generations have divided them, making them two stages of applying one’s effort (*yong gong* 用功), and losing the original natural condition of knowledge and action (*zhi xing ben ti* 知行本體), that I have proposed the theory of their unity and joint advancement. [a4] “Genuine knowledge is what is employed in performing an action. Without action, it is not sufficient to be called ‘knowledge’.” ... [a5] Although this is something I put forward as an urgent measure, to rescue people from a fault, knowledge and action are originally (*ben lai* 本來) like this in their natural condition (*zhi xing zhi ti* 知行之體). I did not follow my own inclinations to promote or demote one of them, temporarily endorsing this theory for its efficacy at this one time. (*IPL* 133, *QJ* 47–48)

In [a2] Wang says that the training (*gong fu*) for knowledge and action are inseparable. In [a3] and [a5] he says that in their “original natural condition” (*ben ti*) or just plain “natural condition” (*ti* 體) knowledge and action are unified. These remarks suggest that Wang sees his doctrine as consisting of two parts, one which describes the training of knowledge and action, and one which describes their original natural condition. And crucially, in response to Gu, he says that his doctrine is both an urgent measure, and an accurate description of knowledge and action in their original natural condition ([a5]).

The terms which I have translated as “training” (*gong fu*) and “original natural condition” (*ben ti*) are semitechnical terms for those working in Wang’s tradition, which are not readily interpretable on their own, so I will pause for a moment to discuss them. Throughout, I will translate *gong fu* as “training” for the sake of uniformity, although in some cases it might be better rendered as “practice” or even as “effort” or

⁵ I will not discuss in this essay the fraught question of how exactly we should understand *li* 理; the reader can treat the notion as a black box.

⁶ My translation of *ji* 即 here as “is a matter of” is controversial; I come back to how it should be translated in detail in Section 4.

“skill.”⁷ In English some forms of “training” have a built-in endpoint: a trainee doctor, for example, might engage in training that would be inappropriate for a doctor at a later stage of their career. *Gong fu* should not be understood as training of this kind. Instead, like “physical training” which is still required for athletes at the top of their game, *gong fu* may be required even for those who are already excellent in a given domain to maintain their level of excellence. The most common uses of *gong fu* that we will encounter are in the phrases the “training of knowledge” (*zhi zhi gong fu* 知之工夫) or “the training of action” (*xing zhi gong fu* 行之工夫). By these expressions, Wang means a regime dedicated to improving or maintaining the level of one’s knowledge or action; the expressions might be more colloquially rendered as “knowledge training” or “action training” on the model of “speed training” or “strength training.”

Our second phrase, “original natural condition” *ben ti*, which is often translated as “original substance,” is in some ways more complex. It is composed of two characters, one that I translate “original” (*ben* 本) and one that I translate “natural condition” (*ti*).⁸ Wang can use *ben ti* in two related but slightly different ways: he can speak of it as something which can be lost and must be restored, or he can speak of it as something which can never be lost.⁹ In my view, in the former cases, *ben ti* means roughly “the state something would be in if nothing was added to it and nothing interfered with it” (hence “original natural condition”), while in the latter it refers to (again, roughly) an internal feature of the thing which is responsible for the fact that it would be in the relevant state, if nothing was added to it or interfered with it (corresponding to one colloquial use of “nature” or “essence”).¹⁰ Below, it will be important to me that there is a use of *ben ti* on which it picks out a condition that can be lost. But otherwise, most of what I will say about the notion should be fairly uncontroversial. In particular, I will not be relying heavily on the glosses I just gave; those who already have views about how to understand the notion should be able to import them here.

⁷ My working hypothesis is that in the passages we will consider the homophonic expression *gong fu* 功夫 (note the variant first character) is an orthographic variant of the same word. So I will translate *gong fu* 功夫 also as “training,” though I always print the Chinese as well for those who wish to track the distinction. I translate the word *gong* 功, written with the first character of 功夫 (the variant), as “effort” throughout, though it too has a broader semantic range, and can mean “effects” or “results” as well.

⁸ “Natural condition” is my gloss on a single word; there is no distinct lexical or semantic item in the expression (*ti*) which corresponds to my “natural.” In particular, there is no use of any cognate for the word commonly (and reasonably) translated as “nature” (*xing* 性); the concept of *xing* and that of *ben ti* have a conceptual relationship, but not an etymological one. When it does not occur in the phrase *ben ti* the character *ti* has a considerably broader semantic range than “natural condition.” Fortunately every use of *ti* we will encounter in this essay—like the one in [a5] above—is most naturally understood as an abbreviation for *ben ti*, so we do not need to engage with this broader use in detail.

⁹ For instance, just considering Wang’s uses of the expression “the original natural condition of the mind” (*xin zhi ben ti* 心之本體), we find Wang saying that it can be “lost” (*shi* 失) in: *IPL* 34, *QJ* 17; *IPL* 222, *QJ* 108–109; cf. *IPL* 204, *QJ* 104. He says that it must be “restored” (*fu* 復), for example, in *IPL* 121, *QJ* 40; *IPL* 127, *QJ* 43 (*xing zhi ben ti* 性之本體); *IPL* 145, *QJ* 66 (repeated in *IPL* 169, *QJ* 81); *IPL* 237, *QJ* 112–113; *QJ* 5.216, compare with Ching 1972: 87 as well as *IPL* 101, *QJ* 34. But in other passages he describes it as something which everyone has, and can only be obstructed, not lost: *IPL* 48, *QJ* 20; *IPL* 152, *QJ* 69; *IPL* 155, *QJ* 70–71; *IPL* 221, *QJ* 108; compare with *QJ* 7.271 (*ben ti zhi zhi* 本體之知).

¹⁰ Mou Zongsan 牟宗三 proposes that the expression be translated “in-itself,” so that the present phrase would be “knowledge and action in themselves” (Mou 1972, see also the English translation of part of the article as Mou 1973); Angle and Tiwald propose “inherent reality” (Angle and Tiwald 2017). Both of these translations (reasonably) emphasize the second use of *ben ti*, while mine emphasizes the first (since that use will be more relevant to us here). But what I say above is meant to be compatible with Mou’s and Angle and Tiwald’s understanding of the term.

As we will see, Wang uses these two terms—“training” and “original natural condition”—comparatively frequently to distinguish between two parts of his doctrine. But Wang also has a second way of marking what seems to be the same distinction. In [a2] above, Wang articulates two problems with his contemporaries: first, a fault connected to effort (*gong* 功) (which is etymologically and conceptually related to training [*gong fu* 工夫/功夫]); second, one associated with “original natural condition” (*ben ti*). He then sums up his own positive doctrine—the antidote to these problems—with the formula “unity and joint advancement” (*he yi bing jin* 合一並進). The parallelism between the foregoing description and this two-part formula, together with the fact that the meaning of “joint advancement” ties it closely to “training,” makes it natural to see the formula as describing the same two aspects of Wang’s doctrine, with “joint advancement” describing the “training” for knowledge and action, and “unity” describing their “original natural condition.” (The shorter formula, “unity of knowledge and action,” presumably uses “unity” to cover both aspects.) In other passages, too, Wang connects the expanded slogan to the two parts of his doctrine: when they occur together, “unity” is associated with original natural condition, and “joint advancement” with training.¹¹

The idea that Wang drew something like this distinction, and even that it was central to his understanding of the doctrine of the unity of knowledge and action, is not new to this essay.¹² I have spent some time rehearsing evidence that he did draw it, for two main reasons. First, not everyone agrees about which terms of Wang’s are associated with which aspects of the doctrine. At several points below, I will defend my interpretation of Wang’s views about the unity of the original natural condition of knowledge and action by dismissing certain passages on the grounds that they do not describe the original natural condition of knowledge and action, but instead describe training. In making these arguments, I will rely on claims I have argued for here, about which terms are associated with which parts of his doctrine.

Second, the distinction between training and original natural condition provides us with a way into our broader question about Wang’s philosophical aims. In [T1], Wang says that he sees his doctrine both as an “urgent method” and as offering a description of the “original natural condition” of knowledge and action. The passage reveals not

¹¹ *IPL* 132, *QJ* 46 ties “joint advancement” to training. In *IPL* 136, *QJ* 52, Wang contrasts “the natural condition of the unity of mind and *li* 理 with the effort of the joint advancement of knowledge and action.” In one passage, Wang even describes “joint advancement” as *leading* to unity: “This is what the ancients’ learning took to be the effort of jointly advancing knowledge and action, so that they are unified” (*QJ* 8.308).

¹² Something like it is recognized by, for example, Tu 1976: 150–153; Lao 1984–86/2019: 3.422–423; L. Chen 1991: 93–108; Shun 2011: 98–99; Wu 2011: 91–97; L.-S. Chen 2015: 7–8.

In support of the claim that the distinction was central to Wang’s view of the doctrine, note that Wang explicitly describes this contrast or implicitly relies upon it in a wide range of further passages in the corpus, which cover a period of many years. In *IPL* 136, *QJ* 52, Wang discusses a contrast between a unity doctrine related to the “natural condition” (*ti*) of mind and *li*, with a “joint advancement” doctrine associated with effort (*gong*). In *IPL* 165, *QJ* 78, he contrasts understanding “knowledge and action” in regard to “applying effort” (*yong gong* 用功) with understanding them in regard to their “original natural condition,” cautioning his correspondent to “examine the two characters closely.” *QJ* 6.232 contrasts the way the ancients spoke of training with the way knowledge and action are unified in their “natural condition and structure” (Ching 1972: 107). In other passages Wang highlights just “training” in connection to knowledge and action (*IPL* 26, *QJ* 15 [*gong fu* 功夫]; *IPL* 132, *QJ* 46 [*gong fu* 工夫]) or just “natural condition” ([T4]). He also draws the full contrast in connection to “inner and outer” (an issue related for him to the unity of knowledge and action) in *IPL* 204, *QJ* 104.

only that he does recognize a possible distinction between these two projects, but also that he understands the distinction between these two modes of teaching in part as a distinction between teaching aimed at practical applications and teaching aimed at describing knowledge and action. In [a5], in particular, when Wang says that his doctrine is not *only* an urgent measure, and also not *only* something he merely endorsed for its efficacy at one time, it is hard to resist a reading where he means to say that, in addition to having these practical aims, his doctrine also accurately describes knowledge and action.

In the rest of this section, I will support my claim that Wang's views about the original natural condition of knowledge and action are intended to accurately describe the way things are. This discussion is important to my overall project in the essay. But it will be something of a digression from the main line of argument, which is more narrowly focused on the notion of unity. So those who are eager to get to the discussion of unity may wish to skip to the next section at this point.

I will support my claim that Wang means his doctrine to accurately describe knowledge and action by responding to an important objection to it. The objection is based on the following passage, a record of a conversation between Wang and his favorite student XU Ai 徐愛 (1487–1517). I will quote the passage at some length, because it gives us further evidence about how Wang uses the distinction between training and original natural condition, and also because it will be important not only here, but in later sections of the essay as well (especially Section 4):

- [T2] [b] [b1] 先生曰：「…… 此便是知行的本體，不曾有私意隔斷的。聖人教人，必要是如此，方可謂之知。不然，只是不曾知。[b2] 此却是何等緊切着實的工夫！[b3] 如今苦苦定要說知行做兩箇，是甚麼意？某要說做一箇，是甚麼意？[b4] 若不知立言宗旨，只管說一箇兩箇，亦有甚用？」
- [c] 愛曰：「古人說知行做兩箇，亦是要人見箇分曉，一行做知的功夫，一行做行的功夫，即功夫始有下落。」
- [d] 先生曰：「此却失了古人宗旨也。……」
- [e] [e1] 「今人却就將知行分作兩件去做，[e2] 以為必先知了然後能行，[e3] 我如今且去講習討論，做知的工夫，待知得真了，方去做行的工夫，故遂終身不行，亦遂終身不知。此不是小病痛，其來已非一日矣。[e4] 某今說箇知行合一，正是對病的藥。又不是某鑿空杜撰，知行本體原是如此。[e5] 今若知得宗旨時，即說兩箇亦不妨，亦只是一箇；若不會宗旨，便說一箇，亦濟得甚事？只是閒說話。」

[b] [b1] The Master said: "... This then is the original natural condition of knowledge and of action, which have not been divided by selfish inclinations. The sage taught people that only a person in this state can be said to know. If they are not in this state, then they do not yet know. [b2] But (*que* 卻) how urgent, practical and effective this training [*gong fu*] is! [b3] What is the goal [*yi* 意] of those today who strenuously and persistently say that knowledge and action are two things? What is the goal [*yi*] of my saying that they are one thing? [b4] If you do not know [*zhi* 知] my purpose [*zong zhi* 宗旨] in setting forth these words, and only care about saying that they are one or that they are two, what use would it have?"

