

The Introspective, Perceptual, and Spontaneous Response Models of Wang Yangming's Philosophy

Philip J. Ivanhoe*

Abstract

This essay presents and evaluates Harvey Lederman's Introspective Model as an interpretation of Wang Yangming's philosophy. According to the Introspective Model, one achieves genuine knowing when pure knowing assays one's occurrent thoughts and discovers that there are no doxastic conflicts among them. Lederman compares the Introspective Model to what he calls the Perceptual Model and argues for the superiority of the former over the latter as an interpretation of Wang's philosophy. I argue that neither the Introspective nor the Perceptual Model offers a plausible account of Wang's philosophy; both go awry by construing Wang's thought as primarily concerned with epistemology while largely ignoring its metaphysical and moral dimensions and historical context. I provide a sketch of an alternative, the Spontaneous Response Model, and argue that it avoids the shortcomings of the other two models and enables one to understand Wang's philosophy in a more comprehensive and systematic way that sees it as intimately related to the philosophical debates of its time. I also show the virtue of this alternative model as a basis for translating Wang's writings, which of course is itself an act of interpretation.

Keywords: pure knowing, genuine knowing, introspective, perceptual, spontaneous response, epistemology, doxastic conflict, knowing and acting, self-cultivation

* Philip J. Ivanhoe is Professor and Chair of the Department of East Asian Languages and Cultures at Georgetown University. He is also Editor-in-Chief of the *Journal of Confucian Philosophy and Culture*. E-mail: aiwenhe1954@gmail.com

** Thanks to Eirik L. Harris, Eric L. Hutton, Daryl Ooi, Justin Tiwald, Jean Tsui, Bryan W. Van Norden, and two anonymous reviewers for the *Journal of Confucian Philosophy and Culture* for helpful comments and suggestions on earlier drafts of this essay.

I. The Introspective and Perceptual Models of Genuine Knowing

Harvey Lederman has argued for a new interpretation of Wang Yang-ming's notion of genuine knowing (*zhen zhi* 真知),¹ independent of its relationship to action, based upon what he calls the Introspective Model (see Lederman 2022).² According to the Introspective Model, one achieves genuine knowing, which is an elevated or enhanced form of knowing, when pure knowing (*liang zhi* 良知) assays one's occurrent thoughts and there are no doxastic conflicts among them.

I will develop a new interpretation of genuine knowledge, according to which Wang characterizes it as an elevated form of knowledge, independently of its relationship to action. In my view, Wang holds that a person has genuine knowledge if and only if they are free from a particular form of doxastic conflict.³ (Lederman 2022, 170-71).

A person who possesses genuine knowing is not self-deceived or of two minds;⁴ they know in a clear, uncluttered, and unambiguous way. If they know the good in this way, they harbor no thought of it as being bad,

¹ I follow David S. Nivison in translating this term as "genuine knowing" (and not genuine knowledge) and *liangzhi* 良知 as "pure knowing" (and not pure knowledge) for reasons we both have described in several earlier publications.

² At the time this article was submitted, Lederman has published one article on Wang Yangming's philosophy and has another forthcoming. He has also submitted a third essay, circulated at least two additional manuscripts, and presented various versions of all these works. My description of his position is based upon a reading of all these resources.

³ This quote makes clear that Lederman believes his interpretation represents Wang's own view, that Wang himself saw freedom from doxastic conflict as the necessary and sufficient condition for genuine knowing.

⁴ On some occasions, Wang does talk about self-deception and being of two minds but he does not present these as epistemological problems; they are problems concerning a lack of sincerity (*cheng* 誠), which is a notion and approach with a long and rich history in the Chinese philosophical tradition. Wang thinks that all examples of moral failure entail a kind of self-deception, and his account of this is grounded in his metaphysics of pattern-principle and *qi* 氣: we fail to understand our true selves—our original nature—and appreciate the underlying oneness that exists between ourselves and all things, because our heart-minds are obscured by unclear or imbalanced *qi* and self-centered thoughts and desires. This general scheme is part of the shared intellectual context of Wang's philosophy.

they love it without a trace of hating it; if they know the bad in this way, they harbor no thought of it as being good, they hate it without a trace of liking it; and so, they are free of conflicting motivations as well as beliefs.

The Introspective Model is offered as an alternative to what Lederman refers to as the Perceptual Model, which comes in a variety of forms.⁵ The most distinctive feature of the Perceptual Model is that episodes of perception of the environment are a component of at least some episodes of genuine knowing or are components of the exercise of the disposition of genuine knowing. As its name suggests, the Perceptual Model is inspired by the similarity between the exercise of pure knowing and the exercise of various sense faculties. Some might be attracted to the Perceptual Model because in the course of his writings, Wang regularly appeals to a variety of comparisons between the exercise of pure knowing and the possession of genuine knowing⁶ and the natural functioning of various senses. One advantage of this model is that perceiving often is portrayed as spontaneous and incorporates motivational features and even action, as does the example from the *Great Learning* (*Daxue* 大學) of smelling a bad odor, which we shall discuss in greater detail below. In this example, knowing the odor immediately establishes that it is bad, one dislikes it wholeheartedly, and one's natural response is to want to move away from it as quickly as possible; these different mental states are presented as if they are a single, extended psychological phenomenon; the relation between them isn't mediated by deliberation or effort, an issue which we shall return to below. At this point, we can say that the Introspective Model is centrally concerned with the consistency of our beliefs, while the Perceptual Model involves to some degree our perceptions of the environment. In this essay, I contend that neither the Introspective Model nor the Perceptual Model offers a proper way to understand

⁵ One issue that distinguishes different versions of the Perceptual Model is whether all or only some episodes of genuine knowing concern phenomena external to the self. Both "Introspective Model" and "Perceptual Model" are names and concepts original to and proposed by Lederman.

