

13

Cognitive science and religious thought

The case of psychological interiority in the *Analects*

Edward Slingerland

One of the most commonly assigned secondary texts in university classes on early Chinese religious thought is Herbert Fingarette's classic *Confucius: The Secular as Sacred* (Fingarette 1972).¹ This is not only because of its brevity and the lucidity of its prose, but also because Fingarette's book marked a sea change in the manner in which Western philosophers approached early Chinese texts. Fingarette (1972) played a central role in inaugurating an era of much more nuanced, culturally sensitive interpretations of the *Analects*, as well as other early Chinese texts, in philosophical circles: an era in which Confucius no longer appeared as a watered-down Christian or "Axial Age" Kantian who occasionally liked to play dress-up and perform some strange rituals, but rather demanded serious philosophical attention as a unique thinker in his own right. Fingarette was one of the first Western philosophers to recognize that the early Confucian model of the self fundamentally challenges a particular understanding of the ethical self, and the self vis-à-vis culture and society, that remains quite prominent in modern Western philosophical and popular discourse. Taken seriously on its own terms, the *Analects* presents a vision wherein the individual is not an autonomous atom, freely pursuing its own rational self-interest, but is rather always already embedded in a web of familial, social and cultural connections. Cultural training is not perceived as an optional add-on to an otherwise self-sufficient and fully-developed individual, but rather as fundamentally constitutive of anything that could be acknowledged as genuine human personhood. Knowledge is not limited to abstract "knowing that" or mere assent to a set of principles, but is rather portrayed as a kind of embodied, and largely implicit, "know-how". Radical choice and self-conscious individualism are not, as Kant and his evil existentialist twins would have it, the very foundation of the ethical self, but rather symptoms of a historically and

globally quite anomalous modern Western *anomie*. *Confucius: The Secular as Sacred* can be seen as inaugurating an important trend in the study of early Chinese religious thought that holds up early Chinese conceptions of the self and society as important correctives to various excesses and blind-spots in modern Western philosophy.²

This trend has been, on the whole, a salutary one; a helpful antidote to Western cultural myopia, and particularly to the quite impoverished and psychologically implausible model of the self, rationality and culture that characterizes much of modern Western philosophical discourse.³ However, as with any medicine, an overdose can be even more harmful than the original malady. I wish to argue here that the philosophical trend inaugurated by Fingarette has in several respects gone too far in emphasizing the uniqueness of early Chinese thought, crossing the line between an appreciation of genuine difference and a quite harmful form of cultural exoticization that might be labelled “reverse Orientalism”.⁴

I will attempt to illustrate this point with a specific and quite dramatic claim that Fingarette made in his 1972 work: that Confucius of the *Analects* completely lacked any notion of psychological interiority. Fingarette makes it clear that he means this in the strongest possible sense:

I must emphasize that my point here is not that Confucius’ words are intended to exclude reference to the inner psyche. He could have done this if he had had such a basic metaphor in mind, had seen its plausibility, but on reflection had decided to reject it. But this is not what I am arguing here. My thesis is that the entire notion never entered his head. The metaphor of an inner psychic life, in all its ramifications so familiar to us, simply is not present in the *Analects*, not even as a rejected possibility.

(Fingarette 1972: 45)

Although over thirty-five years have passed since Fingarette originally made this claim, and despite that fact that it has come under criticism from several different angles in recent decades,⁵ it remains a viable position in the field. It is still maintained by Fingarette himself,⁶ and related stances, such as the claim that early Confucian thought is concerned entirely with role performance rather than any type of inner psychological individuality, are widely asserted by prominent scholars of early Chinese thought.⁷

In this chapter I would like to put this argument finally to rest by means of a two-part critique, the first employing more traditional religious-studies methodologies and the second illustrating the manner in which cognitive science can make original and significant contributions to debates in the study of religious thought. In part one, I will begin with more traditional textual

evidence strongly suggesting that, *pace* Fingarette, Confucius clearly must have had access to a concept of psychological interiority because this concept plays a prominent role in a textual tradition that he was very much dedicated to, that of the *Book of Odes* (*shijing* 詩經). In addition, I will argue that we can find clear evidence of the concept of psychological interiority in the *Analects*; that, in fact, major themes in the *Analects* make no sense without such a concept. In part two, I will explore two ways in which evidence from cognitive science is relevant to the debate: our modern understanding of the cognitive reality of metaphor suggests that we must take the interiority metaphors we see in early Chinese texts seriously, and our best current understanding of human cognition makes it highly unlikely *any* psychologically healthy member of the species *Homo sapiens*, anytime or anywhere, has lacked a concept of inner–outer boundaries to the self or psychological interiority. I will argue that these two pieces of evidence so radically change the burden of proof for Fingarette’s argument that it can no longer be plausibly maintained. I will then conclude with some thoughts on the important role that cognitive scientific evidence can serve as a hermeneutical limit-setter, and how adopting an empirically viable, embodied model of human cognition can significantly alter the interpretative landscape within which sinologists, philosophers and religious-studies scholars do our work.