[c] Xu Ai said: “In saying that knowledge and action are two things, the ancients also wanted people to see distinctly that one thing (*yi xing* 一行) is the training (*gong fu*) of knowledge, and that another thing (*yi xing*) is the training of action, since only in this way will one’s training have direction.”

[d] The Master said: “But this is to lose [sight of] the purpose [*zong zhi*] of the ancients....

[e] [e1] “People now rely on the claim that knowledge and action are distinguished into two things when they act. [e2] They think that one must first know and only then is one able to act, [saying] [e3] ‘For the time being I will discuss and debate, engaging in the training (*gong fu*) of knowledge. Only after my knowledge has become genuine, will I engage in the training of action.’ For this reason, straight to the end of their lives they do not act, and straight to the end of their lives they do not know. This is not a minor disease, and it has not come about in a single day.

[e4] My doctrine of the unity of knowledge and action is medicine to treat this disease. But it is also not a baseless fabrication: the natural condition of knowledge and action are fundamentally like this. [e5] If now you know my purpose, then even if you say that they are two it does not matter, since they are just one thing. If you do not understand my purpose, then if you say they are one, what affair could it help with? You will just be talking idly.” (*IPL* 5, *QJ* 4–5)

The objection I will consider starts from Wang’s remarks in [b3–4] and in [e5]. There, Wang says that he does not care especially whether a person believes the truth about knowledge and action; all that matters to him is whether they understand the aim of his doctrine. The objector claims that these remarks show that Wang either does not accept the idea that there are facts about knowledge and action, or at least does not care to articulate them. As a result, Wang’s claims about the original natural condition of knowledge and action should not be understood as describing knowledge and action, but as aimed at some other target.

In response, I agree with the objector that Wang strongly prioritizes practice over theory, and that the goal of this passage is to emphasize the priority of practice. But I disagree that Wang denies that a part of his doctrine can be understood to accurately describe the way things are. In fact, I believe that in this passage Wang shows quite clearly that he understands his doctrine about the original natural condition to accurately describe an aspect of knowledge and action.

Before I argue for this claim, I want to establish a few facts about how Wang deploys the distinction between training and original natural condition in our passage. The passage is the final section of a longer discussion of the original natural condition of knowledge and action (see below [T3]). The text in [b1] marks the formal conclusion of Wang’s comments on that topic. In [b2] Wang transitions to a discussion of training, indicating the shift in topic with a contrastive particle (*que* 卻). In [c], Xu Ai picks up the distinction Wang has just drawn, saying that in distinguishing knowledge and action, the ancients “also” (*yi* 亦) wanted to propose a doctrine about training. In the sequel, Wang responds to Xu by focusing his own comments on training both explicitly in text from [d] elided above, and also in [e2] and [e3]. In [e4], he sums up the discussion by reiterating that his doctrine is both promulgated as a medicine to treat a

disease, and also intended to describe the original natural condition of knowledge and action. The passage is thus a beautiful example of how Wang and his interlocutors rely on and utilize the distinction between training and original natural condition quite systematically.

With this paragraphing before us, we can confront the objection head on. Wang says in [e4], that his doctrine is not only therapeutic, but *also* that the original natural condition of knowledge and action are like this; he clearly holds that his doctrine accurately describes how they would be in this state. *Pace* our objector, Wang does not deny the distinction between correct description and practical application; on the contrary, he explicitly draws that distinction. Moreover, [e5] provides a further striking piece of evidence for my interpretation and against the objector. There, Wang says—as he said also in [b4]—that understanding the purpose of his doctrine is key to using it for practical purposes. In fact, he says, understanding this purpose is so important, that even someone who denies the letter of the doctrine can still achieve a virtuous state where their knowledge and action would be unified, provided they understand its purpose. In this comment, even as Wang emphasizes the importance of practical consequences over theoretical correctness, he still recognizes that when this person says that knowledge and action are two, they *will be speaking incorrectly*, because their knowledge and action will in fact be one. Far from obliterating the distinction between practical mistakes and mistakes in one’s beliefs (or speech), Wang carefully toes his way around this distinction in the way he describes the special case of someone who understands the purpose of his doctrine, but rejects its letter.

Everyone should agree that Wang strongly emphasizes practice over theory. But he does also recognize a distinction between assessing claims for usefulness and assessing them as correct descriptions of the way things are. Moreover, the texts we have examined strongly suggest that he intends at least his claims about the original natural condition of knowledge and action to fall in the latter category. These points will be important later on, when I defend my interpretation of Wang’s understanding of “unity” against various objections. But they are also of interest in their own right, as evidence for a conservative understanding of Wang’s philosophical aims.

3 Unity

In the previous section, we saw that the doctrine of the unity of knowledge and action has two parts: one about the training for knowledge and action, and one about their original natural condition. The main goal of the remainder of the essay will be to argue for an analysis of how Wang understands the notion of “unity” as applied to the original natural condition of knowledge and action.

Before we turn to that topic, however, it is worth saying something about how Wang understands the notion of “unity” on the training side of his doctrine. In this connection, Wang writes: “Knowledge and action are originally (*yuan* 原) two words which describe one training (*gong fu*). This one training must be described by these two words. Only then will you have described it completely without defects” (*QJ* 6.233; Ching 1972: 107–108). In this passage, it is fairly clear that Wang understands identity as at least one operative notion of “unity” for the training

aspect of his doctrine; he holds that the training for knowledge and the training for action are one and the same.¹³

Unfortunately, however, there is no similarly crisp statement in the surviving texts of Wang's views about the notion of unity operative for the original natural condition of knowledge and action. The rest of this section will be devoted to teasing his views on this topic out of the texts. I will first argue that Wang can describe knowledge in its original natural condition as “genuine knowledge” and that he thinks of action in its original natural condition as action exhibiting some virtue, for instance filial piety (*xiao* 孝, hereafter “filiality”) or fraternal respect (*ti* 悌, hereafter “respect”) (Section 3.1). I will then argue that for a range of traditional virtues like filiality and respect, Wang endorses claims of the form:

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

I will argue for attributing this principle (and related ones for other virtues) to Wang in two steps.¹⁴ First, I will argue that Wang endorses the right-to-left direction of Unity,

¹³ In some other passages Wang frames the “unity” of knowledge and action in training (*gong fu*) slightly differently. There, he seems to understand the “unity” on the training side as the idea that phases of “knowing” and phases of “acting” alternate in such a way that it is impossible to distinguish a first “knowing” phase of training from a second “acting” phase. In famous examples from the letter to Gu Dongqiao (*IPL* 132, *QJ* 46–47 and *IPL* 136, *QJ* 51–52), where Wang is more or less explicit that he is discussing training and not the original natural condition of knowledge and action, Wang has this second point most firmly in view. In the first of these passages, Gu claims that one knows food and soup before eating them, knows clothes before wearing them, and knows a road before traveling it. Wang responds that in each case a desire, that is, an inclination, which is a part of action, precedes some relevant knowledge—presumably knowing the food, soup, clothes, and road by sight—so this knowledge does not entirely precede action. Wang also argues that an important form of knowledge—bodily experience of the taste of the food and soup, and of the feel of the clothes and road—comes after the physical action, so one only gains complete knowledge after some action. The same idea about the alternation or interpenetration of stages of knowing and acting is in the background in *IPL* 136, *QJ* 51, where Wang discusses how one must engage in actions to learn how to shoot a bow or how to write. His point is not that the nature of the relevant knowledge is such as to require antecedent action, but rather that some action is involved in typical examples of acquiring the relevant form of knowledge. This latter idea illustrates his main thought, which is that phases of action and phases of knowing must both occur on the path to the ideal state of virtue.

There are a number of ways of understanding this claim about alternation, so that it is compatible with my claim in the main text that the training of knowledge and action are identical. Perhaps Wang uses the terms “knowledge” and “action” somewhat differently here: he is not describing the training in terms of the ideal state of knowledge or action at which it aims—neither form of knowledge or action is knowledge or action in their original natural condition—but rather in terms of how the activities involved in the training might be conventionally described. Alternatively, he may use the word “training” differently in the two sets of passages. A third idea (compatible with the first two, but independent of them) is that he means to use these examples to motivate his main idea (that the training for knowledge and action are identical), by illustrating that they do not occur in phases which can be neatly separated. In either case, it seems to me, this different way of describing training can be seen as complementary to, and not in tension with, the central claim that the training of knowledge and action is identical.

¹⁴ I believe Wang also endorses variants of this principle for other virtues: for conscientiousness (*zhong* 忠), humaneness (*ren* 仁), and compassion (*ce yin* 惻隱). When I use the name “Unity” (and also the names of similar principles, which will only be stated explicitly for filiality) I will sometimes mean not just the claim displayed above in particular, but the whole family of principles. For respect, see *IPL* 5, *QJ* 4. For conscientiousness and humaneness, see *IPL* 139, *QJ* 56. For compassion, see *IPL* 8, *QJ* 7; *IPL* 135, *QJ* 50–51. I follow tradition in translating *ce yin* 惻隱 as “compassion” but the term might be better rendered as “being pained by” or “unable to bear”; see for example Shun 2018: 90 for discussion.

that is, (AK) if a person is acting filially, they genuinely know filiality (Section 3.2). Second, I will argue that Wang endorses the left-to-right direction of Unity, that is, (KA) if a person genuinely knows filiality, they are acting filially (Section 3.3). The conjunction of AK and KA is equivalent to Unity; so, if Wang endorses both of them, he endorses Unity.

My discussion throughout the section will be centered on one key passage. Since the details of this passage will be important throughout, I will quote it in full here at the start. Prior to the passage, Xu Ai asks Wang about an apparent difficulty with the doctrine of the unity of knowledge and action, and Wang asks him to give an example illustrating his point. Xu replies:

For instance, today everyone knows that they should be filial to their parents, and that they should be respectful to their older brothers, but they are unable to be filial, and unable to be respectful. So in this case, knowledge and action are separated, and are clearly two things. (*IPL 5, QJ 4*)

Xu's example is related to examples of *akrasia*; the people he describes know that they ought to perform an action, but they voluntarily fail to do it nevertheless. Xu takes this kind of case to threaten the unity of knowledge and action, presumably because these people's actions are not responsive to what they know about what they should do. In his reply, Wang defends the doctrine by introducing a distinction between the knowledge and action of the people Xu describes, and knowledge and action in their original natural condition. Wang's reply occupies the text immediately preceding [T2]; the last sentences of this excerpt ([h5]) are the first sentences of that one ([b1]):

[T3] [f] [f1] 此已被私慾隔斷，不是知行的本體了。[f2] 未有知而不行者。知而不行，只是未知。[f3] 聖賢教人知行，正是安復那本體，不是着你只恁的便罷。
 [g] [g1] 故《大學》指個真知行與人看，[g2] 說「如好好色，如惡惡臭」。[g3] 見好色屬知，好好色屬行。只見那好色時已自好了，不是見了後又立個心去好。[g4] 聞惡臭屬知，惡惡臭屬行。只聞那惡臭時已自惡了，不是聞了後別立個心去惡。[g5] 如鼻塞人雖見惡臭在前，鼻中不曾聞得，便亦不甚惡，亦只是不曾知臭。
 [h] [h1] 就如稱某人知孝、某人知弟，必是其人已曾行孝行弟，方可稱他知孝知弟，不成只是曉得說些孝弟的話，便可稱為知孝弟。[h2] 又如知痛，必已自痛了方知痛，[h3] 知寒，必已自寒了；[h4] 知饑，必已自饑了；[h5] 知行如何分開？此便是知行的本體，不曾有私意隔斷的。聖人教人，必要是如此，方可謂之知。不然，只是不曾知。

[f] [f1] In this case, knowledge and action have already been divided by selfish desires; they are no longer in the original natural condition (*ben ti*) of knowledge and action. [f2] No one has ever known but failed to act. If one knows but does not act, one simply does not yet know. [f3] In regard to knowledge and action, the sages and worthies taught people to stabilize and restore that original natural condition; they did not order people to do any old thing and then just stop.