⁶ On my interpretation, pure knowing is a kind of faculty—a faculty of moral sapience; genuine knowing is the result of the unimpeded exercise of this faculty.

Wang Yangming's philosophy because both misconstrue his central focus and concern as being epistemological in nature. Instead, we should understand Wang's philosophy in terms of what I will call the Spontaneous Response Model, an approach that takes him as pursuing the proper aims of every Confucian: moral agency, self-cultivation, and the ideal of effortless action.

II. Assessing the Two Models

At times, it is not clear whether the Introspective Model entails that one must be or is in any way aware of being free of doxastic conflict or whether this is simply an objective feature of and requirement for the possession of genuine knowing.⁷ If one is not in any way aware of the consistency of one's moral beliefs, then it is not clear how or even if one can know that a given case of one's own knowing is indeed genuine. If one is aware of being free of doxastic conflict or, what is different, if moral agents consciously aim to achieve this state, this would present additional challenges for the Introspective Model as an interpretation of Wang's philosophy. Such self-appraisal would seem to entail one thought too many and perhaps an infinite regress of appraisals;⁸ it would be to add something to the natural functioning of the heart-mind, which would interfere with its spontaneous response; this is something Wang explicitly warns against.

The Master once said to his students, "Not a single thought should be allowed to stick to the heart-mind in itself just as not the slightest trace of dust or dirt should be allowed to stick to the eye. It does not take much [dust or dirt] for the entire eye to see nothing but darkness."

⁷ Lederman seems to hold the latter view but the central role played by freedom from doxastic conflict in his account and statements such as "a person's *liangzhi* always knows the ethical qualities of their mental events" (Lederman 2002, 182) seem to point to a version of the former. The resolution of this issue would not affect most of the core arguments advanced in this essay.

⁸ Thanks to Eirik L. Harris for suggesting the second possibility.

He added to this saying, “The thought need not be a self-centered thought. Not even a good thought should be allowed to stick. If you put some gold or jade dust in the eye; the eye still will not be able to open.” (See Wang, *Record*, sec. 336)⁹

As we shall discuss more fully below, such moral self-assessment is anathema to Wang’s views about the heart-mind, pure knowing, pattern-principle, and oneness; in particular, it disrupts the natural functioning of genuine moral knowing and acting. Contrary to what the Introspective Model claims, introspection is only a method that one must at times employ in order to allow pure knowing to operate fully and unimpeded.¹⁰

Apart from the question of whether one must or can be aware of being free of doxastic conflict, a major challenge for the Introspective Model is that Wang *never* talks about doxastic conflict as the source of moral failure. For Wang, moral failure is the result of the interference posed by impure and unbalanced *qi* 氣 and self-centered thoughts and desires to the natural and spontaneous functioning of pure knowing. Advocates of the Introspective Model might respond by denying that their interpretation entails that eliminating doxastic conflict is a conscious aim of moral agents or even that they are or can be aware that they are free of doxastic conflict (though claims like the one cited in note nine above present challenges for the latter response); instead, they might insist that a moral agent is focused only on getting rid of bad beliefs, with bad beliefs understood as self-centered thoughts and desires that conflict with the inclinations of pure knowing. This may still involve one thought too many and constitute being attached to a particular ideal of purity of thought—namely, to be free of “bad beliefs.”¹¹ Wang certainly wants us to be attentive and aware of the

⁹ All translations of Wang’s works were made by the author of this essay.

¹⁰ Thanks to Eric L. Hutton for suggesting this way to describe our different interpretations of Wang’s view.

¹¹ Wang was not deeply concerned with beliefs—good or bad. Of course, he was concerned, to some extent, about how people do things like interpreting the classics, which involve various beliefs (though in this regard he usually talked more about “theories” or “what people say” than beliefs). But Wang did not typically describe lesser forms of knowledge in terms of wrong beliefs. For example, he thought just about everyone in his time

arising of self-centered thoughts and desires, but he never talks about eliminating bad *beliefs* or doxastic conflict. He teaches that if we remain on guard, if we are attentive and aware of the arising of self-centered thoughts and desires, our pure knowing has the power to naturally and immediately identify and eliminate them. We should not harbor any conception of some further aim or dwell on our achievements. The ideal heart-mind is to operate like a clear, bright mirror, functioning effortlessly and spontaneously in accord with pattern-principle.

If we set this concern aside for the sake of argument then, on the view under consideration, overcoming doxastic conflict is merely a by-product of the effort to eliminate self-centered thoughts and desires; it simply describes the mental state required for genuine knowing to occur. However, if the elimination of doxastic conflict and the harmonization of beliefs is just a side-effect of the operation of pure knowing, it is at best odd that Wang never highlights or explains this, as such a view would represent a claim unique among and uncharacteristic of neo-Confucian thinkers. Moreover, it is not clear why Wang or anyone would think that being in a mental state devoid of such doxastic conflict produces the harmonization of desires and feelings that are characteristic of spontaneous moral response or how such a state would guarantee genuine knowing. A person indifferent to a moral situation they are facing would be free of doxastic conflict but clearly would lack the harmonization of desires and feelings characteristic of spontaneous moral response and would lack genuine knowing.¹² If I know that a certain direction is north and do not believe that it is also south or that any other direction is north, then I am free of doxastic conflict. Nevertheless, I may simply be wrong. Consistency of belief does not generate either harmonization of feelings or guarantee correctness

What was wrong with their knowledge was that it lacked the right affective dimension or charge. Of course, his issue is complicated by his view that knowing and acting are united. To believe something for Wang is to feel something and this is to be inclined to do something. This, in brief, is why he saw self-centered thoughts and desires as fundamentally interrelated. Thanks to Daryl Ooi for pressing me on this point.