Textual evidence

Psychological interiority in the Odes

Confucius of the *Analects* places enormous importance on the *Shijing* 詩經 or *Book of Odes*, which for him seemed to embody all of the aesthetic excellence and moral wisdom of the ancients. Although there has long been scholarly disagreement about the precise dating of the *Book of Odes*, it is the general consensus that the received text represents largely Western Zhou or earlier materials. That our received text has not been passed down to us unchanged since that time is made clear by the fact that Warring States texts often cite “lost” Odes, that is, verse that is attributed to the *Odes* but not present in our extant Mao edition. On the strength of this evidence alone it is difficult to specify with any precision or confidence the exact content of the text that Confucius so treasured. Recent archaeological evidence has demonstrated that there may have been more diversity in various editions of the *Odes* circulating in Warring States China than was previously thought (Kern 2005); nonetheless, it is clear that something broadly resembling our received version played a major role in Confucius’ life.

This being the case, indications that psychological interiority plays an important role in the *Odes* would suggest that Confucius of the *Analects* at least had

access to the concept, even if he then decided to ignore or reject it. Let us consider briefly a few particularly relevant *Odes*. In Mao 35 (“Valley Breeze” or *Gufeng* 谷風), a virtuous wife being sent away by her husband, who has apparently tired of her, says “I go along the road slowly, slowly/ In my innermost heart-mind reluctant” (行道遲遲／中心有違). Although *xin* 心 here is probably best rendered as “heart”, since it is emotions that are being emphasized, I will render it consistently as “heart-mind” because, as we shall see, it is the seat of both emotions and thoughts. The term that I have translated as “reluctant” (*wei* 違) means literally “opposed” or “going against”, in the sense that the poetess’s physical or outward behavior (travelling down the road, away from her former home) is in conflict with her “innermost heart-mind” (*zhongxin* 中心). One could not have clearer expression of conflict between inner psychological state and outer behaviour. Consider a similar sentiment expressed in Mao 65 (“The Wine-Millet Bends” or *Shuli* 黍離), where a poet filled with sorrow compares his bowed head and sunken posture to a millet stalk overlaid with grain:

Slowly I moved about,
 In my innermost heart-mind all-agitated.
 Those who know me,
 Say that my heart-mind is worried;
 Those who do not know me,
 Ask what it is that I am looking for.

行邁靡靡、中心搖搖。
 知我者、謂我心憂、
 不知我者、謂我何求。

The poet’s physical posture, with his bowed head and slow gait, suggests someone searching the ground for a lost object, hence those who are unaware of his inner sorrow ask what he is looking for. Those who know him, though, realize that he is not looking for anything: his gait and posture reflect, in fact, the metaphorical weight of profound sorrow and worry. Here we see again the idea that inner feelings are not necessarily obvious from one’s outward behaviour, with the additional implication that it is therefore difficult for outside observers to know for certain what is going on “inside” another person.

The term *zhongxin* 中心 (lit. “innermost heart-mind”, “heart of hearts”) appears sixteen times in the *Odes*, and clearly involves container logic, *zhong* 中 meaning “middle”, “inside” or “centre”. This innermost heart-mind contains one’s most intimate personal thoughts and feelings, which, because the heart-mind is encompassed and therefore masked by the outer container of the external body, makes it difficult for interior thoughts and feelings to be perceived from the outside. This leads to the possibility of a disjunction

between inner psychology and outer behaviour, although outer behaviour can be used as a clue to infer indirectly the contents of the heart-mind. This would seem to correspond quite closely to what we have in mind when we talk about “psychological interiority”.