[g] [g1] For this reason, the *Great Learning* points to genuine knowledge and action for people to see. [g2] It says they are “like loving lovely sights and hating hateful odors.”¹⁵ [g3] Seeing a lovely sight belongs to knowledge, while loving a lovely sight belongs to action. But when someone sees a lovely sight, he already at that time automatically loves it. It is not that after seeing it he additionally makes up his mind to love it. [g4] Smelling a hateful odor belongs to knowledge, while hating a hateful odor belongs to action. When someone smells a hateful odor, he already at that time automatically hates it. It is not that after smelling it he separately makes up his mind to hate it. [g5] It’s like a person with his nose blocked: even if he sees something with a hateful smell in front of him, in his nose, he has not smelt it. So while he doesn’t really hate it, this is only because he does not yet know the odor.

[h] [h1] The same goes for saying that someone knows filial piety or that someone knows fraternal respect. They must have at some point acted filially or acted respectfully, before they can be said to know filial piety or fraternal respect.¹⁶ If a person merely knows how to say some filial or respectful words, that’s not enough for it to be acceptable to say that they know filial piety or fraternal respect. [h2] Knowledge of pain is also like this. [h3] One must have been in pain oneself to know pain. [h4] One must have been cold oneself to know cold. One must have been hungry oneself to know hunger. How then can knowledge and action be separated? [h5] This then is the original natural condition of knowledge and action, which have not been divided by selfish inclinations. The sage taught people that only a person in this state can be said to know. If they are not in this state, then they do not yet know. (*IPL 5, QJ 4*)

3.1 Genuine Knowledge and Virtuous Action

My first task will be to make a little progress on what Wang takes the original natural condition of knowledge and action to be.

Wang says (in [f]) that the knowledge and action in Xu’s example are no longer knowledge and action in their original natural condition. He goes on (in [g1]) to describe this form of knowledge and action as “genuine knowledge and action.” In [f2] Wang uses “know” and “act” without any qualification, but in both [f1] and [f3] he is explicit that he is considering the original natural condition of knowledge and action, so it is clear that in [f2] and in the passage as a whole he has this “genuine knowledge and action” (and not just plain knowledge and action) in mind. He does not intend to claim that the people in Xu’s example have no knowledge *at all*, anymore than he

¹⁵ A more literal translation might be “like loving a beautiful sight, and hating a bad odor.” I have opted for the translation in the main text in an attempt to imitate the fact that “love” is written with the same character (*hao* 好) as the adjective I have translated “lovely” (although they are pronounced differently) and the verb “hate” is written with the same character (*wu* 惡) as the adjective I have translated “hateful” (although they too are pronounced differently).

¹⁶ I defend the “objectual” reading of these examples in detail in Lederman 2022: §2, cf. Lederman forthcoming: §3.

means to claim that they fail to act at all. Instead, he seems to accept that they exhibit some form of knowledge and action, and to deny only that they exhibit knowledge and action in their original natural condition.

In [g1] Wang describes both knowledge and action as genuine, but it is clear in other passages that Wang takes “genuineness” to be most importantly an attribute of an ideal form of knowledge, and not of an ideal form of action. Later in this same passage, Wang uses “genuine” in application to knowledge alone (this was quoted above in [T2] [e3], *zhi de zhen* 知得真, there translated as: “knowledge has become genuine”). In [T1] [a], in a set phrase which Wang repeats several other times (see Section 4 below for extensive discussion), “genuine” is on the list of attributes of “knowledge,” not of action. In [T1] [d] Wang describes “genuine knowledge” (*zhen zhi* 真知) specifically as “what is used in performing an action” (*yi wei xing* 以為行). And finally, in a striking passage where Wang says that it is as hard to convey what it is like to be in an ideal state of virtue as it is for a mute person to describe the taste of bitter melon, XU Ai—who happens to be standing by—says that “it is only in this way that it is genuine knowledge, which is [a matter of] acting” (*ru ci cai shi zhen zhi, ji shi xing yi* 如此才是真知, 即是行矣) (IPL 125, QJ 42). Together these passages provide strong support for the claim that Wang could use “genuine knowledge” as a technical term to describe the form of knowledge relevant to the unity of knowledge and action. By contrast, there are no passages at all in which Wang speaks of “genuine action” on its own. So, while Wang does use “genuine” here to describe knowledge and action together, it seems clear that he typically associates genuineness with the elevated form of knowledge relevant to the unity of knowledge and action, not with an elevated form of action.

The idea that Wang emphasizes “genuine knowledge” (and not “genuine action”) is further supported by the fact that this expression had a venerable history predating Wang, some of which would have been well known to him. The expression is used as early as the *Zhuangzi* 莊子 (3rd c. BCE) (“The Great and Most Honored Master,” 1), but much more importantly for Wang, Song-dynasty authors used the term fairly extensively. In a pair of famous passages, CHENG Yi 程頤 (1033–1107), for example, says that someone who has previously been mauled by a tiger will change his countenance at the news that a tiger is roaming the countryside, whereas people who have never encountered a tiger may know that tigers are to be feared, but they will not change their appearance. In the first of these passages, Cheng describes this difference in people’s response as illustrative of the difference between genuine knowledge (*zhen zhi* 真知) and ordinary knowledge (*chang zhi* 常知), stating that it is the former which is key to virtuous action (Wang 2004: 2A.16). In the second, Cheng adds a second example to illustrate the idea, saying that rich people who have tasted roasted fish respond to its smell differently than poorer people who have never tasted it (Wang 2004: 18.188). Cheng’s hugely influential successor ZHU Xi also uses this term in many places, in some cases even referring directly to Cheng’s examples.¹⁷ Wang would certainly

¹⁷ For just a few of the many instances of “genuine knowledge” in ZHU Xi, see for example Li and Wang 1986: 15.302, 303, 309, the last of which discusses the example of the tiger. For further discussion of these precedents in English see Shun 2010: 188; Angle 2018: 166; and Huang 2015: ch. 3.

have known about Cheng and Zhu's use of the expression "genuine knowledge," so their use provides further circumstantial evidence that Wang could use "genuine knowledge" to describe the ideal form of knowledge relevant to the unity of knowledge and action (i.e., the original natural condition of knowledge). Since these authors did not use the expression "genuine action" in a similarly high-profile way, their usage also provides some evidence that Wang would in the first instance have thought of "genuineness" as distinctively associated with the elevated form of knowledge, not the elevated form of action.

What about the original natural condition of action, which Wang here (but not in general) describes as "genuine ... action"? In the passage above, Wang begins to give a direct positive discussion of Xu's examples, of filiality and respect in [h]. In [h1] he speaks of knowing filiality and knowing respect on the one hand, and enacting filiality and enacting respect on the other. These statements are clearly intended to describe the original natural condition of knowledge and action, and to codify the key relationship underlying his doctrine; Wang describes not just any old knowledge of filiality and respect, but rather *genuine* knowledge of filiality and respect. On the action side, Wang has previously given no indication of how specifically to qualify the ideal form of action. His discussion here seems designed to fill in this gap, telling us that the action relevantly related to (genuine) knowledge of filiality is filial action and that the action relevantly related to (genuine) knowledge of respect is respectful action. So the passage suggests that this aspect of his doctrine connects (genuine) knowledge of filiality or respect on the one hand, with filial and respectful action on the other. More generally, Wang's comments in [h] suggest that Wang understands the elevated form of action relevant to the unity of knowledge and action as virtuous action.

3.2 AK

There is a case to be made that knowledge and action in their original natural condition are on the one hand genuine knowledge, and, on the other, virtuous action. How does Wang understand the relationship between these two? I will now argue, on the basis of the examples in [g], that Wang holds that they begin at the same time. This will pave the way for my argument that Wang endorses AK, that is, that if a person is acting filially, they genuinely know filiality.

Wang interprets the two examples from the *Great Learning* (quoted in [g2]) as saying that loving a lovely sight begins no later than seeing it ([g3]) and that hating a hateful odor begins no later than smelling it ([g4]). He furthermore says that seeing and smelling "belong" to knowledge, while hating and loving "belong" to action. In what follows, I will describe seeing, smelling, loving, and hating as "stand-ins" for knowledge and action respectively. This terminology is intended to be neutral on whether Wang holds that these examples are themselves examples of the elevated kind of knowledge and action he has in mind, or whether he intends them merely as analogues for this elevated form of knowledge and action, not instances of it. Whichever way we understand the examples, Wang certainly holds that facts about the stand-ins help to illustrate his views about the relationship between knowledge and action, and this will be enough for our purposes here. In [g3–4], he says twice that the relevant stand-in for action begins no later than the

relevant stand-in for knowledge, suggesting that he holds that virtuous action similarly begins no later than the relevant genuine knowledge.¹⁸

In [g5], Wang presents a different way of spelling out his example of the odor. He describes someone who sees an object which has a bad smell, but who cannot smell it themselves because their nose is blocked. On my preferred reading of the example, Wang imagines that the person knows that the object they see has a bad smell, perhaps on the basis of induction, inference, or testimony. Read in this way, the example gives a fairly exact parallel to XU Ai’s cases: this person exhibits what might seem to be a stand-in for knowledge (they know that there is a bad odor, and hence that it is worthy of hate), but they do not exhibit the stand-in for action (since they do not experience sensations of hatred). And Wang gives an exactly parallel diagnosis: he says that the person does not in fact have the kind of knowledge which interests him, presumably because they do not know it in the right way, by smelling it. The goal of Wang’s analysis—in particular, his identification of what counts as knowing the odor in the right way—is to defend the idea that the stand-in for action begins no later than the stand-in for knowledge. So, in this third example too Wang holds that the stand-in for action begins no later than the stand-in for knowledge, providing further evidence that he would accept the claim that virtuous action begins no later than the relevant genuine knowledge.¹⁹

¹⁸ My own view is that Wang means the examples only to be analogues for knowledge and action. I do believe that Wang himself thought that psychological or affective responses are or can be actions in their own right (e.g., *IPL* 226, *QJ* 109–110, cf. *QJ* 32.1292–1293, *QJBB* 323; *IPL* 132, *QJ* 46–47; see below, n35). But it seems to me improbable that in the quick remarks here he means to be alluding to this relatively esoteric aspect of his thought. If his goal had been to emphasize the nonobvious idea that the affective responses of loving and hating are actions, we would expect him to develop this idea in the context of filiality and respect as well. But he does not say “a person who knows filiality is acting filially because their knowledge is a form of action”; instead he talks about enacting filiality and enacting respect in a seemingly ordinary way that would be misleading if he had his doctrine about mental action in mind. Note that reading the examples as analogues is compatible with many different ways of understanding “belong to” (*shu yu* 屬於), for instance, if we take it to mean “are” or “are a way of.” A person who is describing an allegorical work of visual art may say “Here, the swan is the family’s prosperity,” to indicate that the swan represents the family’s prosperity. In saying this, the person is clearly not committing themselves to the barely intelligible claim that a swan is prosperity. Similarly, on this way of understanding *shu yu*, we can take Wang to be saying that the loving and hating are actions or ways of acting *given the way he is understanding the example*; this expression on its own cannot be used to resolve the question of whether Wang means to assert here that they are (or are not) literally actions.

¹⁹ On a different reading of the example, the person sees the object, but does not know that the odor exists. On this reading, Wang would also clarify that the person does not have the stand-in for action only because they do not have the stand-in for knowledge, so he would still emphasize that action begins no later than knowledge. Since the upshot is the same, the difference between these readings does not make a difference to the points made in the main text. But there are at least three reasons to favor my preferred reading. First, and most importantly, if Wang’s point is simply that someone who has no knowledge of an odor at all may fail to react to the odor, it is unclear why he would specify that the person sees the object and has a blocked nose. Why not say instead that the person is far away and has no idea about the odor? Second, the alternative reading leaves Wang’s discussion of the example open to an obvious objection. While it is true that a person who does not know that there is an odor at all may fail to exhibit a relevant affective response, clearly a person could know that there is a bad smell without having the relevant affective response (if someone told them there was an odor), and in this sense they would have some form of knowledge of the odor without hating it. If Wang did not distinguish between different ways in which one could know about the odor (e.g., by smell as opposed to by testimony), he would leave it open that in this example too the stand-in for knowledge and the stand-in for action could come apart. Third, the alternative reading disrupts the parallel with XU Ai’s cases, because the person would not have a form of knowledge parallel to the knowledge the people in Xu’s cases had. This person would know the object which produces the odor, but they would not know that they should hate the odor because they do not know that there is an odor. Since it disrupts this parallel, the alternative reading makes the example less relevant to the discussion.