¹² On my understanding of Wang's philosophy, he would say that the pure knowing of the indifferent person still knows what to do but this knowing is obscured by impure *qi* and self-centered desires. The point is, it is hard to see how such a situation involves entertaining beliefs.

of belief. This is a problem for the Introspective Model, but it is not a problem for Wang because, as we have implied above and will make clear below, he does not hold that the elimination of doxastic conflict is the justification or proof of the possession of genuine knowing. Genuine knowing depends upon engaging and responding to the world by living in the light of pure knowing, which naturally detects, follows, and brings one into accord with pattern-principle. The Introspective Model holds that morality is purely a matter of having the right kinds of mental states. This is consistent with the widespread view, among contemporary Anglo-American philosophers, that moral facts depend upon and inhere exclusively in the mind, which is true, in different ways, for Kantians and well as Humeans. In general, Confucians hold that in order to genuinely possess and manifest virtue, one's heart-mind must be in an appropriate state, but they also hold that these internal states are responses to objective, external states and standards. Wang Yangming holds a version of this characteristically Confucian view.¹³ Ultimately, he grounds genuine knowing in the objective standard of pattern-principle, which provides the underlying basis for the unity or oneness of self and world. The Introspective Model largely avoids mentioning or relying upon such ideas or any of Wang's metaphysical beliefs, and in particular his signature claim that "heaven, earth, and the myriad things are one body"; this is the primary source of a number of difficulties.

Eliminating doxastic conflict is a concern of Socrates, at least as portrayed in the early Socratic dialogues, and this is one way to understand his use of the *elenchus*, by means of which he elicits views from his interlocutors and reveals inconsistencies among them, leading them to gain new knowledge. This, though, is not the characteristic approach of Wang Yangming. Wang is not trying to lead his disciples to proper awareness of propositional knowledge; he is seeking to awaken and enlighten them to the robust moral guide—pure knowing—complete, perfect, and standing ready within them. He is seeking to eliminate things, such as self-centered thoughts and desires, that interfere with the spontaneous extension of pure knowing

¹³ Thanks to Bryan W. Van Norden for noting the importance of this set of issues.

(*zhi liangzhi* 致良知).¹⁴ This view about the nature and function of pure knowing arguably is Wang's most original and distinctive contribution to Confucian philosophy.¹⁵ The Introspective Model claims to provide an account of genuine knowledge "independently of its relationship to action," but like Mengzi before him, whenever Wang discusses choices that reveal the operation of pure knowing, the cases concern choosing one *action* over another. For example, Wang regularly cites examples such as Emperor Shun marrying without first informing his parents or King Wu launching a military operation before burying his father.¹⁶ In this respect, Wang is simply following Mengzi, who, in other famous passages, asks us to choose between saving our drowning sister-in-law or observing ritual propriety or to choose to live or to uphold righteousness.¹⁷ None of these have anything to do with eliminating doxastic conflict through an introspective process that is independent of action.¹⁸

Another challenge for the Introspective Model is that Wang often invokes examples and metaphors that clearly concern the perception, construal, assessment, and response to things outside the self. An example of this type, though one that Wang does not invoke for reasons we shall explore below, is found in one of the most famous passages of Song dynasty neo-Confucianism, in which Cheng Yi explained the difference between ordinary knowing (*chang zhi* 常知) and genuine knowing (*zhen zhi* 真知) by presenting a story about a villager who recently had been mauled by a tiger. Cheng Yi says that everyone who heard that there was a tiger in the area attacking people reacted with alarm and fear, but this one villager, who had actually been mauled by

¹⁴ Since pure knowing combines both cognitive and affective elements, when it is fully deployed it will result in action.

¹⁵ As we shall see when we discuss the Spontaneous Response Model below, the extension of pure knowing is intimately related to Wang's metaphysical belief that heaven, earth, and the myriad things are one body. The practical, moral realization of this metaphysical truth is a comprehensive sense of care for all people, creatures, and things. For examples of the relationship among these ideas, see Wang (*Record*, sec. 179, 272). Thanks to an anonymous reviewer of an earlier draft of this essay for raising this point.

¹⁶ These are two common examples among many. See Wang (*Record*, sec. 139).

¹⁷ See *Mengzi* 4A.17 and 6A.10 respectively. See also 7A.15, where the term pure knowing (*liang zhi* 良知) first appears, which clearly concerns knowledge about *what to do*.

¹⁸ I have benefitted tremendously from comments and suggestions by Eric L. Hutton in thinking through and formulating the points made in this paragraph.

a tiger, reacted differently. This one man, Cheng Yi claims, experienced genuine knowing; the other villagers merely possessed ordinary knowing—the kind of knowledge an inexperienced child possesses—about tigers being dangerous and warranting fear. This inferior kind of knowledge does not involve mistaken belief or doxastic conflict, it concerns the affective quality of one’s knowledge. To the best of my knowledge, almost all neo-Confucians and their modern interpreters have taken Cheng Yi’s story as a paradigmatic case of genuine knowing: experiencing appropriate fear of tigers in the visceral way the mauled villager did exemplifies the kind of robust, affectively informed response characteristic of genuine knowing. If Wang indeed holds the Introspective Model, he almost certainly would reject Cheng Yi’s example, for it appears to involve a perception and affective reaction to something out there in the world.