The role of psychological interiority in the Analects

The *Odes* are by their very nature rather terse and suggestive. By the time we get to the *Analects* itself,⁸ the concept of psychological interiority can be found widely and clearly expressed, and indeed underlies some of the core themes and anxieties expressed in the text. Moreover, the container metaphor *zhong* (“centre” or middle”) that played such a dominant role in the *Odes* is joined by another vivid container metaphor, that of “inner” (*nei* 內) versus “outer” (*wai* 外). A nicely representative example is *Analects* 5.27, where Confucius laments “I should just give up! I have yet to meet someone who is able to perceive his own faults and then take himself to task inwardly” (已矣乎吾未見能見其過而內自訟者也). The phrase translated as “to take oneself to task inwardly” (*neizisong* 內自訟) means literally “to internally file a legal complaint against oneself”; the translator Simon Leys sacrifices literal fidelity in order to preserve the metaphorical thrust of this phrase in his translation: “exposing [his faults] in the tribunal of his heart” (Leys 1997: 23). It represents about as strong a sense of psychological interiority as one could wish: within the self is to unfold a metaphorical lawsuit in which one takes oneself to task. A very similar sentiment is expressed in *Analects* 4.17, where the aspiring gentleman is urged to “look within” himself (*neizixing* 內自省):

The Master said, “When you see someone who is worthy, concentrate upon becoming their equal; when you see someone who is unworthy, use this as an opportunity to look within yourself.”

子曰見賢思齊焉見不賢而內自省也

The result of this process of “looking within oneself” is an accurate measure of one’s one state of moral self-cultivation, which in turn can give one confidence in one’s own virtue even in the face of social disapproval or external difficulties. In *Analects* 12.4, the disciple Sima Niu asks Confucius to characterize the gentleman. The Master replies:

“The gentleman is free of anxiety and fear.”

Sima Niu said, “Free of anxiety and fear’ – is that all there is to being a gentleman?”

“If you can look inside yourself and find no faults, what cause is there for anxiety or fear?”

子曰君子不憂不懼曰不憂不懼斯謂之君子已乎子曰內省不疚夫何憂何懼

We see here the idea that introspection gives one access to what we might call the “true self”: what is on the “inside” is the genuine self, which may or may not be reflected on the “outside” of the person.

Indeed, throughout the *Analects* we see a suspicion of the information that can be gleaned from the outside of the self. Confucius was famously dubious, for instance, of the reliability of a person’s verbal assertions.⁹ He often coupled his concerns about the unreliability of words with a suspicion of what could be determined from a person’s countenance or facial expression (*se* 色), literally “colour”, essentially the outside surface of the container of the self. In 1.3, Confucius famously declares that “a clever tongue and beguiling countenance are rarely signs of *ren*”¹⁰ (巧言令色，鮮矣仁). This suspicion of glib speech and superficial appearance is found throughout the *Analects*. The saying found in 1.3 is repeated in 17.7,¹¹ and in 15.11 the danger presented by “glib people” (*ningren* 佞人) is compared to the derangement of morals brought about by the music of Zheng. David Nivison (1999: 751) has made a very interesting observation that may explain Confucius’s hatred for clever, ingratiating people: in archaic Chinese (AC), *ning* was pronounced **nieng*¹² and is actually a graphic modification of its cognate *ren* 仁 (AC **nien*). The original meaning of *ren* was something like “noble in form”, and it would appear that *ning* was its counterpart in the verbal realm: “attractive or noble in speech”. In giving *ning* a negative meaning in the *Analects*, Confucius drives a wedge between the two qualities: *ren* now becomes “true”, that is, inner nobleness or virtue, whereas *ning* represents the false, external counterfeit of *ren*. This is no doubt the sentiment behind such passages as 12.3 (“The Good person is sparing of speech”) and 13.27 (“reticence is close to Goodness”), as well as Confucius’ general suspicion of language and outward show.

We see concerns about hypocrisy explicitly linked to the container metaphor of “inner” in 17.12, where Confucius declares “To assume a severe countenance while being weak inside – is this not, to take an analogy from the common classes, like breaking into a home in order to commit burglary?” (色厲而內荏，譬諸小人，其猶穿窬之盜也與). “Lower classes” is here a rendering of *xiaoren* 小人 (lit. “small people”), more typically translated as “petty person”. Here it is clearly being used in its socio-economic sense in order to make the point that while poverty-struck commoners commit transgressions in order to steal physical objects, the “petty people” among the aristocratic and educated classes, who, being well-off materially, have no need to literally

commit burglary, steal metaphorically, the object of their “burglary” being a good reputation or worldly renown.

This idea of hypocrisy as metaphorical thievery, to “lack the substance but steal the name”, as the Song Dynasty commentator Zhu Xi puts it (1218), also features in the following passage, 17.13, where Confucius cryptically declares that “the village worthy is the thief of virtue.” Probably the best commentary on this passage is by one of Confucius’ Warring States followers, Mencius, found in *Mencius* 7:B:37. Here Mencius quotes 17.13, and then is asked for further explanation by the disciple Wan Zhang:

“What sort of person is this, who is referred to as a ‘village worthy?’”