As we have seen, Wang is explicit in these examples that the stand-in for action begins no later than the stand-in for knowledge. He does not say explicitly that the stand-ins for action also begin no earlier than the stand-ins for knowledge, but it is fairly clear that he would say that they do. In all three examples Wang seems to be interested in the form of “love” and “hate” that cannot antedate a person’s appropriate acquaintance with the object of that love or hate. He does not consider the idea, for instance, that people are born with a love of lovely sights in general or a hatred of hateful odors in general, and that this love or hatred (a stand-in for action) precedes their perception of any relevant sights or odors (here a stand-in for knowledge). He also does not consider the idea that, in seeing the source of the odor or learning about the odor on the basis of testimony, the person who knows that it is a bad odor already hates it. Instead he emphasizes how the relevant love or hatred require the right kind of perception of the object. Perhaps even more strikingly, Wang does not say directly that the stand-in for action begins *earlier* than the stand-in for knowledge, but only that it is not true that it begins later. This way of putting the point seems designed to lead the reader to infer that action does not begin earlier, either.

So, to sum up: Wang says explicitly, on more than one occasion, that the stand-in for action begins no later than the stand-in for knowledge. In the examples he uses to illustrate this idea, it is also clear that the stand-in for action begins no earlier than the stand-in for knowledge. If one thing begins both no later and no earlier than something else, then the two must begin at the same time. So here the stand-ins for knowledge and action begin at the same time, suggesting that Wang also holds that genuine knowledge and virtuous action must also begin at the same time. This claim on its own does not give us the full strength of Unity; Unity requires in addition that knowledge and action end at the same time. But Wang’s remarks here do seem to support a broader picture on which he endorses that claim, as well.

So far, it might seem, so good. But Wang’s next remarks (in [h1]) might seem immediately to undercut any antecedent support for attributing Unity to him. Here Wang says that one can describe a person as (genuinely) knowing filiality or respect only if they have in the past acted filially or respectfully. If one parses this sentence in a literal-minded way, Wang would say explicitly that a person must have first acted filially before they acquire the relevant knowledge, flatly contradicting the claim that the relevant knowledge and the relevant action begin at the same moment, and hence also contradicting Unity itself.²⁰

But this literal-minded reading is not forced on us, and I believe it would be a mistake to endorse it. Suppose I say that you can’t appreciate the columns of Zhangjiajie 張家界 until you’ve seen them. The sequence of tense in this statement might lead a robot to conclude that I meant that the appreciation would start after the seeing started. But most people would not draw this conclusion. Instead, they would recognize that my main claim is that it is by seeing the columns that one comes to appreciate them. On this natural reading, I would minimally not be taking a stand on whether the first moment of seeing

²⁰ In light of mathematical discoveries made since Wang’s time, we now know that this would not be a contradiction without further assumptions. For instance, if the temporal extent of knowledge and action were modeled by a single interval in the real numbers that is open on the left, then every moment of knowledge could be preceded by a moment of action, even though every moment of action was also simultaneous with a moment of knowledge. But Wang would not have had such ideas at his disposal, and I think we should see him as holding that actions must have a first moment. Given this further assumption, the literal-minded reading would indeed contradict Unity, even given what we now know.

comes before the first moment of appreciating; and in fact I would naturally be read as committing myself to the claim that they start simultaneously. On this alternative, more human reading, Wang’s focus is not a claim about temporal precedence, but instead a claim about how genuine knowledge is acquired, that is:

Means KA People come to genuinely know filiality only by acting filially.

And this claim is consistent both with Unity and—what is more relevant in the immediate context—with the claim that genuine knowledge and virtuous action begin at the same time.²¹

Which should we prefer, the literal-minded reading, on which action precedes knowledge, or the more human one, on which Wang asserts Means KA? Three arguments point toward the latter. First, as we have seen in the examples which precede this comment, the stand-ins for action and the stand-ins for knowledge start at the same time. Wang does not provide clear examples in which action begins before knowledge, as one might have expected if his goal were to assert the temporal claim the literal-minded reading attributes to him here.

Second, the three examples which follow these comments—of pain, cold, and hunger in [h2–h4]—fit well with a view on which Wang endorses Means KA, but not with one on which he holds that action precedes knowledge. Here, Wang says that one can know pain, cold, and hunger only if one has been in pain, been cold, or been hungry. Although Wang uses temporal language again in these three examples, the examples themselves seem designed to rule out the literal-minded construal of this language: it would be at best an unusual case (if it is even possible) for someone to be in pain first, before knowing pain, or for them to be cold or hungry first, before knowing cold or hunger. By contrast, it is natural to think that being in pain, being cold, and being hungry typically begin at the same time as one’s knowledge of pain, cold, and hunger. These examples seem designed to highlight the claim that the experience of being in pain, being cold, or being hungry are the means by which one acquires the relevant knowledge, and that this knowledge begins simultaneously with the experience. Since, in the context, Wang seems to take being pained, being cold, or being hungry as stand-ins for action, the examples fairly directly make the point that one acquires relevant knowledge by acting (i.e., by having the relevant experience). At the same time, they suggest—contrary to the literal-minded reading—that Wang does not hold that the relevant action begins before the relevant knowledge.²²

²¹ Another way of reconciling this remark with Wang’s earlier claim that knowledge and action begin at the same time would be to hold that, while Wang does make the temporal claim that knowledge comes after action, he is not committed to a universal generalization, but only to a generic claim. What he wants to say is that, typically, we say that people know filiality only if they have acted filially. The idea would be that Wang is simply setting aside “first moments”; if pressed he would agree that his claim does not apply to them.

²² Again, we do not need to take a stand on whether Wang would have endorsed the claim that being pained, being cold, or being hungry are actions *in propria persona*; to illustrate the general idea all he needs is for the reader to come along with him in carving up the example *for present purposes* in a particular way. The broad idea in these examples that genuine knowledge requires something like first-hand experience can also be seen in CHENG Yi’s discussion of the farmer and the tiger, as well as in Wang’s suggestion that some knowledge is as incommunicable as a mute person’s knowledge of the bitterness of bitter melon (IPL 125, QJ 42). Elsewhere Wang speaks of “personally understanding” the doctrine (*zi zhi de* 自知得) on the basis of “personal experience” (*jiu shen xin shang ti lü* 就身心上體履; QJ 6.232, Ching 1972: 106).

Third and finally, Wang's slogan itself provides us with a reason to prefer to read him as asserting Means KA, and not as claiming the action begins before knowledge. If Wang had wanted to claim that action comes before knowledge, why would he have described his doctrine as "the unity of knowledge and action," and not "action first, and knowledge later"? The latter would have made for an equally striking contrast with the "knowledge first, action later" slogan he attributes to ZHU Xi and his followers. The slogan Wang did choose suggests that he did not want to insist on the temporal priority of action, but that he simply used temporal language to make the point that knowledge is acquired by acting.

I conclude that Wang's apparently temporal statement in [h1] is best understood as a claim about the role of action in acquiring knowledge, that is, as Means KA. In endorsing Means KA, Wang seems not only to be saying that action is necessary for the acquisition of relevant knowledge; he seems to be saying also that action suffices for the acquisition—and hence possession—of this knowledge. While his focus in the six illustrative examples in the passage (sight, smell, blocked nose, pain, cold, hunger) is on the claim that knowledge and action *begin* at the same time, the claim that acting suffices for acquiring relevant knowledge precludes the possibility that action could continue after a person has stopped knowing (since the continued action would itself be sufficient for reacquiring the knowledge). So, our discussion so far points to the claim that Wang is committed to:

AK If a person is acting filially, they genuinely know filiality.

3.3 KA

This completes my case for AK. It is now time to turn to the other half of Unity, KA, that is: if a person genuinely knows filiality, they are acting filially.

To introduce my arguments that Wang accepts KA, note first that Means KA on its own does not provide an adequate response to XU Ai's challenge. XU Ai is concerned that people can know that they should be filial, without now being filial. Means KA rules out the possibility that a person could now know filiality without *having been filial in the past*. But it does not say anything about the relationship between a person's current knowledge and their current ethical state: it leaves it open that a person could now know filiality, because they acted filially in the past, even though they are now no longer even disposed so to act.

Wang's response to XU Ai suggests that he holds that there is a further connection between genuine knowledge and virtuous action, which goes beyond Means KA, and which rules out this possibility. (If Wang did not endorse such a further principle, his response would make little sense: he should have simply conceded that Xu's cases were reasonable and possible, and clarified the doctrine of unity of knowledge and action in some other way.) Two natural candidates for such a further claim are:

General KA If a person genuinely knows filiality, they will act filially whenever they are faced with a situation where filial action is appropriate.

KA If a person genuinely knows filiality, they are acting filially.

These two principles rule out the problematic cases in different ways. The first says that the possession of genuine knowledge guarantees that a person will act filially if an

opportunity arises. The second says that the only times when a person has genuine knowledge are the times when the person is in fact acting filially.

I will now argue that Wang is committed to KA. First, I will argue for KA as opposed to General KA as an interpretation of this passage. Second, I will provide independent evidence that Wang endorses KA. And finally, I will respond to an objection to the claim that Wang endorsed KA.

My first argument centers on the claim that KA makes much better sense of Wang's focus on Means KA in our passage than General KA does. Means KA is naturally understood to imply that in the first moment or moments that a person has genuine knowledge of filiality, they must be acting (since action is required for the acquisition of the relevant knowledge). Moreover, it is natural to generalize Means KA to apply to times beyond the first moments of action or knowledge by saying that people not only *acquire* knowledge by acting, but in general that they know filiality—that is, persist in possessing this knowledge—only by acting filially. It is a short step from this generalization of Means KA to KA: since one continues to possess knowledge only by acting filially, if one knows filiality, one is acting filially. So, if the doctrine of the unity of knowledge and action rules out XU Ai's examples because it involves a commitment to KA, it is easy to make sense of Wang's focus in this passage on Means KA. He presents an analysis of the first moments of knowledge and action which, if generalized to later moments as well, would rule out XU Ai's examples.

By contrast, if Wang's goal were to assert General KA, it is unclear why he would focus on this claim about the *acquisition* of genuine knowledge. There is no obvious, natural way of strengthening Means KA so that it would imply General KA. To see this, suppose that you can only know the color red by seeing it. It does not follow that, later on, since you still know the color red, you will be guaranteed to see it whenever it is presented to you. You might have now lost the capacity for sight altogether. And there is also no obvious, natural way of strengthening General KA so that it implies Means KA. Even if genuine knowledge guarantees acting well in the appropriate circumstances, it does not follow that one must have acted well in the past to acquire the psychological dispositions (or whatever it might be) that guarantees acting well in the future: glasses do not have to be broken to become fragile. These logical points do not prove that there is no felt relationship between the principles; they are only meant to sharpen an antecedent sense (which I hope the reader will share) that the exact relationship between these principles is at best not obvious.

This argument is contrastive: it is an argument for KA *as opposed to* General KA. The argument leaves it open that there might be a third principle which makes better sense of the passage. My second argument is a more direct argument in favor of KA, based on independent evidence from Wang's letter to Gu Dongqiao:

[T4] [j] [j1] 吾子謂:「語孝於温清定省,孰不知之?」[j2] 然而能致其知者鮮矣。[j3] 若謂粗知温清定省之儀節,而遂謂之能致其知,則凡知君之當仁者皆可謂之能致其仁之知,知臣之當忠者皆可謂之能致其忠之知,則天下孰非致知者邪?[j4] 以是而言,可以知致知之必在於行,而不行之不可以為致知也明矣。知行合一之體,不益較然矣乎?