Drawing on the work of Chen Lai, Lederman notes that Wang does not cite Cheng Yi’s story about the tiger or use the term “ordinary knowing” in his writings (see Lederman 2019). While true and important for reasons we will discuss below, this observation in no way supports the claim that Wang did not recognize the difference between lesser and higher forms of knowing or that he did not understand genuine knowing in the way that Cheng Yi promoted in this story and his other writings.¹⁹ Wang clearly does employ the concept of lower or ordinary forms of knowing throughout his work,²⁰ and he uses the phrase “genuine

¹⁹ In his more recent work, Lederman makes clear that Wang accepts and employs Cheng Yi’s conception of pure knowing. See, for example, note 22, in Lederman 2022. In forthcoming work (Lederman forthcoming 2022, note 16) he claims that Cheng Yi’s famous example of the man mauled by a tiger doesn’t directly support the idea that episodes of genuine knowledge have episodes of perceiving as parts” but (note 47) that this poses no problem for his interpretation because Wang “did not hesitate to diverge” from even eminent predecessors. But since, as Lederman himself recognizes, Wang *didn’t* diverge from Cheng Yi’s conception of pure knowing, it seems odd to think he rejected such an intimately related idea as genuine knowing without the slightest explanation or comment.

²⁰ For example, in section five of his *Record for Practice*, Wang says, “One cannot say she knows filial piety or brotherly respect simply because she *understands* how to say something filial or brotherly” (不成只是曉得說些孝弟的話, 便可稱為知孝弟). In earlier published work, I translated 曉 as “knows,” but I now see this as another example where Wang is intentionally avoiding the term 知; he clearly is denying that this person possesses genuine knowing. Thanks to Eric L. Hutton for pointing this out to me.

knowing” with the same sense that Cheng Yi highlights in this famous story: an enhanced type of knowing that has necessary implications for perception, construal, assessment, response, and acting. He seems to have other reasons for not citing the story about the tiger or using the term “ordinary knowing,” which is what we shall argue below.

Another passage, from chapter three of the *Great Learning*, mentioned above, presents additional challenges for the Introspective Model.

Seeing something beautiful is a case of knowing, while liking something beautiful is a case of acting.²¹ As soon as one sees something beautiful, one likes it. . . Smelling a bad odor is a case of knowing, while disliking a bad odor is a case of acting. As soon as one smells that bad odor, one naturally dislikes it. It is not as if you first smell it, and only then, intentionally, you decide to dislike it. Consider the case of a person with a stuffed-up nose. Even if he sees a malodorous object right in front of him, the smell does not reach him, and so he does not dislike it. This is simply not to know the bad odor.²²

Wang insists that a person with a stuffed-up nose *does not know* the odor of a malodorous object right in front of them—though they might well “know” that it is foul-smelling (an example of a lesser form of knowing)—because they do not know it in the right, visceral way; their perception does not involve affective and conative responses and constitute the beginning of acting. When Wang discusses this passage, he does not talk about ordinary knowing but simply says that “this is simply *not to know* the bad odor” (*bu hui zhi chou* 不會知臭). Here, Wang makes an intentionally strong and dramatic claim designed to achieve

²¹ Some interpret this passage as describing things that elicit spontaneous reactions of love and hate, but I understand it to be about things we naturally like/are attracted to (*hao* 好) and dislike/naturally turn away from (*wu* 惡). Among other things, this strikes me as more in line with what Mengzi says about our nascent moral sprouts. Zhu Xi as well as other commentators interpret the passage in the former way and understand “something beautiful” (*se* 色) as “an object of sexual desire.” Thanks to Bryan W. Van Norden for pressing me on this issue and in particular to explain why I do not follow Zhu Xi (though I believe his sense of the passage is for the most part correct and close to my own).

²² Wang discusses this passage in four places in his *Record for Practice*, the first being section 5 in conversation with Xu Ai, which is quoted here.

a rhetorical effect aimed at motivating proper practice—something he often does, as we shall argue below. He denies that the person has anything *worthy* of the name knowledge and that such knowledge is not worth pursuing. His choice to avoid the term “ordinary knowing,” noted above, and his preference, later in life, to talk about “the extension of pure knowing” (*zhi liangzhi* 致良知) when discussing this set of issues²³ are highly significant. By singling out the faculty of genuine knowing with a distinctive, special name—derived from Mengzi and made famous by thinkers like Cheng Yi—and giving it a central role in self-cultivation, Wang raises it up not only as a *more valuable* kind of knowing but—as his own words strongly imply—the only kind of knowing worth having or pursuing.

As Lederman notes (his note 47 cited in note 19 above), on the Introspective Model, Wang cannot accept the passage from the *Great Learning* as an example of genuine knowing, and yet Wang clearly offers it as a paradigm for genuine knowing several times in the course of his recorded teachings. The entire passage makes clear that genuine knowing occurs only when the heart-mind properly construes and responds to phenomena in the actual world. One who is not being stimulated by the world (e.g. someone with a stuffed-up nose) cannot construe and respond properly (by disliking it and moving away from it) and so cannot possess genuine knowing. If all genuine knowing is introspective, one cannot have genuine knowledge of things like colors, sights, or smells. Lederman claims that there is no evidence that Wang saw these examples from the *Great Learning* as cases of genuine knowing; instead, he interprets them as presenting “analogues” of such cases.²⁴ Nevertheless, Wang’s own words here and elsewhere, the connections this passage has to the rest of his philosophical system, and the ubiquitous presence and understanding of this passage among neo-Confucian philosophers, make it difficult to accept this denial.

²³ Lai (2021, chap. 5) points out that several of Wang’s disciples report that Wang didn’t arrive at the doctrine of the extension of innate knowledge until after his big enlightenment experience at Longchang. As noted above, this teaching arguably is Wang’s most distinctive and original contribution to Confucian philosophy.

²⁴ See Lederman (2019), where Lederman explains what he means by this and notes that Antonio Cua holds a similar view.