“He is the type of person who says, ‘Why be so grandly ambitious?’ His words have nothing to do with his actions, and his actions have nothing to do with his words. Such a person then goes on to declare, ‘The ancients, the ancients, why were they so standoffish and cold? When you are born in an age, you should accommodate yourself to it. As long as you do so skilfully, this is acceptable.’ Someone who, in this way, tries to surreptitiously curry favour with his contemporaries – this is the ‘village worthy.’”

“If everyone in a village praises a man as being worthy, and nowhere can you find someone who does not consider him worthy, what did Confucius mean by calling such a person a ‘thief of virtue?’”

“Those who try to censure him can find no basis; those who try to criticize him can find no faults. He follows along with all the vulgar trends and harmonizes with the sordid age. Dwelling in this way he seems dutiful and trustworthy; acting in this way, he seems honest and pure. The multitude are all pleased with him – he is pleased with himself as well – and yet you cannot enter with him into the Way of Yao and Shun. This is why he is called the ‘thief of virtue’. Confucius said, ‘I despise that which seems to be but in fact is not. I despise weeds, for fear they will be mistaken for domesticated sprouts. I despise glibness, for fear it will be mistaken for rightness. I despise cleverness of speech, for fear it will be mistaken for trustworthiness. I despise the tunes of Zheng, for fear they will be mistaken for true music. I despise the colour purple, for fear it will be mistaken for vermilion [*Analects* 17.18]. I despise the village worthy, for fear that he will be mistaken for one who truly possesses virtue.’”

The village worthy is thus one who carefully observes all of the outward practices dictated by convention and so attains a measure of social respect,

but who lacks the *inward* commitment to the Way that characterizes the true Confucian gentleman. Confucius refers to him as the “thief of virtue” because from the *outside* he seems to be a gentleman, and so lays a false claim to virtue. By serving as counterfeit models of virtue for the common people, the village worthy is in effect a false prophet, not only blocking the development of true virtue in himself but also leading others astray.

This issue of potential hypocrisy is a central theme in the *Analects*, and, as in the passages we have examined above, is often linked to the potentially misleading nature of container surfaces (facial expression, mere words, outer behaviour), whereas true virtue is consistently linked with the “inside” of the container self. In his discussion of the *Analects*, Fingarette at times acknowledges this emphasis on sincerity or genuineness, but systematically elides the connection between genuineness and interiority. For instance, in his discussion of 3.12, where Confucius declares that “If I am not fully present at the sacrifice, it is as if I did not sacrifice at all”, Fingarette acknowledges that true ritual only works because “the individuals involved do it with seriousness and sincerity”, concluding that “beautiful and effective ceremony requires the personal ‘presence’ to be fused with learned ceremonial skill” (1972: 8). As P. J. Ivanhoe observes regarding this comment, “The scare quotes around ‘presence’ cannot disguise what [Fingarette] has here admitted. There is a clear reference to an inner self that plays a critical role in ideal ritual interaction” (Ivanhoe 2008: 47). It is thus exceedingly hard to see, even limiting ourselves only to traditional textual analysis, how we might understand the *Analects* without attributing to Confucius a fairly robust sense of psychological interiority.

Perspectives from cognitive science

The importance of metaphor

At one point Fingarette does briefly address the presence in the *Analects* of the metaphor of “inner” (*nei* 內), but quickly dismisses the three occurrences of this word as “vague allusions” entirely lacking in conceptual importance (1972: 46). In fact, as the textual evidence examined above clearly demonstrates, these references to “inwardness” are anything but unimportant, and are deployed together with other similar metaphors throughout the text in a consistent and conceptually important manner. At a much more general level of analysis, the ease with which Fingarette dismisses the specific metaphors used in the text is symptomatic of a broader tendency of philosophers to give short shrift to the importance of metaphors. As Mark Johnson has observed (Johnson 1981), the Western philosophical tradition has long been characterized by a view of meta-

phor as philosophically superfluous: a decorative rhetorical device expressing a thought capable of being fully reduced to some literal equivalent, and therefore merely entertaining at best, and potentially misleading at worst. Scholars of early Chinese thought trained in analytic philosophy departments are typically heirs to this attitude, dismissing the metaphorical specificity of arguments in early Chinese thought in the belief that what really matters is extracting their abstract, logical and propositional essence.¹³