[j] [j1] You say: “who does not know to say that filiality consists in warming and cooling [one’s parents bed] and settling them and inquiring after [their health]?”²³ [j2] But those who can extend this knowledge [*zhi qi zhi* 致其知] are few. [j3] If we describe someone who roughly knows the detailed rites for how to warm and cool, to settle and inquire after, and for this reason say that they can extend their knowledge, then it would be admissible to say that anyone who knows that the ruler should be humane can extend his knowledge of humaneness, and admissible to say that anyone who knows that the subject should be conscientious is able to extend his knowledge of conscientiousness. Who in the world would not extend his knowledge? [j4] If we consider the matter from this perspective, we know that extending knowledge must consist in acting, and it is clear also that if a person is not acting then they cannot be regarded as extending their knowledge. Is the natural condition of the unity of knowledge and knowledge and action not still more evident now? (*IPL* 139, *QJ* 56)

In [j3], Wang speaks of the “natural condition of the unity of knowledge and action.” This is not the same as saying “the unity of the natural condition of knowledge and action.” But the remark suggests that Wang has the doctrine about the original natural condition of knowledge and action in view, as opposed to his doctrine about training.

The phrase “extension of knowledge” (*zhi zhi* 致知) appears in the *Great Learning*—a canonical text for those working in Wang’s philosophical tradition—on a list of eight stages in or aspects of the development of an ethical state. Wang has a distinctive interpretation of a number of items on this list, including this one. In fact, the standard translation “extension of knowledge” is based on ZHU Xi’s understanding of this idea; a better rendering of the phrase given Wang’s interpretation might be the “perfection of knowledge.” But this aspect of Wang’s thought will not be important for us here. What will be important is that in many passages, Wang closely associates this extension or perfection of knowledge with the unity of knowledge and action. In these passages it is plausible that he uses the expression to describe the very same elevated form of knowledge which he elsewhere describes as the “original natural condition of knowledge” or “genuine knowledge.”²⁴ Here, too, Wang seems to use “extended knowledge” (*zhi zhi* 致知) in [j2], [j3], and [j4] to describe this elevated form of knowledge.

If one accepts my translation, in [j4], Wang says explicitly that extending knowledge consists in acting, and that if a person is not acting they cannot be regarded as extending their knowledge. On this translation, and given my claim that “extended knowledge” denotes the same elevated form of knowledge as “genuine knowledge,” Wang asserts KA (or, more properly, but equivalently, the contrapositive of KA): he says that if a person is not acting, they do not have extended knowledge.

²³ The passage says only “warming and cooling, settling and inquiring,” but it clearly alludes to a famous passage in the *Book of Rites*, “Summary of the Rules of Propriety (Qū Li Shang 曲禮上),” which traditional commentators (whose works Wang would have known) understood in this way.

²⁴ Perhaps most notably, see *QJ* 5.211, Ching 1972: 68–69 (cf. *QJ* 27.1100 where the same point is made almost verbatim, and also *IPL* 139, *QJ* 56; *IPL* 140, *QJ* 58; *IPL* 321, *QJ* 137; *QJ* 8.308). Note that, while I think it is natural to take instances of “extended knowledge” to denote genuine knowledge when Wang explicitly associates this term with the unity of knowledge and action, I do not think that every use of “extended knowledge” denotes this elevated form of knowledge.

But the translation on which this argument relies will be controversial. Someone might reject that translation and instead render the first sentence as “extending knowledge depends upon acting,” and the second as “if a person has not acted, they cannot be regarded as having extended their knowledge.” This alternative translation requires an alternative conception of extended knowledge than the one I have suggested. Perhaps the most natural option would be a view on which extended knowledge is a stable, long-lived state or capacity. But Wang’s remarks in [j3] reveal that in this passage he does not think of extended knowledge in this way. In [j3] Wang repeatedly contrasts someone who merely knows how to say the right words about virtue with the state of a person who “*can* extend his knowledge” of humaneness or conscientiousness. In these remarks, the relevant stable state is not one of having extended one’s knowledge but rather of *being able* to extend it. This discussion suggests that while a virtuous person is able to extend their knowledge, they will not do so all the time, but only when the occasion calls for it. In this context, the translation I’ve given for [j4], according to which the occasion for a person’s extending their knowledge is the occasion of acting virtuously, is far more natural than the alternative. And, as I have said, on this translation, and given my earlier claim that “extended knowledge” here denotes the same elevated form of knowledge which Wang elsewhere calls “genuine knowledge,” Wang endorses KA.²⁵

So the hypothesis that Wang endorses KA not only makes better sense of his emphasis on Means KA in [T3] than the hypothesis that he endorses General KA, it also receives independent support from [T4]. The case that Wang endorses this principle is strong. But there is an important objection which might seem to undermine this case. The objection is based on the fact that in an absolutely crucial passage ([T3] [h1]) Wang says that a person knows filiality only after they have acted filially, and that they know respect only after they have acted respectfully. These claims (along with related ones made in [h2]–[h4]) seem designed to allow the reader to infer that there is a form of knowledge which persists *after* the initial action, and thus that there are times when a person knows but does not act, contradicting KA.

This is an extremely important objection. If I did not believe the interpretation I will advance, the reason would be because of it. But I think there is a reasonable way of responding to the objection. The response is based on the suggestion that there are two ways of speaking about knowledge relevant to this passage. Compare a person who is now experiencing hunger with a person who is now recalling what it feels like to be hungry (or who is merely capable of recalling this sensation). There is a reasonable way of using “know hunger”—which I will call the “expansive sense”—on which it applies to both of these people. But there is also a reasonable way of using this expression—which I will call the “restrictive sense”—on which only the person who is currently hungry counts as knowing hunger. (Some may find the example of “know poverty” more compelling: a person who is no longer poor *has* known poverty in the relevant sense, but they do not know it any longer.) In my view, when Wang says that a person knows filiality only after they have acted filially, inviting the reader to infer that knowledge persists after action, he

²⁵ In the face of this second argument alone, one might question my claim that “extended knowledge” can denote the same elevated form of knowledge as “genuine knowledge,” claiming that Wang uses genuine knowledge in the way described in General KA, and extended knowledge in the way described in KA. But this position would still face the challenge raised in my first argument for KA as opposed to General KA, since that argument applied to General KA *as a description of genuine knowledge*.

should be understood as using “know” in the expansive sense: people continue to know filiality in the sense of being able to recall the sensation of being filial, after they have acted filially. But he can also use this expression in the restrictive sense, and in this sense, which is the one relevant to KA, knowledge of filiality does not persist after filial action.

While this interpretation allows us to respond to the objection, I want to be clear that it does have significant costs. It requires postulating a distinction in Wang’s usage that Wang himself does not make explicit. Moreover, it requires claiming that Wang is not talking about the most important form of knowledge for the unity of knowledge and action in [h1], arguably *the key* remarks on how genuine knowledge relates to virtuous action. But I believe that these costs must be paid. Every reasonable alternative interpretation I have been able to think of fails to satisfy other, more important demands which I have emphasized already. For example, while General KA can make sense of the idea that a person continues to have knowledge in Wang’s most favored sense after having acted filially, we have seen that Wang’s comments in the passage simply do not explain how a person’s first experience of filial action would be required for (or, even more oddly, sufficient for) the acquisition of something like a disposition to act filially in appropriate circumstances. As I have said, it is not by breaking that glasses become fragile. Moreover—and equally importantly—in [T4], Wang endorses an analogue of KA for “extended knowledge,” which I have suggested is the same form of elevated knowledge he elsewhere calls “genuine knowledge.” Any view which rejects KA in [T3] will struggle to explain Wang’s endorsement of an analogue of KA for this closely related notion. So, while I acknowledge that my interpretation of Wang’s comments in [h1] has significant costs, it is still the best overall interpretation of his remarks that I have been able to find. Since no other interpretation seems to me to do better, I believe we should attribute KA to Wang.

I have argued that, in [T3], Wang is best understood as committed to AK and KA. Since the conjunction of these principles is equivalent to Unity, if he is committed to them, he is also committed to Unity. As presaged in the introduction, in arguing for the claim that Wang is committed to Unity I have largely treated “knowledge” and “action” as placeholders. I have not taken a stand on many controversial questions about how Wang substantively understands these notions. But the conclusion that Wang is committed to Unity is a striking result, which tightly constrains how one might develop a fuller interpretation of Wang’s views about knowledge and action. To see this point, note that on the usual way of thinking about (propositional) knowledge, such knowledge is a long-lasting state. If Wei knows that the capital of Tang 唐 was Chang’an 長安, he still knows this even when he is not thinking about this question, and in fact even if he is asleep or unconscious. But clearly if Wei is unconscious and not moving, he is not acting in any familiar sense. So, given KA, he cannot have genuine knowledge. Genuine knowledge thus must be quite different from propositional knowledge understood in this usual way, if Wang accepts Unity. Unity imposes substantive constraints on the interpretation of Wang’s views about (genuine) knowledge and action.

The arguments I have given for this substantive constraint do not employ premises about the broader interpretation of Wang’s views on knowledge and action. They are based on claims of detail about particular passages. These arguments can be accepted by those who endorse a wide array of “big picture” views about the aims of Wang’s doctrines. By the same token, however, those who wish to challenge the arguments cannot do so only by offering a big picture which conflicts with Unity; they must engage as well with the detailed arguments based on these particular texts.

4 Identity

The English word “unity” might naturally suggest to some readers that Wang must have in mind not only the idea that the original natural condition of knowledge occurs if and only if the original natural condition of action does (that is, the thesis I have dubbed “Unity”), but a much stronger claim as well: that knowledge and action are in some sense the very same thing. We can make this idea more precise, using the key terms which also appear in Unity, as follows:

Identity To genuinely know filiality just is to act filially.²⁶

Here I use “just is” to denote a symmetric relation, so that Identity is equivalent to the claim that to act filially just is to genuinely know filiality. Unity says that a person genuinely knows filiality if and only if they act filially. Identity says that genuinely knowing filiality and acting filially are the very same thing. If genuinely knowing filiality is the very same thing as acting filially, then a person will genuinely know filiality if and only if they are acting filially; Identity entails Unity. But Unity does not entail Identity: it does not follow from the fact that one thing occurs if and only if another does, that the two are identical.

Proponents of Identity should see my arguments that Wang endorses Unity as for the most part friendly: Unity gets us part of the way to Identity. But they will want to argue that Unity is not enough, and that the full extent of the relationship Wang postulates between the original natural conditions of knowledge and action can only be captured by Identity. In this section I will argue against this view. I will first present two direct arguments against the claim that Wang endorses Identity, and then I will consider two ways one might attempt to argue *for* Identity, and argue that neither is successful.²⁷

²⁶ Huang 2017: 75 is one of the most explicit endorsements I know of.

²⁷ My discussion of Identity is inspired in part by Warren Frisina’s discussion of the principle, primarily in Frisina 1989 (which is essentially reprinted in Frisina 2002). I will quote Frisina below in support of my suggestions for how an interpretation which attributes this principle to Wang might look. But I want to emphasize from the outset that it is not clear that Frisina is himself a wholehearted proponent of Identity, or even the spirit behind that principle. Frisina introduces the project of his 1989 paper by saying: “My aim is to demonstrate that Wang intended his phrase literally. Knowledge and action are, *in all their forms, really one thing*” (Frisina 1989: 419, emphasis his). In the preface of Frisina 2002, he writes: “... I was inspired by the philosophy underlying the Neo-Confucian scholar WANG Yang-ming’s famous slogan *chih hsing ho-i* (the unity of knowledge and action). Though it sounds strange at first I have come to believe that he meant us to take this slogan literally. For Wang, knowledge is a way of acting. To say that we know something is really a statement about how we interact with it...” (Frisina 2002: 4). But, in spite of these remarks, I am unsure whether Frisina endorses Identity. When he spells out his thesis in more detail, he does not stick to the letter of the claim that knowledge and action are, “in all their forms, one thing.” At the end of the previous quotation, for instance, Frisina seems to say only that knowledge is constituted not by a single action but by a pattern of actions, precluding the claim that any particular action could be identical to knowledge of an object. Frisina also makes this point in the paper, writing: “Knowledge is not a ‘representation the world,’ but a pattern of behavior, a way of being in the world. To know something is to react to it in a way that takes it into account” (Frisina 1989: 420). Here knowledge of objects is identified with “a pattern of behavior,” not with any particular action: not all actions are identical to knowledge. So although I will describe Frisina below as attributing this principle to Wang, I do not think it is obvious that he does believe that Wang endorses the principle.