In section three of his *Questions on the Great Learning* (*Daxuewen* 大學問), Wang explains that the highest good is “the nature endowed by heaven”; when it is luminously and clearly manifested, it is pure knowing and then,

. . . whatever is good will appear as good and whatever is bad will appear as bad. We will respond to the weighty or the light, the substantial or the insubstantial as these affect us. . . (see Wang, *Record*, 272)

Here Wang talks about how things *appear* to us, *being affected* by our perception of the different qualities of things in the world, and how such stimuli elicit *the proper responses* from our faculty of pure knowing. Later in this passage, he goes on to liken pure knowing to the use of rulers, compasses, and squares; this is an analogy he makes in other passages as well. For example,

Now pure knowing is to various circumstances as compasses, carpenter’s squares, and measuring lines are to circles, squares, and different lengths. . . . And so, when compasses and carpenter’s squares are truly set, they cannot be deceived in regard to circles and squares, yet can be applied to all the circles and squares in the world. When measuring lines are truly laid out, they cannot be deceived in regard to different lengths, yet can be applied to all the different lengths in the world. When pure knowing is truly extended, it cannot be deceived in regard to various circumstances, yet it can respond to all the various circumstances in the world. . . (see Wang, *Record*, sec. 136).

These passages do not avail themselves of the range of concerns that characterize the Introspective Model. Nowhere does Wang talk about avoiding doxastic conflict in any sense; instead, he talks about pure knowing properly engaging with, construing, and responding to the different qualities of things out there in the world and how the heart-mind, which is the conscious, knowing aspect of pattern-principle, provides an unerring standard for moral assessment, motivation, and action and the basis for realizing the underlying unity of heaven, earth, and the myriad things. These passages describe and illustrate by example the proper perception of and response to things and events.

The passages and arguments presented above make it implausible to claim that pure knowing and genuine knowing exclusively concern introspection and the consistency of one's beliefs. But does this mean that pure knowing and genuine knowing are primarily about perception? As seen and noted above, Wang does often liken the operation of pure knowing to the natural functioning of various senses: e.g. sight and smell. These explicitly involve not just perception but construal, assessment, and response as well. Such perception, construal, assessment, and response can be about our internal states like thoughts and feelings or things outside the self, though, as we shall argue below, given Wang's overall philosophy, this distinction between inside and outside is neither helpful nor sustainable. While the Perceptual Model offers several advantages over the Introspective Model, for a variety of reasons, to which we now turn, neither the Introspective nor the Perceptual Models as described above are plausible accounts of Wang's philosophy.

One general problem that both models share is that they are primarily about epistemology, roughly, a theory of knowledge. This imposes a characteristically Western philosophical concern and approach to Wang's philosophy and takes as its object of analysis a phenomenon that can in principle be divorced from action, motivation, and emotion. Of course, in itself, this does not constitute a valid objection to either model, but a fuller account of Wang's philosophical system reveals why such an approach is mistaken.

First, it should strike anyone familiar with Wang's view as deeply problematic to interpret him as interested primarily in knowledge when one of his most famous signature teachings is the *unity of knowing and acting*.²⁵ For Wang, any view that clearly separates knowing from acting fails as an account of knowing or acting and *a fortiori* as an account

²⁵ Lederman explicitly argues that Wang provides just such an account of knowledge, i.e. one that stands "independently of its relationship to action." See the quote in the opening section of this essay. In addition to the article in which this quote appears, which is devoted to describing Wang's theory of knowledge, Lederman focuses on epistemology in all of his other writings and presentations on Wang; for example, this is the core theme of his article "Conceptions of Genuine Knowledge in Wang Yangming," which is forthcoming in *Oxford Studies in Epistemology*.

of genuine knowing or acting. As we noted in our discussion of the story about the villager mauled by a tiger and the passage from the Great Learning about having a stuffed-up nose, Wang insists that all cases in which knowing and acting are separated fall short of genuine knowing, and he refuses to acknowledge that they are knowing in any worthwhile sense. Both the Introspective and Perceptual Models, as described above, violate Wang's signature teaching concerning the *unity of knowing and acting*.

Second, each of the models described above privileges, to varying degrees, one or the other side of the self. The Introspective Model explicitly claims that pure knowing is wholly an internal affair, while the Perceptual Model insists that it, at the very least, at times concerns the exterior world and that perception of the exterior world is the paradigm for how pure knowing works. In any event, both models depend upon an interior/exterior distinction, but this very distinction violates another of Wang's signature teachings: *that heaven, earth, and the myriad things form one body*. The Introspective Model is particularly problematic in this regard as it insists that the exercise of pure knowing and genuine knowing do not take any part of the exterior world as their intentional objects.

III. An Alternative Model for Understanding Wang's Philosophy

Wang's teachings about the *unity of knowing and acting and that heaven, earth, and the myriad things form one body* both depend on a set of robust metaphysical claims, and this "heroic"²⁶ metaphysical system reflects a general scheme shared not only by almost every neo-Confucian but also, as we shall argue below, that was simply part of the background beliefs of any educated person in the Ming dynasty. Any viable interpretation of pure and genuine knowing must be able to account for and is likely to depend in considerable measure upon this shared set of metaphysical

²⁶ I refer to such a system as heroic not only because from a modern point of view it takes something like courage to embrace such a metaphysical system but also because such a commitment entails a profoundly challenging and majestic moral vision.

beliefs. Like Zhu Xi, Wang believes that the universe is composed of pattern-principle (*li* 理) and *qi* 氣, that pattern-principle provides the norms and structures that underlie the world's physical and moral order, while *qi* constitutes its physical form and provides a more palpable shared basis unifying the things of the world. While Zhu and Wang do differ on certain aspects of this general scheme, often, their differences are more matters of degree than kind. This is not to deny that these disagreements at times lead to distinct views and disputes not only about metaphysics but how best to cultivate oneself.²⁷ For example, Wang emphasizes much more strongly and dramatically than does Zhu Xi that the heart-mind is pattern-principle and elaborates upon this by claiming that the heart-mind is the conscious and aware aspect of pattern-principle and so that it is the heart-mind of the universe. None of these claims represent a dramatic departure from the standard neo-Confucian view, which he shared with Zhu, but they do help us understand some of the most characteristic and arresting aspects of his teaching. In particular, these features of Wang's philosophical system help us understand his distinctive teaching about the extension of pure-knowing, as described above, and are why he is regarded as a member of the Learning of the Heart-Mind (*xin xue* 心學) school of neo-Confucian philosophy.