From the perspective of cognitive science, this attitude would seem to be empirically quite ill-advised. There is a growing body of evidence that human thought, far from involving exclusively amodal concepts linked to each other in a propositional manner, is rather primarily image-based and modal in character, that is, deriving its structure from sensory-motor patterns. Among cognitive scientists, this image-based view of human concepts has been perhaps most systematically developed by Lawrence Barsalou and his colleagues, who argue for a “perceptual symbol” account of human cognition. According to this model, the symbols manipulated in human thought are understood, not as pictures, but as “records of neural activation that arises during perception” (Barsalou 1999: 583). These records can be abstracted from and combined in various ways in areas of the brain “upstream” from the sensory-motor cortices, but they always remain to some extent grounded in sensory-motor systems. There is a huge and constantly growing body of evidence in favour of at least some version of the perceptual symbol account,¹⁴ but perhaps the strongest argument in its favour is that it avoids two fundamental problems that plague amodal symbolic accounts: the transduction problem (how perceptual signals could get “translated” into amodal symbols) and the grounding problem (how arbitrary, abstract symbols could ever come to refer to something in the world); and it fits better with what we know about how the brain in general works.

This idea of bodily-based, concrete schemas serving as essential conceptual templates for our understanding of abstract, or less clearly-structured, domains is also the basic insight behind conceptual metaphor theory, which the philosopher Mark Johnson and the linguist George Lakoff have done the most to develop. They were pioneers in formulating a comprehensive and coherent model of cross-domain projection and, most significantly, demonstrating the pervasiveness of these projections in all aspects of human conceptual life.¹⁵ Against theories of metaphor that portray it as a relatively rare and somewhat “deviant” mode of communication thrown in to add rhetorical spice, Lakoff and Johnson argue that “conceptual metaphor” is in fact a ubiquitous and fundamental aspect of human cognition. Conceptual metaphor, as they understand it, involves the recruitment of structure from a concrete or clearly organized domain (the *source* domain) in order to understand and talk about another, usually more abstract or less clearly structured domain (the

target domain). This is the basic conception of metaphor as a cross-domain mapping introduced above, which encompasses similes and analogies as well as metaphors in the more traditional sense.

The most basic of these projective mappings are a set of “primary metaphors,” which are the result of relatively abstract target domains becoming associated with some basic schema source domains (PATH or SCALE, for instance) through experiential correlation. Lakoff and Johnson (1999: 50–54) provide a short list of representative primary metaphors such as AFFECTION IS WARMTH, IMPORTANT IS BIG, MORE IS UP and so on, specifying their sensory-motor source domains and the primary experience correlations that give rise to them. Although they argue that *all* such primary metaphors develop gradually through experiential correlation, it is likely that at least some basic cross-domain associations are the result of fixed synaesthetic cross-wiring,¹⁶ such as the correlation of tones with verticality, or textures such as sharpness with tones or tastes (“E-sharp” or “sharp cheddar”).

However these primary metaphors are developed, all individuals have a huge store of them at their disposal by the time they are able to become productive users of language. These accumulated metaphorical associations then become one of the individual’s primary tools for reasoning about him- or herself and the world, especially when it comes to relatively abstract or unstructured domains, as well as for communicating thoughts to others. While concepts such as “time” or “death” may have a skeleton structure that is represented conceptually in relatively amodal terms, in most cases this amodal structure is not rich or detailed enough to allow us to make useful inferences. Therefore, when we attempt to conceptualize and reason about relatively unstructured realms, this skeleton is fleshed out (usually automatically and unconsciously) with additional structure provided by primary metaphors derived from basic bodily experience, often invoked in combination with other primary metaphors to form complex metaphors or conceptual blends. When primary or complex source domains are activated in such cases and mapped onto the target domain, most aspects of the source domain’s conceptual topology, that is, inference patterns, imagistic reasoning pattern, salient entities and so on, are preserved, thereby importing a high degree of structure into the target domain.

Image schemas and conceptual metaphors have been shown to play a foundational structuring role in everything from basic human categorization and grammatical structures to religious and philosophical discourse, scientific theorizing and legal reasoning.¹⁷ Simple documentation of the pervasiveness and systematicity of conceptual metaphor in human cognition goes a long way toward demonstrating that such schemas play more of a role than as mere figures of speech. In addition to the more general experimental evidence for the imagistic basis for concepts discussed above with regard to the perceptual

symbol theory, there is also now a veritable mountain of linguistic and psychological evidence for the claim that conceptual metaphors in fact represent conceptually active, dynamic, language-independent structures that play an inevitable and fundamental role in embodied human cognition.¹⁸ To be sure, the empirical science of metaphor is still in its infancy and many outstanding problems remain, including how precisely metaphors are instantiated neurobiologically and how they interact with relatively abstract or amodal propositions or conversational intentions. One may also, of course, question the details of specific metaphor analyses, or claims as to the extent to which particular metaphorical entailments are driving a given argument. What is emphatically *not* in doubt, however, is that conceptual metaphors are cognitively real, that is, metaphorical linguistic expressions do activate corresponding image schemas in the sensory-motor regions of the brain, and that these activated schemas play an important role in perception, semantic and syntactic processing, and at least certain sorts of reasoning processes.