My first argument against attributing Identity to Wang is a conceptual one. If, as Identity says, to genuinely know filiality is to act filially, then presumably if a person acts filially, their acting filially must be identical to their genuinely knowing filiality. But this last claim is implausible. The sage-king Wu famously raised an army without mourning his parents, an action that was considered filial in spite of its apparent unfiliality (for Wang's discussion of the example see *IPL* 139, *QJ* 57). Was this filial action identical to Wu's genuinely knowing filiality? It is just incredible to say that it was. Wu's action of raising the army consisted of many component actions: of sending messengers to rouse the troops, of riding from place to place with his generals, or of marching with the army. These actions are not identical to anything understood as knowing in any ordinary sense of that term. One can understand a conception of action on which his actions had various mental states or activities as *parts* (and also on which the mental events themselves were actions, see n35), but surely his actions involved a great deal more than this: he had to move his mouth, leap onto his horse, and move his legs as he marched.

Arguments based on the intrinsic plausibility of a given position cannot be decisive in determining what WANG Yangming believed; it is possible that Wang endorsed implausible positions. But the obviousness of this style of counterexample does place the burden of proof squarely on the shoulders of proponents of Identity. The textual evidence would have to be strongly in favor of Identity if we were to attribute it to Wang, given how evident the counterexamples to it are. If Wang did endorse this radical position, one would at the very least expect him to have described why he thought such counterexamples are not in fact counterexamples, as a way of explaining the position he did hold. But we find no such explanation in the surviving texts.

Someone might seek to resist this argument by claiming that it misrepresents Wang's aims as a philosopher. They might say that Wang's goal in endorsing Identity is not to offer a theory of the nature of knowledge and action, but instead to reform the concepts we use to think about our ethical lives. On this picture, the fact that Wang does not consider the kind of example I have just described could be seen as evidence in favor of the claim that Wang does not aim to describe knowledge and action. Such examples, based on the commonsense notions of knowledge and action, are irrelevant to Wang's project, which precisely involves reforming those commonsense concepts.

But, first, while this response may escape the letter of the argument, it does not escape the spirit of it. Even if the response were to succeed in showing that such examples are not counterexamples to Wang's doctrine properly understood, it would not succeed in showing that the examples are irrelevant to Wang's project. Given the obviousness of examples like the one above, where King Wu's actions are not knowledge in any sense, Wang would still owe his students an explanation of how, on his proposed conceptual reform, this kind of example is no longer concerning—how exactly are knowledge and action now to be understood so that these examples either do not count as actions or do count as knowledge? In my view, this first point is sufficient to show that the response fails. But the response also has a second problem, since the methodological remarks I discussed in detail in Section 2 undermine the idea behind it. As we saw, Wang there contrasts the fact that his doctrine was developed as a therapeutic measure, with the fact that his doctrine concerns the original natural condition of knowledge and action. By far the most natural reading of those remarks is that Wang endorses a straightforward

distinction between the primarily practical aspects of his doctrine, and the aspects of it which are concerned with describing reality. If Wang’s doctrine about the original natural condition of knowledge and action is meant as a primarily practical one, a conceptual reform which is intended to remedy people’s moral defects by a radical form of mental therapy, the contrast Wang draws in those passages would make no sense.

A second argument against attributing Identity to Wang is textual. It is based on the text of paragraph [d] in [T2], which I elided above, but will now quote:

[T5] [d] 先生曰：「此卻失了古人宗旨也。[d1] 某嘗說知是行的主意，行是知的功夫；[d2] 知是行之始，行是知之成。[d3] 若會得時，只說一個知己自有行在，只說一個行己自有知在。……」

[d] The Master said: “But this is to lose [sight of] the purpose [*zong zhi* 宗旨] of the ancients. [d1] I have said that knowledge is the main goal (*zhu yi* 主意) of action, and that action is the training (or: “effort” *gong fu* 功夫) of knowledge, [d2] that knowledge is the beginning of action, and action is the completion of knowledge. [d3] If you understand correctly, then as soon as you speak of knowledge, action will automatically already be included, and as soon as you speak of action, knowledge will automatically already be included....” (*IPL 5, QJ 5*; cf. *IPL 26, QJ 15*)²⁸

In context, it is clear that Wang’s remarks about knowledge being the beginning and action coming later concern the original natural condition of knowledge and action. These remarks come immediately after XU Ai’s question in [T2] [c], which, as I argued at the end of Section 2, divides the doctrine of the unity of knowledge and action fairly explicitly into one part concerning the original natural condition and another concerning training. In response to this question, Wang starts by saying that one must be clear about the ancients’ purpose. But before clarifying their purpose (in the passage immediately following this), he pauses parenthetically to restate his own view about the original natural condition of knowledge and action.²⁹

So in the two key sentences, Wang ascribes properties to knowledge in its original natural condition, which action in its original natural condition lacks (being the beginning, being the guiding aim), and ascribes properties to action in its original

²⁸ There are three ways of taking the beginning of the passage ([d1]), depending on how we understand *yi* 意 there. On a first interpretation (represented in my translation), *yi* is taken as “goal” or “aim” (Wang uses the word in this way twice in [T2] [b3], which comes just before our passage). His idea is that one’s action aims at knowledge, and acting is the process of fulfilling that aim, so that the best forms of knowledge and action are jointly realized. On a second interpretation, *yi* means “inclination” and Wang speaks here of one or more mental events that last through the action and guide it. (So, L. Chen 1991: 101, and also Angle and Tiwald 2017: 131, who translate this “intent of acting.”) On a third interpretation, my “main goal” (*zhu yi*) is instead understood as “master” (*zhu zai* 主宰), which is in turn associated with *liangzhi* 良知 (Z. Wu 2018: 21). The difference between these interpretations will not matter much in what follows. My main reason for preferring the first here is that Wang elsewhere also draws a contrast between the “main goal” *zhu yi* and “training” *gong fu* 功夫/工夫 in a context where it clearly has nothing to do with a particular person’s psychological state, and where it is also clear that *zhu yi* is not to be taken as “master” (*IPL 25, QJ 15*; *IPL 168, QJ 80–81*).

²⁹ The best construal of the related remarks in *IPL 26, QJ 15* is similar: the first remarks there concern original natural condition, while the later ones concern training.

natural condition (being the completion, being the training) which knowledge in its original natural condition lacks. But if two things have different properties, they are not identical. So, knowledge and action are not identical.³⁰

It might seem, however, that the passage threatens more than just Identity. Wang's remarks about "beginning" and "completion" on their own could be read as describing a temporal relationship (knowledge comes first, and action later), or a conceptual one (knowledge is in some sense conceptually prior; action is in some sense conceptually posterior). If the temporal reading were the correct one, this passage would be incompatible with Unity as well as Identity.³¹ But in context, the latter, conceptual reading is clearly to be preferred. First, Wang immediately goes on to say that, whenever one speaks of knowledge, action is present, and whenever one speaks of action, knowledge is present. These remarks conflict with a temporal interpretation of Wang's earlier comment: if knowledge came first, and action later, then one could imagine the knowledge pre-existing and being stopped prior to the action, so that there would be circumstances in which one could speak of knowledge without speaking of action. Second, Wang elsewhere presents his main opponent as someone who holds that knowledge comes first and action later (*IPL* 137, *QJ* 54; *IPL* 140, *QJ* 58; *QJ* 8.309; *QJ* 32.1331). It would be odd if he then asserted here that in fact knowledge comes first and action later. As CHEN Lai 陳來 points out (though he defends the temporal reading), we would be left to wonder what was distinctive about Wang's own view (L. Chen 1991: 101). Third, and most powerfully, in every single example Wang considers in [T3] (which comes earlier in this same passage)—the examples of the odor, sight, filiality, respect, pain, cold, and hunger—he emphasizes explicitly that action begins *no later* than knowledge. It would be bizarre if, in closing this passage, he undid all of his earlier work, and affirmed that knowledge does, after all, come before action. In short, there is strong reason to understand his comments as indicating conceptual, not temporal priority, and on this reading the comments are consistent with Unity.³²

³⁰ For this argument, see also L. Chen 1991: 99; Lee 1994: 423. Frisina recognizes the problem but does not offer a reading of the passages that fits with his interpretation (Frisina 1989: 421). There are, however, some substantive responses in the literature. Yong HUANG suggests an ingenious reinterpretation of [d2], on which it should be read as "[the beginning of] knowledge is the beginning of action; [the completion of action] is the completion of knowledge" (Huang 2017: 73). On this view, [d2] would be consistent with Identity. But the reading requires what is in my view a fairly extreme interpolation, and in any case, it does not explain away the remark in [d1], where Wang also attributes further different properties to knowledge and action (being the guiding aim vs. being training). Wu Zhen offers a different kind of diagnosis, suggesting that in his later years, Wang would no longer have accepted this (earlier) remark, and that in those later years he did endorse something like Identity (Z. Wu 2011: 103). This position deserves further consideration, but as indicated in n. 4 I am here exploring how we should understand Wang's views on the assumption they were consistent after 1509.

³¹ For the temporal reading see, for example, Lao 1984–86/2019: 3.422, who claims that all Wang intends is to defend the claim that the source of knowledge and action are the same (which he calls "the source meaning," *genyuan yiyi* 根源意義). For criticism, see L.-S. Chen 2015: 9, and now L.-S. Chen 2019: 126–132.

³² One might attempt to save the temporal reading of "beginning" and "completion" by elucidating a sense of conceptual containment which does not require temporal coincidence. One could imagine a philosopher saying that speaking of a seed involves speaking of a tree, even though the tree does not exist yet. Similarly, one might think that Wang could be saying that speaking of knowledge would in some sense involve speaking of action, even though the action does not exist yet. But this response, while ingenious, still falls foul of the second and third argument in the main text. CHEN Lisheng 陳立勝 develops a version of this view, on which the sense of knowledge relevant here is a capacity which develops into action (L.-S. Chen 2015: 7, 9; now L.-S. Chen 2019: 126–132). It is unclear to me whether Chen ultimately sees this view as supporting Identity, or whether he in fact accepts that even Unity should be rejected (since he holds that time-differences between knowledge and action are admissible provided they are "homogeneous").

In any case, if this passage is inconsistent with Unity, then it is *a fortiori* inconsistent with Identity. So even if one does not accept the foregoing arguments that the passage should be read in a way that is consistent with Unity, one should still agree with me that the passage provides evidence against Identity. In fact, I think the passage is about as clear as they come in the surviving works of WANG Yangming. Together with the first conceptual argument against attributing Identity to Wang, it presents a powerful case that Wang did not endorse Identity.

These arguments are strong, but they might not be conclusive. If there were strong countervailing evidence in favor of the claim that Wang does endorse Identity, one might be led by that evidence on balance still to attribute Identity to Wang, and to find some way of explaining away the passages I have just been discussing. In the remainder of the section, then, I want to consider what seem to me the two strongest arguments that Wang endorses Identity. I will argue that neither of these arguments succeeds.

The first argument is perhaps the more prominent of the two, but it is also the weaker. It is based on the claim that the *literal* reading of “unity” in the slogan “the unity of knowledge and action” is as “identity.”³³ I agree that the expression *heyi* 合一 can mean “identity.” But the expression has a much broader semantic range; it can also something more like “correspondence,” “correlation,” or “co-occurrence.” So on its own the fact that Wang uses *heyi* is not evidence that Wang endorses Identity in addition to Unity. Both Unity and Identity accord with literal readings of the slogan (cf. L. Chen 1991: 96 for a similar point).

A second argument is more challenging. It is based on a family of passages in which Wang repeats a set phrase, which we saw above in [T1]. As I will discuss, there are two natural interpretations of this phrase, one of which directly supports Identity, and the other of which does not. In the remainder of the section, I will consider this phrase in detail and argue that the evidence favors the second interpretation, which does not support Identity. My argument for this conclusion will be a little involved; those who are content with the arguments I have already given against Identity may wish to skip to the next section at this point.

Below I print this difficult set phrase, with two different translations. The (A) translation corresponds to an interpretation which supports Identity; the (B) translation corresponds to my preferred alternative (printed in my translation of the phrase in [T1] above):

[T6] 知之真切篤實處，即是行；行之明覺精察處，即是知。

(A) Insofar as knowledge is genuine, practical, earnest, and effective, it is action; insofar as action is lucid, perceptive, focused, and discriminating, it is knowledge.

(B) The genuine, earnest, practical, and effective aspects of knowledge are [a matter of] action; the lucid, perceptive, focused, discriminating aspects of action are [a matter of] knowledge.

³³ Frisina’s repeated insistence that his attribution of something like Identity to Wang derives from a “literal reading” of WANG Yangming seems an instance of this argument. “My aim is to demonstrate that Wang intended his phrase literally” (Frisina 1989: 419; cf. Frisina 1989: 420, 421, 424, 433, 442, all using the word “literal”).