Some scholars have argued that the metaphysical differences between Zhu and Wang are rooted in something more substantial and profound. For example, several appeal to Wang's claim that "the heart-mind is pattern-principle" (*xin ji li* 心即理) and present it as the deep metaphysical insight that distinguishes Wang's views from fellow neo-Confucians. But, in fact, the Cheng brothers said similar things and Zhu Xi could in principle agree with this claim, depending on how one understands it.²⁸ In claiming that "the heart-mind is pattern-principle," Wang was not staking out an unprecedented and unique metaphysical

²⁷ For an argument in support of this claim, see Ivanhoe (2016).

²⁸ For example, Cheng Hao declares "the heart-mind is pattern-principle; pattern-principle is the heart-mind" (心是理, 理是心) in Cheng and Cheng (*Extant Works*, chap. [juan 卷] 5) and Cheng and Cheng (*Collected Works*, chap. 13). For a valuable discussion of this general issue, see Angle and Tiwald (2017, note 13). Thanks to Justin Tiwald for help in regard to this issue, including the reference to Cheng Hao's work.

position but emphasizing how knowledge of the pattern-principles of things must be drawn from one's own heart-mind and not by attempting to discover—through the “investigation of things”—pattern-principle from things in the exterior world (an issue to which we shall return below). Like his claim that those who do not possess genuine knowing do not really know *at all*, such dramatic claims are made primarily for their therapeutic effects and are aimed at inspiring proper practice, and not because he takes himself to have a fundamentally different metaphysics than Zhu Xi.

This is true in regard to another of Wang's most dramatic claims: that there are no pattern-principles outside the heart-mind (*xin wai wu wu* 心外無物) (see Wang, *Record*, sec. 6). Many have mistakenly taken him to be making a metaphysical claim about where pattern-principles exist and saying that they can *only* be found within the heart-mind. This would make him a particular kind of idealist, similar in important respects to thinkers like Berkeley. But this is not what Wang is claiming. Here again, he is not denying the standard neo-Confucian view that pattern-principle underlies and provides the world's physical and moral order; if he were, that would warrant direct, careful, and extensive treatment and create a number of problems for his overall view. He is instead insisting that the heart-mind *does not lack* any of these pattern-principles.²⁹ Had Wang intended to deny that there is pattern-principle outside the mind, he would have been making a claim that would have sounded outrageous not only to fellow philosophers but to almost all of his contemporaries; this offers another good example of how context must inform and in some cases restrain our efforts to accurately interpretate an historical philosophical text.³⁰ Rather than a denial of the existence of pattern-principle outside the heart-mind, the claim that “there are no pattern-principles outside the heart-mind” is akin to the Biblical Noah saying, as the rain began to fall, “there are no animals not in the ark.” Of course, there are many animals that are not in the ark (unfortunately, most of them are about to drown), but Noah's claim is a

²⁹ For the definitive defense of this interpretation of Wang, see Tien (2010).

³⁰ For example, one sees the term employed throughout novels such as *Romance of the Three Kingdoms* (*Sanguo yanyi* 三國演義) to describe objective matters of fact about the external world. Thanks to Bryan W. Van Norden for suggesting this point and example.

convenient way of saying that “there are no animals of which there is no paradigm in the ark.”³¹ So, if one’s goal is to become familiar with every animal, one need not look outside those in the ark. One also might reasonably take him to be implying that the only way to ensure that one will gain a comprehensive knowledge of the variety of animals (and avoid drowning oneself) is to look inside the ark.

Wang was saying something even stronger: it is *only* by looking within that one can comprehend pattern-principle. This means, first, as argued above, that the heart-mind contains the complete set of pattern-principles; these do not need to be augmented by additional pattern-principles found outside the heart-mind. It also means, second, that the way to access pattern-principle is not to mine the things and events out there in the world through the “investigation of things” (*gewu* 格物) as advocated by Zhu Xi but to remain aware of and attentive to the pattern-principles within one’s heart-mind *as these come into contact with and respond to* things and events.³² The general account provided above also explains why Wang defined “things” (*wu* 物) as “thoughts” (*yi* 意) and *gewu* 格物 as “the rectification of one’s thoughts.” According to Wang, the proper approach to learning is to ensure that one eliminates all self-centered thoughts and desires; one must pursue this task by monitoring one’s thoughts in the course of one’s daily activities, searching for self-centeredness. Whenever a self-centered thought is detected and clearly identified by one’s pure knowing, it is immediately scotched and eliminated, one’s thoughts are thereby rectified, and one’s knowing will be genuine. Pure knowing is the lynchpin of Wang’s philosophy, which is why one finds him encouraging his disciples to simply *have faith in* (*xin* 信) their pure knowing.³³ In a remarkable passage, Wang describes how such faith depends upon establishing

³¹ I owe this splendid way of putting it to Bryan W. Van Norden.

³² Wang is famous for describing the unsuccessful attempt by him and his friends to discern the pattern-principle of bamboo by concentrating their attention and energy on physical bamboo plants; he also insisted that traditional study of the classics was not necessary and was instead potentially an impediment to moral improvement. This is not to deny that bamboos and the classics contain pattern-principle but to insist that focusing our attention and energy on “investigating” these things is not the way to access the pattern-principle within one’s heart-mind.