This work on metaphor represents one important way in which cognitive science can be of use in the academic study of religion. We do not have direct access to the mind of Confucius or the compilers of the *Analects*. We *do*, however, share with them a common experience of interpersonal struggle, lawsuits and containers, which gives us conceptual access to passages such as 5.27. Another nice Warring States example of this sort of bodily-based access, which also provides a window onto otherwise inaccessible inner psychological experience, is a passage in the Confucian text *Xunzi*, where increasing severities of criticism are conceived of metaphorically as being stabbed with increasingly large weapons: a minor criticism is a “needle”, whereas more serious criticism is a “spear-stab”. We can compare this to such English expressions as “sharp” criticism, “cutting sarcasm”, or gentle “needling”. Here our common physiological responses to being prodded with pointy objects gives us insight into the common psychological pain of enduring criticism from others.

It is this common, embodied experience that can serve as a bridge to the otherwise inaccessible experience of the “Other”, and this bridging function is precisely why we cannot ignore the metaphors employed in texts from other cultures or dismiss them as “vague allusions”. At the same time, the recognition that these experiences are contingent upon bodies and physical environment, that no set of experientially-derived conceptual schemas provides unmediated access to the “things in themselves” and that some degree of cultural variation in schemas is to be expected allows us to avoid the sort of rigid universalism that characterizes Enlightenment-inspired approaches to the study of thought and culture. Ideally, then, conceptual metaphor analysis represents a tool from cognitive science that can give scholars of comparative religion access to a universally shared conceptual grammar, which can then in turn serve as a tool for genuine cross-cultural dialogue.¹⁹

The biological self-container and Theory of Mind (ToM)

It may seem like a rather trivial matter to emphasize, but it must be kept in mind that the early Chinese had *bodies*: they were *Homo sapiens*, a rather unusual but fully embodied species of great ape, and shared with their modern conspecifics a host of basic embodied experiences: ingesting food, expelling waste, coming into contact with other physical bodies and so on. As Antonio Damasio has pointed out, a basic, necessary precondition for *any* form of life is some sort of boundary between inner and outer:

One key to understanding living organisms, from those that are made up of one cell to those that are made up of billions of cells, is the definition of their boundary, the separation between what is *in* and what is *out*. The structure of the organism is inside the boundary and the life of the organism is defined by the maintenance of internal states within the boundary. Singular individuality depends on the boundary. (Damasio 1999: 135–6)

The inner–outer boundary is necessary for physiological homeostasis, that is, assuring that environmental variation does not cause excessive variation within the organism itself. As Damasio notes, biological life simply stops if the profile of the “chemical bath” inside the boundary of the self varies outside very narrow range. He also observes that this necessity of a regulated boundary between inside and outside describes not only the specifications for survival of any organism, but also “some of the biological antecedents of the sense of self – the sense of a single, bounded, living organism bent on maintaining stability to maintain its life” (136). We thus should not be at all surprised to find inner–outer metaphors playing an important role in early Chinese discourse about the self as in our own discourse.

Work in cognitive science also suggests that, when it comes to speculating about what goes on “inside” this container self, the vast majority of human beings will share a set of powerful and automatic intuitions. Cognitive scientists have been arguing for decades for the existence in human beings of a “Theory of Mind” (ToM),²⁰ which causes us to go beyond perceptual data to “paint” mental properties (desires, goals, thoughts) onto the world. It is apparent that, from a very early age, human beings conceive of intentionality as a distinct kind of causality, and distinguish it from both the kind of physical causation that characterizes folk physics and teleological, “vitalistic” causation.