The main difference between the (A) translation and the (B) translation concerns how the expression X *zhi* Y *chu* X 之 Y 處 is rendered (“insofar as X is Y” vs. “the Y aspects of X”). As a consequence of this difference, the two translations differ also in how they take the phrase *ji shi* 即是 “just is.” The (A) translation renders it literally as “is” or “just is,” while the (B) translation takes it a little more loosely as “is a matter of.” This less literal use of *ji shi* has parallels in English. One can well imagine a baseball coach instructing a player while watching an expert pitcher. “Do you see the spin on that curveball? It’s his ring finger.” In the (B) translation, as in this English sentence, the idea would be that Wang is indicating the cause or explanation of the relevant qualities.

Given our earlier discussion it is natural to think that for knowledge to be in its original natural condition just is for it to be genuine, practical, earnest, and effective (recall: “genuine knowledge”), and for action to be in its original natural condition of action just is for it to be lucid, perceptive, focused, and discriminating. On this assumption, the (A) interpretation directly supports Identity. By contrast, the (B) translation does not support Identity. Wang’s idea would be something we might more fully express as “knowledge’s being genuine, practical, earnest, and substantial comes from its association with action; action’s being lucidly aware and precisely discriminating comes from its association with knowledge.”

These translations each seem about equally plausible to me linguistically. Some might think that, owing to the fact that it is more literal, the (A) translation has a default presumption in its favor. But that is not obvious to me: Wang often uses emphatic language nonliterally even in theoretical contexts. We must look to the context of the set phrase for help. And the context for three of the four passages where Wang uses the phrase supports (as I will now argue) the (B) translation rather than the (A) translation. (The fourth, from a letter to ZHOU Daotong 周道通, *QJ* 32.1331, does not point either way.)

Let us begin with [T1], where Wang uses the phrase in the opening of his response to Gu Dongqiao’s challenge ([a1]). First, immediately after using the set phrase, in [a2], Wang says “The training of knowledge and action at its basis cannot be separated.” This suggests that he sees the point he has just made, in the set phrase, as closely related to the fact that the training of knowledge and action is one and the same. The (B) reading of the set phrase makes better sense of this connection: if Wang is saying that achieving or maintaining the ideal qualities of knowledge (respectively, action) is a matter of action (respectively, knowledge) then he is essentially saying that the training for one of these must be training for the other. Second, and much more strikingly, a few sentences later (in [a4]), Wang approvingly quotes Gu (who seems to be quoting Wang himself) as saying “Genuine knowledge is just what is employed in performing an action; if a person does not act it is not worthy to be called knowledge” (真知即所以爲行, 不行不足謂之知). This statement strongly suggests that Wang believes genuine knowledge is distinct from action: “what is employed in performing action” cannot be action itself. Moreover, the quotation supports precisely the picture described by the (B) translation: Wang says that the fact that the knowledge is genuine is a matter of the action the person takes when they have this knowledge. The (A) translation, by contrast, does not fit well with the passage as a whole. If we accept that

translation, Wang starts by asserting an identity, but then takes it back moments later when he says that genuine knowledge is what is employed in performing the action.³⁴

The second and third passages where Wang uses the set phrase come from the same 1526 letter, "In reply to inquiries from a friend." The first use of the phrase in the letter (*QJ* 6. 232; Ching 1972: 106) comes at the end of a discussion of learning (*xue* 學). After expanding on the set phrase in a few sentences, Wang concludes his discussion "originally it is one training" (元來只是一箇工夫). So here too, Wang uses the set phrase to develop a point about how these ideal conditions of knowledge and action are acquired or maintained: his focus is on the unity of the *training* for knowledge and action, not on whether the state of having the ideal conditions is the very same thing. Once again, since the (B) translation fits better with an emphasis on training than the (A) translation does, we should prefer it here.

The second use in the letter bears out the same idea:

[T7] 知之真切篤實處，便是行；行之明覺精察處，便是知。若知時，其心不能真切篤實，則其知便不能明覺精察；不是知之時只要明覺精察，更不要真切篤實也。行之時，其心不能明覺精察，則其行便不能真切篤實；不是行之時只要真切篤實，更不要明覺精察也。

The genuine, practical, earnest, and effective aspects of knowledge are [a matter of] action; the lucid, perceptive, focused, and discriminating aspects of action are [a matter of] knowledge. If when you are [engaged in] knowing, your mind is unable to be genuine, practical, earnest, and effective, then your knowledge will not be able to be lucid, perceptive, focused, and discriminating; it is not that when you are [engaged in] knowing you only need to be lucid, perceptive, focused, and discriminating, but don't also need to be genuine, practical, earnest, and effective. If when you are acting, your mind is unable to be lucid, perceptive, focused, and discriminating, then your action will not be able to be genuine, practical, earnest, and effective; it's not that when you are acting, you only need to be genuine, practical, earnest, and effective, but don't also need to be lucid, perceptive, focused, and discriminating. (*QJ* 6.234; Ching 1972: 108)

The *Zhong Yong* 中庸 (*Doctrine of the Mean*) speaks of earnest *action* (*du xing* 篤行) (not earnest knowledge); it was also common to speak of lucid or discriminating *knowledge* (not luminous or discriminating action) (cf., e.g., Li and Wang

³⁴ This argument is not quite as conclusive as I would like. If Wang had written 真知即所以行, this would uncontroversially mean that genuine knowledge is used in guiding action; it is what is used to act. But what he did write, 真知即所以為行, has three different possible construals. If we read the character 為 as *wèi* here, it means (i) that genuine knowledge is used to promote action. If we read it as *wéi*, it most likely means something like (ii) that genuine knowledge is employed in acting. But it is perhaps *just* possible that the expression linguistically could mean (iii) that genuine knowledge constitutes action. Since this interpretation is less natural than the alternatives, it would still be a mark against attributing Identity to Wang that it requires adopting it.

1986: 14.281). A key aspect of Wang's set phrase is that it reverses these typical attributes of ideal knowledge and action. In the present passage Wang offers a kind of explanation for why he has performed this reversal: he holds that knowledge cannot have its typically valedictory properties unless it is associated with an action which has *its* typical valedictory properties. But if this is right, then Wang's overall point seems to be a familiar one: that the training of knowledge and action cannot be separated, because the ideal form of knowledge can only be acquired and maintained through action, and the ideal form of action can only be achieved through knowledge. While this point is perhaps consistent with the (A) translation, it is better captured by the (B) translation. And there is a further point in favor of the (B) reading here as well, a point which concerns what Wang does not say as opposed to what he does. In the closing sentence of the excerpt, Wang does not say "if when you are acting, your mind is unable to be lucid ... then *since the lucidity of your mind is the lucidity of your action*, your action will not be able to be genuine...." Without this kind of comment, his remark is much more naturally read as indicating that a person must have one set of qualities if they are to have the other, not that having one set of them *just is* having the other. So on balance, this passage too, it seems to me, favors the (B) translation above.

To sum up: the contexts where Wang uses the set phrase favor the (B) translation and interpretation. In those passages, Wang is concerned to argue that one cannot achieve the ideal properties of knowledge without achieving the ideal qualities of action (and vice versa). He does not seem concerned to argue that knowledge is identical to action, by any means—in one of the passages he even makes a remark which is flatly inconsistent with their identity. So these passages do not support attributing Identity to Wang. They certainly do not provide the kind of conclusive evidence in its favor that would be required to overcome the strong presumption against Identity generated by my first two arguments.

5 Conclusion

The doctrine of the unity of knowledge and action has two components: one concerning the training of knowledge and action, and one concerning their original natural condition (Section 2). Wang says that the training of knowledge and action are unified in the sense that they are one and the same. But he does not hold that knowledge and action are identical in their original natural condition (Section 4). Instead, using "genuine knowledge" and various forms of virtuous action for knowledge and action in their original natural conditions, I have argued that he endorses claims of the form (Section 3):

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

I have not argued that Unity is the only claim that Wang accepts as part of his doctrine about the original natural condition of knowledge and action, and indeed I am open to the idea that Wang endorses other claims under this heading, in addition to Unity.³⁵ But I reject the claim that Wang endorses Identity, and I believe, moreover, that Unity is a core part of Wang’s distinctive views about the relationship between knowledge and virtuous action.

In key remarks about the unity of knowledge and action, Wang makes some apparently radical methodological claims. He says that his doctrine was proposed as

³⁵ There are several further claims which I think are especially worth exploring:

Knowledge is Action Every episode of knowledge is an action.

Knowledge is a Part of Action All episodes of genuine knowledge of filiality are part of a filial action.

Action has Knowledge as a Part All filial action has episodes of genuine knowledge of filiality as a part.

The most promising lines of thought in support of attributing these claims to Wang rest on two key ideas. First, in two passages (*IPL* 132, *QJ* 47; *IPL* 226, *QJ* 109–110, cf. *QJ* 32.1292–1293, and *QJBB* 323 with Shu 2017: 2090 and Wu 2018: 16), Wang seems to say that inclinations (*yi* 意) or concerns (*nian* 念) are actions or are parts of actions, suggesting (i) that all inclinations and all concerns are actions. Second, in a family of other passages (*IPL* 6, *QJ* 6; *IPL* 78, *QJ* 27; *IPL* 137, *QJ* 53; *IPL* 174, *QJ* 86–87; *IPL* 201, *QJ* 103), Wang describes knowledge as intimately related to inclinations. In two, he says that “insofar as an inclination is lively and lucid, it is knowledge” (*zhi yi zhi ling ming chu wei zhi zhi* 指意之靈明處謂之知; *IPL* 201, *QJ* 103; cf. also *IPL* 174, *QJ* 86–87), suggesting (ii) that all episodes of knowledge are inclinations. These two claims immediately imply Knowledge is Action. (Indeed, the set phrase discussed in the previous section may provide further support for this claim, since the expression “lively and lucid” which Wang says qualifies those inclinations that are knowledge also appears qualifying action in that phrase.)

For Knowledge is a Part of Action: I argued above that even if Wang holds that some mental events are actions, it is plausible that he does not think that purely mental (nonbodily) actions always suffice for virtuous action: some virtuous actions require an associated physical performance. Given this, episodes of genuine knowledge of filiality are not in general identical with *filial* actions (although they might be identical with actions). But while they may not be identical with actions, they could be guaranteed to be parts of such actions.

For Action has Knowledge as a Part: given Unity, filial action requires that the person exhibit genuine knowledge of filiality. Wang’s idea that inclinations are parts of actions shows that he happily includes mental events associated with an action as part of the action itself. And this general thought naturally leads to the idea that the genuine knowledge that accompanies filiality is a part of the filial action.

So there is some case to be made in support of attributing these principles to Wang. But the case is far from watertight. The passages in which Wang describes the metaphysics of knowledge (most notably *IPL* 174, *QJ* 86–87; *IPL* 201, *QJ* 103) are not explicitly connected to Wang’s doctrines about the unity of knowledge and action. Two of the clearest passages in which Wang associates mental events with action (*IPL* 132, *QJ* 47; *IPL* 226, *QJ* 109–110) are connected to the unity of knowledge and action, but they are explicitly associated with the training part of his doctrine, not with the original natural condition of knowledge and action. (In the first [*IPL* 132] Wang explicitly addresses training; in the second [*IPL* 226] Wang connects his point about concerns (*nian* 念) to the “guiding aim” or “purpose” (*zong zhi* 宗旨) of the unity of knowledge and action, again suggesting his focus is on the practical dimension of the doctrine.) So even if Wang does endorse these additional principles, the evidence favors the idea that he would not have taken them to be part of his doctrine about the original natural condition of knowledge and action. Second, there is no similar line of argument that Wang endorses the claim that all actions are inclinations or concerns, and also no case that Wang holds that all inclinations or concerns are episodes of knowledge. So once again, even if Wang endorses the claims above, he would not be committed to the claim that all actions are episodes of knowledge (i.e., the position would be consistent with rejecting Identity). This claim is important, since some who have made much of Wang’s interest in mental action have suggested that, if Wang did think that an affect-like mental event was identical with a knowledge-like event, this would suffice for him to hold that knowledge and action are the very same thing (most notably, Huang 2017: 76; Huang 2020). But it wouldn’t, unless Wang also held (counterintuitively and without explicit argument) that the *only* actions are mental actions. Third, the evidence in favor of (ii) needs closer examination. It is clear that Wang thinks that every episode of knowledge is somehow associated with an inclination, but it is not clear that he thinks they are identical. The strongest point in favor of (ii) seems to me in *QJ* 32.1292–1293, where Wang says that when a concern (*nian*) arises, it is knowledge, not that all knowledge is like this. I hope to consider this kind of evidence in more detail elsewhere.

a remedy for the ethical sickness of his day, and he says that if a student does not understand the practical implications of his position (or, “its purpose”), it does not matter whether they assert his doctrine or its negation ([T2] [b3]–[b4]). These remarks might seem to suggest that Wang does not view his doctrine as an accurate description of knowledge and action. And this suggestion might lead one to go further: one might hold that Wang does not aim to describe knowledge or action, but instead to develop new concepts which he believed would, if inculcated in the right way, help people to get themselves to be virtuous.