³³ See Qian Dehong’s comments to Wang (Record, sec. 167).

a commitment (*lizhi* 立志) and laments that establishing such a commitment is difficult to describe as it involves “creating something out of nothing” (*wu zhong sheng you* 無中生有) (see Wang, *Record*, sec. 115).³⁴ Of course, in keeping with Wang’s teaching of the unity of knowing and acting, genuine knowing results in action in accord with pattern-principle.

Given the picture of Wang’s philosophy sketched above, we must reject both the Introspective and Perceptual Models and recognize that Wang’s teachings and aims are not plausibly interpreted as primarily concerned with epistemology in general or consistency of beliefs in particular. Instead, they are focused on the proper aims of every Confucian: moral agency, self-cultivation, and the ideal of effortless action. In light of his signature teachings concerning the *unity of knowing* and acting and that *heaven, earth, and the myriad things form one body* and the metaphysical beliefs described and analyzed above, we should understand his teachings about pure and genuine knowing in terms of a seamless process of perceiving, construing, assessing, and acting.³⁵ Drawing upon the description of his philosophy sketched in the course of this essay, Wang’s approach to learning is grounded on claims characteristic of the standard neo-Confucian metaphysical scheme. The heart-mind has access to all the pattern-principles in the universe and is the only place where these pattern-principles can be conscious and aware (see Wang, *Record*, sec. 118).³⁶ Our heart-minds seek to understand pattern-principle and naturally will do so as we work to eliminate self-centered thoughts and desires, which violate our original nature, separate us from the underlying unity of the world, lead us to care less or not at all for parts of the world, and cut us off from and distort our innate and ever-available ability to know. As pure knowing purges our thoughts of every hint of self-centeredness, self-centered thoughts and desires dissipate, our *qi* becomes limpid and pure, and

³⁴ While his overall system is quite different, Wang is here grappling with a problem first discussed by Mengzi in *Mengzi* 6A.15.

³⁵ This is how Nivison and Ivanhoe describe this core teachings. See, for example, Nivison (1996, 228), Ivanhoe (2002, 99-100), and Ivanhoe (2011, 282).

³⁶ Such a view also captures the standard neo-Confucian teaching, shared by Zhu Xi, that the heart-mind is made to know.

the pattern-principle of the heart-mind shines forth and accords with the pattern-principle of the myriad things and events, revealing the intimate connection between the self and world or, as Wang puts it, the fact that heaven, earth, and the myriad things are one body. This general view about how to attain understanding of things and events is captured in the modern Chinese term *lihui* 理會 which literally means “pattern-principles meeting” and has the sense of “to understand.” There is no need to aim at, entertain, or invoke thoughts about avoiding doxastic conflict; such a concern would only erect a further impediment to the free and spontaneous flow of pure knowing and acting. For Wang, every occasion of genuine knowing is the beginning of acting: a heart-mind free of self-centered desires and thoughts functions like a clean, bright mirror: it spontaneously and unselfconsciously reflects accurately and without prejudice whatever it encounters, and its reflections entail responding fully and appropriately to each and every situation and circumstance, leaving no trace behind.

Such an interpretation captures all the characteristic features of Wang’s philosophy and avoids the range of problems generated by the two epistemologically-focused models explored above. Moreover, it offers a scheme through which one can naturally and properly translate and understand Wang’s writings. The Introspective Model does not allow one to do so and requires one to distort and misrepresent the text in various ways.³⁷ For example, Lederman renders some famous lines from Wang’s writings in the following way.

When a [good] motivating concern arises, the *liangzhi* of your mind already knows that it is good. Suppose you do not wholeheartedly love it but instead turn away from it and diminish it. You would then be taking what is good to be bad and obscuring your *liangzhi* which knows that it is good.³⁸ (Lederman, 2022, 187)

³⁷ I think the same can be said of the Perceptual Model, though I will not explore that issue here as I do not know of anyone who holds to this model as the basis for translating Wang’s works.

³⁸ “意念之發，吾心之良知既知其為善矣，使其不能誠有以好之，而復背而去之，則是以善為惡，而自昧其知善之良知矣” (Wang, *Record*, sec. 278).

In general, I prefer to translate *yi* 意 as “thought” rather than “motivating concern”³⁹ and here the translation calls for something like “thought or idea” since 意 appears alongside *nian* 念.⁴⁰ “Motivating concern” obscures the fact that many of our thoughts and ideas arise in response to our encounters with the external world, focuses attention inwardly, and creates what appears to be evidence for the Introspective View. Two additional features of Lederman’s translation illustrate even more clearly some of the ways in which the commitment to read Wang as chiefly concerned with epistemology leads him to mistranslate Wang’s text and create what appears to be evidence for the Introspective View.⁴¹ First, he translates *qu* 去 as “diminish” when the word most commonly has the meaning of “to cast out,” and, I would say, clearly does in this context.⁴² This is to take the word as describing a mental attitude when in fact it implies a physical action, as does *bei* 背 “turn your back.” Second, he translates *yi shan wei e* 以善為惡 as “taking what is good to be bad” thereby interpreting it as a case of doxastic conflict (i.e. believing X to be good while also believing it to be bad). But this is not the most natural way to read the phrase in light of its grammar and context and

³⁹ While by no means a fully adequate equivalent, “thoughts” does a better job of capturing several of the different senses of *yi*, which include being episodic, entailing both cognitive and conative aspects, and occurring in response to things in the world as well as internal ideas and feelings. Thanks to Stephen C. Angle and Justin Tiwald for helpful discussions of this term of art in Wang’s philosophy.