The literature on ToM is vast; the reader is referred to Bloom (2004) for a helpful and quite readable review. Here I will merely note that there is increasingly clear evidence that the tendency to project agency onto the world appears to emerge quite early in development (Spelke *et al.* 1995; Bloom

2004; Phillips & Wellman 2005); has a largely automatic and perceptual component in addition to cognitive components emerging later in development (Scholl & Tremoulet 2000; Tager-Flusberg 2005; Senju *et al.* 2009); is present cross-culturally (Barrett *et al.* 2005; Cohen 2007); is vulnerable to selective and at least partial damage in conditions such as autism (Baron-Cohen 1995; Tager-Flusberg 2005); and would appear to be distributed in human populations in a spectrum ranging from autism (deficient ToM) to schizophrenia (excessive ToM) with a clear genetic basis (Crespi and Badcock 2008; Crespi *et al.* 2009). The accumulation of evidence concerning ToM in human cognition motivates Paul Bloom's famous argument that mind-body dualism is not an accidental philosophical legacy of Plato or Descartes, but rather a universal feature of human "folk" cognition (Bloom 2004).

The fact that seeing other agents as motivated by invisible, interior mental states appears to be an evolved, universal human cognitive default strikes me as a final and fatal bit of evidence against Fingarette's argument. All biological organisms are characterized by boundaries between inner and outer, and humans in particular automatically and effortlessly populate these interiors with psychological entities of various sorts. This being the case, it would really be quite shocking if such concepts did not inform the thought of Confucius of the *Analects*, and even more shocking if he did not even consider them as a possibilities. When we combine this evidence from cognitive science with the textual evidence long available to sinologists that strongly suggests that Confucius *did*, in fact, reason in terms of container logic and psychological interiority, the "no interiority" argument is faced with an insurmountable barrier.

Conclusion

As anyone engaged in the project of studying texts is aware, textual interpretation is not an analytic science: one cannot "prove" that one's interpretation of a text is correct with the same degree of confidence that one can demonstrate a geometrical or logical proof. Although I believe that the passages from Warring States texts that I have cited above are best understood as reflecting a sense of psychological interiority, one could conceivably try systematically to read all of these passages in a "non-interiority" fashion; as, indeed, Fingarette and his defenders attempt to do. In the final analysis, all that one can do when defending a particular line of interpretation is lay out one's textual evidence, add to it whatever extra-textual evidence one feels relevant and let the felt weight of this evidence do its work on one's audience.

In religious studies, the relevant extra-textual evidence has typically consisted of historical or archeological evidence. What I hope to have demonstrated

here is that a rich and powerful new source of evidence, evidence concerning likely human cognitive universals, can also throw its weight onto the hermeneutic scale, often with decisive effect. Moreover, I would also like to argue that, at a broader level, the model of human commonality that arises from an embodied, evolutionary-informed approach to the human self can and should transform the very interpretative landscape in which our hermeneutical debates occur. As I have argued in great detail elsewhere (Slingerland 2008), humanistic inquiry in Western academia has, especially over the last half century or so, been dominated by disembodied models of human cognition. Whether rationalistic and universalist or social constructivist and radically particularistic, these models have been based on the assumption that the basic architecture of human thought arises in a manner completely independent of our evolved, biological embodiment. Such a position is no longer empirically tenable. The human mind is inextricably embodied, and like all embodied minds is the product of evolutionary processes. In the case of humans, these evolutionary processes occur in both biological (genetic) and cultural forms,²¹ but neither one has the effect of magically extracting us from the physical world in which we are embedded. The manner in which a hermeneutic journey unfolds depends very much upon its point of departure. In the academy today, that point of departure is typically the assumption of radical cultural difference, which in turn is based upon a disembodied, culturally or linguistically constructed model of human cognition. Adopting an embodied perspective dramatically alters the point from which we enter a text from another culture, with important implications for the manner in which the hermeneutical process will then subsequently unfold. This represents an important contribution of cognitive science to the academic study of culture, including religious culture.

As we have seen above in the case of the *Analects*, adopting the embodied perspective radically shifts the burden of proof onto those who would deny psychological interiority to Confucius, a burden that, as evidence reviewed suggests, the text cannot bear. This does not mean that early Confucian thought did not differ in important and revealing ways from that of, say, Descartes or Kant; it also does not mean that texts like the *Analects* do not challenge many basic elements of modern Western conceptions of the self, elements that very much deserve to be brought into question. What is *does* mean is that such conceptual variation needs to be contextualized within a framework of basic human cognitive universals. Indeed, it is this very framework that allows texts or thinkers from another era or cultural context to be comprehensible in the first place. It is important to recognize that a fully exoticized “Other” cannot engage us at all, and that the religious or philosophical challenge, the corrective force, of texts such as the *Analects* can only be felt against a background of cognitive universality.