I have argued that—at least in the case of the unity of knowledge and action—such a radical conception of Wang’s aims not only has no textual basis, but is in fact undermined by the very methodological remarks which might have inspired it. Wang draws a stark contrast between the fact that his doctrine was proposed in order to remedy the ills of his day, and the fact that it describes knowledge and action in their original natural condition. He is clear that he takes his doctrine to accurately describe this original natural condition.

Of course, it is possible that Wang fails to live up to this aim. It is possible that in putting forward Unity, Wang analyzes something as “knowledge” or “action” which does not resemble anything a reasonable person could mean by either of those terms. Perhaps most obviously, what I have said here leaves it open that Wang understands “genuine knowledge” as an aspect of the mind completely different from what might be reasonably called “knowledge.” Elsewhere I argue that Wang does mean something reasonably understood as “knowledge” by “genuine knowledge” (Lederman 2022, [forthcoming](#)). But even if he does not, that fact would not show that he did not *aim* to describe knowledge and action; people often fail to live up to their aims.

In discussing Wang’s aims as a philosopher I have focused on whether he aimed to describe the nature of knowledge and action, or whether he was engaged in some kind of conceptual engineering. But this question bears on a second one: whether Wang was in any sense a systematic thinker. Tu Wei-ming is characteristically eloquent in presenting the problem:

It is not to be wondered that there exist many constructions of WANG Yangming’s philosophy. One is easily led to believe that the datum itself is of such a plastic nature that without much artificial effort it can be shaped into a variety of designs according to the intentions of the designers. How much this is attributable to Yang-ming’s own decision to discourage any systematization of his ideas remains an open question. But the act of constructing a philosophical edifice for Yang-ming cannot be justified merely on grounds that such an intellectual game has been widely and continuously practiced.

...I am increasingly suspicious of the claim that the reconstruction of Yang-ming’s philosophy involves no more than the procedure of integrating his basic precepts into a rational system. Nor do I accept the presumption that the content of Yang-ming’s thought can be encapsulated in a more or less rigorously related set of formulas. For I cannot endorse the optimism that, once the blue-print of Yang-ming’s philosophical edifice is dissected and analyzed, the task of reconstruction can be painlessly accomplished by a group of professional builders. (Tu 1973: 187)

Tu is careful here not to commit himself to the claim that there is no interesting systematic theory underlying WANG Yangming’s thought. But his remarks strongly

suggest skepticism that Wang’s thought had any “inner logic.” Tu seems to see this as a good thing: he sees the lack of an inner logic as indicating that WANG Yangming was much more ambitious than a mere theoretical philosopher. It suggests to him that Wang did not seek to present a cold skeleton of abstract principles, but instead (and what is in Tu’s view more important) a living invitation to engage in practical self-improvement. Others, however, may not share Tu’s excitement. For, if Tu is right, it would vindicate an already prevalent, but somewhat disparaging, verdict on Wang as a philosopher. Wang would be seen as a charismatic teacher, who made suggestive remarks on a range of topics, but who lacked the discipline to see these remarks through. Wang’s writings may be inspiring, but that is where their interest ends, for there is simply no worked-out idea behind the showy surface of his alluring style.

Tu does not say explicitly here that those who aim at systematic reconstruction of Wang’s thought are committed to denying that Wang saw his practical aims as paramount. But his remarks strongly suggest that he believes this is true. Like GU Dongqiao, Tu seems to think that Wang’s doctrine is either designed for a practical purpose or focused on a theoretical claim about the nature of knowledge and action—it cannot be both. But this is precisely the idea that Wang rejects in the section of his response to GU Dongqiao which I quoted at the very start of the essay, as [T1]. Wang tells us that his doctrine of the unity of knowledge and action is both a remedy for a disease, and also a description of the natural condition of knowledge and action. Once we distinguish these aspects of Wang’s thought, we can see a simple, bold theoretical claim underlying part of Wang’s doctrine. Wang certainly understood the therapy he administered as his most important philosophical aim. But he took that therapy to be grounded in his own discoveries in basic science, about the natural condition of knowledge and action.

Acknowledgments Thanks to Suhyun Ahn, P. J. Ivanhoe, Justin Tiwald, and Hwa Yeong WANG for detailed comments. Thanks also to two referees and especially HUANG Yong as editor of this journal for many suggestions and criticisms.

References

- Angle, Stephen. 2018. “Buddhism and ZHU Xi’s Epistemology of Discernment.” In *The Buddhist Roots of ZHU Xi’s Philosophical Thought*, edited by John Makeham. Oxford: Oxford University Press.
- Angle, Stephen C., and Justin Tiwald. 2017. *Neo-Confucianism: A Philosophical Introduction*. Cambridge: Polity.
- Chan, Wing-tsit. 1963. *Instructions for Practical Living and Other Neo-Confucian Writings by WANG Yang-ming*. New York: Columbia University Press.
- . 陳榮捷. 1983. *Wang Yangming’s Instructions for Practical Living with Detailed Notes and Collected Discussions* 王陽明傳習錄詳註集評. Taipei 臺北: Taiwan Xuesheng Shuju 臺灣學生書局.
- Chen, Lai 陳來. 1991. *The Boundary between Being and Non-Being: The Spirit of WANG Yangming’s Philosophy* 有無之境: 王陽明哲學的精神. Beijing 北京: Renmin Chubanshe 人民出版社.
- Chen, Lisheng 陳立勝. 2015. “What Kind of ‘Unity’? How to ‘Unify’? A New Interpretation of WANG Yangming’s ‘Unity of Knowledge and Action’ 何種‘合一’? 如何‘合一’?——王陽明知行合一說新論.” *Journal of Guiyang University (Social Sciences)* 貴陽學院學報(社會科學版) 3: 2–9.
- . 2019. *The Pivot Between Human and Saint* 人聖之機. Beijing 北京: Sanlian Shudian 三聯書店.

- Cheng, Chung-yi 鄭宗義. 2018. "A New Interpretation of WANG Yangming's Doctrine of the Unity of Knowing and Acting 再論王陽明的知行合一." *Academic Monthly* 學術月刊 50.8: 5–19.
- Ching, Julia. 1972. *The Philosophical Letters of WANG Yang-ming*. Canberra: Australian National University Press.
- _____. 1976. *To Acquire Wisdom: The Way of WANG Yang-ming*. New York: Columbia University Press.
- Cua, Antonio S. 1982. *The Unity of Knowledge and Action: A Study in WANG Yang-ming's Moral Psychology*. Honolulu: University Press of Hawaii.
- 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.
- _____. 2002. *The Unity of Knowledge and Action: Toward a Nonrepresentational Theory of Knowledge*. Albany: State University of New York Press.
- Huang, Yong. 2015. *Why be Moral?* Albany: State University of New York Press.
- _____. 2017. "Knowing-That, Knowing-How, or Knowing-To?" *Journal of Philosophical Research* 42: 65–94.
- _____. 2020. *Belief, Desire, and Besire*. London: Bloomsbury Publishing.
- Ivanhoe, Philip J. 2002. *Ethics in the Confucian Tradition: The Thought of Mengzi and WANG Yangming*. Indianapolis: Hackett Publishing.
- _____. 2009. *Readings from the Lu-Wang School of Neo-Confucianism*. Indianapolis: Hackett Publishing.
- Lao, Sze-kuang 勞思光. 1984–1986/2019. *A New History of Chinese Philosophy* 新編中國哲學史. Beijing 北京: Sandian Shudian 三聯書店.
- Lederman, Harvey. 2022. "The Introspective Model of Genuine Knowledge in WANG Yangming." *The Philosophical Review* 131.2: 169–213.
- _____. Forthcoming. "Conceptions of Genuine Knowledge in WANG Yangming." *Oxford Studies in Epistemology*.
- Lee, Ming-huei 李明輝. 1994. "WANG Yangming's Doctrine of the 'Unity of Knowing and Acting' in the Light of Kant's Practical Philosophy 從康德的實踐哲學論王陽明的'知行合一'說." *Bulletin of the Institute of Chinese Literature and Philosophy* 中國文哲研究集刊 4: 415–440.
- Li, Jingde 黎靖德, and Wang Xingxian 王星賢, eds. 1986. *Categorized Sayings of ZHU Xi* 朱子語類. Beijing 北京: Zhonghua Shuju 中華書局.
- Mou, Zongsan 牟宗三. 1972. "The Differentiation and Development of the Learning of WANG Yangming 王學的分化與發展." *New Asia College Academic Annual* 新亞書院學術年刊 14: 89–131.
- _____. 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.
- Shi, Weimin. 2017. "The Quest for Ethical Truth: WANG Yangming on the Unity of Knowing and Acting." *Comparative Philosophy* 8.2: 46–64.
- Shu, Jingnan 東景南. 2017. *Extended Biography of WANG Yangming* 王陽明年譜長篇. Shanghai 上海: Shanghai Guji Chubanshe 上海古籍出版社.
- Shun, Kwong-loi. 2010. "ZHU Xi's Moral Psychology." In *Dao Companion to Neo-Confucian Philosophy*, edited by John Makeham. Dordrecht: Springer.
- _____. 2011. "WANG Yang-ming on Self-Cultivation in the *Daxue*." *Journal of Chinese Philosophy* 38.s1: 96–113.
- _____. 2018. "On the Idea of 'No Self'." *Proceedings and Addresses of the American Philosophical Associations* 92: 78–107.
- Tiwald, J., and Bryan W. Van Norden. 2014. *Readings in Later Chinese Philosophy*. Indianapolis: Hackett Publishing.
- Tu, Wei-ming. 1973. "Subjectivity and Ontological Reality: An Interpretation of WANG Yang-ming's Mode of Thinking." *Philosophy East and West* 23.1/2: 187–205.
- _____. 1976. *Neo-Confucian Thought in Action: WANG Yang-Ming's Youth*. Berkeley: University of California Press.
- Van Norden, Bryan. 2019. "WANG Yangming." In *The Stanford Encyclopedia of Philosophy*, Fall 2019 ed., edited by Edward Zalta. <https://plato.stanford.edu/entries/wang-yangming/> (last accessed in July 2020).
- Wang, Xiaoyu 王孝魚, ed. 2004. *Collected Works of the Two Chengs* 二程集. Beijing 北京: Zhonghua Shuju 中華書局.
- Watson, Burton, trans. 1993. *The Lotus Sutra*. New York: Columbia University Press.
- Wu, Guang 吳光, QIAN Ming 錢明, DONG Ping 董平, and YAO Yanfu 姚延福, eds. 2011. *Collected Works of WANG Yangming* 王陽明全集. Shanghai 上海: Shanghai Guji Chubanshe 上海古籍出版社.
- _____. 2018. "The Theory of 'The Unity of Knowledge and Action' As *Liangzhi* Ethics 作為良知倫理學的'知行合一'論——以'一念動處便是知亦便是行'為中心." *Academic Monthly* 學術月刊 5: 14–24.

- Yu, Zhenhua 郁振華. 2014. “On Knowing How in Morality and Metaphysics: An Exploration of Gilbert Ryle and WANG Yangming 論道德—形上學的能力之知——基於賴爾與王陽明的探討.” *Social Sciences in China* 中國社會科學 12: 22–41.
- Zheng, Zemian. 2019. “An Alternative Way of Confucian Sincerity: WANG Yangming’s ‘Unity of Knowing and Doing’ as a Response to ZHU Xi’s Puzzle of Self-Deception.” *Philosophy East and West* 68.4: 1345–1368.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.