⁴⁰ In classical sources, the word *nian* 念 tends to mean thought or idea and especially repeated reflection or thought. Numerous clear examples of this usage can be found in the *Book of History* (*Shujing* 書經). See Legge ([1865] 1970; 55, 58, 324, 329, 390). Buddhists used the term with the special sense of momentary episodes of thought. Their ideal was to have thoughts arise and then quickly fade away; one was not to try either to suppress thoughts or allow them to become attached to the mind. Rather, thoughts should naturally and spontaneously come and go in response to whatever comes before them, as a mirror responds to the things brought before it. This is an idea that is prominent in the *Platform Sutra* (*Liuzu tanjing* 六祖壇經) and that greatly influenced Wang Yangming.

⁴¹ It is not altogether clear what motivates Lederman to interpret Wang’s philosophy as primarily concerned with epistemology. The examples discussed here are offered as illustrations of the potential hazard of beginning with and remaining within the grip of this hermeneutical commitment.

⁴² Cf. Wang’s use of this word in passages such as section 39 of a *Record for Practice*: “As soon as a single [self-centered] thought begins to stir, one must conquer and cast it out” (纔有一念萌動，即與克去). Note how closely “conquer and cast it out” mirror the grammar of section 278, “turn your back and cast it out.”

Wang's teaching about the unity of knowing and acting. Like the case of 去 above, what is implied is action rather than mental states like beliefs. To *cast out* the good is to *treat what is good as if it were bad*. And so, a more accurate translation would be,

Whenever a [good] thought or idea comes forth, the pure-knowing of your heart-mind knows that it is good; if you do not sincerely like it, but instead turn your back and cast it out, then you treat what is good as bad and obscure your pure knowing, which knows that it is good.

These contrasting translations offer more than just different renderings of the text, they represent profoundly different interpretations of Wang's philosophy. The former, which I contend requires contorting the original text, supports the Introspective Model, the latter, which offers a natural and straightforward rendering of the text, supports the Spontaneous Response Model. Only the latter properly captures Wang's account of the natural, effortless operation of pure knowing and the realization of genuine knowing.

IV. Conclusion

The primary aim of this essay is to present and assess Harvey Lederman's Introspective Model as an interpretation of Wang Yangming's philosophy. According to the Introspective Model, one achieves genuine knowing, which is an elevated or enhanced form of knowing, when pure knowing assays one's occurrent thoughts and discovers that there are no doxastic conflicts among them. Lederman compares the Introspective Model to what he calls the Perceptual Model, which he attributes to a number of other contemporary scholars, and argues for the superiority of the former over the latter as an interpretation of Wang's philosophy. I have argued that neither the Introspective nor the Perceptual Model offers a plausible account of Wang's philosophy; both go awry by construing Wang's thought as primarily concerned with epistemology and largely ignoring its metaphysical and moral dimensions and broader historical context. As part of this critical project, I have provided a sketch of an

alternative, the Spontaneous Response Model, and argued that it avoids the numerous and varied shortcomings of the other two models and enables one to understand Wang's philosophy in a more comprehensive and systematic way that locates it securely in the philosophical debates of his time and place. I have also shown the virtue of this alternative model as a basis for translating Wang's writings, which of course is itself an act of interpretation. The presentation of the Spontaneous Response Model offered here is only a sketch and leaves many issues in need of greater and more careful treatment; such work is well worth doing but doing so would take us far beyond the primary aim of this essay.

REFERENCES

- Angle, Stephen C., and Justin Tiwald. 2017. *Neo-Confucianism: A Philosophical Introduction*. Cambridge, UK: Polity.
- Chen, Lai. 2021. *The Spirit of Wang Yangming's Philosophy: The Realms of Being and Nonbeing*. Los Angeles: Bridge 21 Publications.
- Cheng, Hao 程顥, and Cheng Yi 程頤. *Er Cheng ji* 二程集 (Collected Works of the Two Chengs).
_____. *Henan Chengshi yishu* 河南程氏遺書 (Extant Works of the Two Cheng Brothers from Henan).
- Ivanhoe, Philip J. 2002. *Ethics in the Confucian Tradition: The Thought of Mengzi and Wang Yangming*. Indianapolis: Hackett Publishing.
- _____. 2011. "Moral Perception in McDowell, Wang, and Mengzi." *Dao: A Journal of Comparative Philosophy* 10 (3): 273-90.
- _____. 2016. *Three Streams: Confucian Reflections on Learning and the Moral Heart-Mind in China, Korea, and Japan*. Oxford, UK: Oxford University Press.
- Lederman, Harvey. 2019. "The Introspective Model of the Unity of Knowledge and Action." (Unpublished essay provided by the author.)
- _____. 2022. "The Introspective Model of Genuine Knowledge in Wang Yangming." *Philosophical Review* 131 (2): 169-213.
- _____. Forthcoming. "Conceptions of Genuine Knowledge in Wang Yangming." In *Oxford Studies in Epistemology*. Oxford, UK: Oxford University Press.
- Legge, James, trans. (1865) 1970. *The Shoo King, or the Book of Historical Documents*. Vol. 3 of The Chinese Classics. Hong Kong: Hong Kong University Press.
- Mengzi 孟子. *Mengzi* 孟子.
- Nivison, David S. 1996. "The Philosophy of Wang Yang-ming." In *The Ways of Confucianism*, edited by Bryan W. Van Norden. Chicago: Open Court.
- Tien, David. 2010. "Metaphysics and the Basis of Morality in the Philosophy of Wang Yangming." In *Dao Companion to Neo-Confucian Philosophy*, edited by John Makeham. Dao Companions to Chinese Philosophy Series. New York: Springer Publishing Company.
- Wang, Yangming 王陽明. n.d. *Chuanxilu* 傳習錄 (A Record for Practice).
_____. n.d. *Daixuewen* 大學問 (Questions on the Great Learning).