Notes

1. Recently translated into French and reprinted by University of Montréal Press (Fingarette 2004).
2. See especially the work of Roger Ames (e.g. Ames 1991) and Henry Rosemont Jr. (e.g. Rosemont 1991).
3. For arguments concerning the superior empirical adequacy of certain early Chinese conceptions of the self and self-cultivation vis-à-vis Enlightenment models of the self, see Munro (2005) and Slingerland (2011a forthcoming [published?]).
4. I refer to this trend as “reverse Orientalism” because it shares with classic Orientalism a monolithic conception of the “East” as opposed to an equally monolithic “West”, as well as many of the same specific claims about the nature of “Eastern” thought, but reverses the normative evaluation: the East is no longer negatively portrayed as inferior or servile, but rather positively lauded as a “holistic” cure to the social and philosophical ailments of modern Western life. See Slingerland (in preparation) for a more thorough discussion of this trend in modern Western scholarship.
5. In an early response to *The Secular As Sacred*, for instance, Henry Rosemont Jr. criticized Fingarette’s lack of interiority argument for resting on merely negative evidence (Rosemont 1976: 471), and Benjamin Schwartz similarly argued that the lack of lexical equivalents to “subjectivity” or “psychic states” does not mean that such concepts do not play an essential role in the text (Schwartz 1985: 71–5). Also see Ruskola (1992) for an important critique of relevant aspects of Fingarette’s position.
6. In a recent book chapter, for instance, Fingarette notes that translations of the *Analects* have been distorted as a result of the “psychologizing of Confucius, particularly its subjective orientation. We in the West take subjective, ‘Inner’ life so much for granted that reading Confucius this way is quite unselfconscious, and hence all the more prejudicial” (Fingarette 2008).
7. See, for example, Rosemont & Ames (2009) on “role ethics” in early Confucian thought.
8. Traditionally, the *Analects* has been viewed as a coherent and accurate record of the teachings of the Master, recorded during his lifetime or perhaps shortly after his death in approximately 480 BCE, but the current consensus among contemporary scholars is that our received version is a somewhat heterogeneous collection of material from different time periods, assembled by an editor or series of editors, probably considerably after the death of Confucius, but likely completed by the late fifth century or early fourth century BCE.
9. See particularly 5.5, 11.25, 12.3, 13.27, 15.11, 16.4 and 17.18.
10. *Ren* (仁), often translated as “Goodness” or “humanity” is, for Confucius, the highest of the virtues, the “master virtue” of being a proper human being.
11. Cf. 5.5, 11.25, 12.3, 16.4.
12. Generally the modern Mandarin pronunciation of Chinese characters is given, the Mandarin dialect being the standard form of modern spoken Chinese. When relevant, however, it is the practice to provide postulated archaic pronunciation (reconstructed indirectly by historians of phonetics) denoted with an asterisk.
13. See, for instance, Shun (1997: 103–7) or Hutton (2002: 169). Interestingly, in the study of early Chinese thought one also often finds the mirror image of this attitude: the idea that Chinese thought, or East Asian or even Eastern thought more generally, is metaphorical through and through, in a manner that qualitatively distinguishes Eastern thought from logical, literal Western thought. For a critique of this counter-extreme, the reader is referred to Slingerland (2011b forthcoming [published?]).
14. For reviews see the essays collected in Pecher and Zwaan (2005). Another important

- recent statement of the argument for mental images as foundational for human cognition is Kosslyn *et al.* (2006), which also includes a helpful review of the empirical evidence.
15. Lakoff & Johnson (1980, 1999) and Gibbs (2006) provide helpful introductions to conceptual metaphor theory.
 16. For more on the relationship between synaesthesia (the unusual blending of two or more senses) and metaphor, see Slingerland (2008: 156–62).
 17. See Slingerland (2008: 170–72 for extensive references.
 18. For reviews of various convergent lines of linguistic and experimental evidence, see McNeill (1992), Lakoff & Johnson (1999: 81–9), Coulson (2001: 75–83), Rohrer (2005) and Gibbs (2006).
 19. For more on this topic, see Slingerland (2004).
 20. ToM is “theory”-like in that it goes beyond the available data to postulate the existence of unobservable, causal forces or principles. There is a lively debate concerning the appropriateness of the word “theory” when it comes to ToM. Some, such as Gopnik & Wellmann (1994), defend the position that theory of mind *is* a sort of implicit theory. The defenders of the “simulation” position (Gordon 1992; Gallese & Goldman 1998), on the contrary, argue that the achievements of ToM are the result of sensory-motor simulation, relying upon our mirror-neuron system. A third position is carved out by Shaun Gallagher with his claim that ToM is the result of perception-based “body-reading” (Gallagher 2005: 227).
 21. For more on “dual inheritance theory”, see Richerson & Boyd (2005) and Henrich & McElreath (2007).