Xin: Being Trustworthy

Winnie Sung

As published in *International Philosophical Quarterly* September 2020

ABSTRACT: This essay analyses the Confucian conception of *xin*, an attribute that broadly

resembles what we would ordinarily call trustworthiness. More specifically, it provides an

analysis of the psychology of someone who is *xin* and highlights a feature of the Confucian

conception of trustworthiness: the trustworthy person has to ensure that there is a match

between her self-presentation and the way she is. My goal is not to argue against any of the

existing accounts of trustworthiness but to draw on Confucian insights so as to shed light on

features of trustworthiness that are overlooked in current discussions. I hope to show that the

Confucian conception of trustworthiness puts more emphasis on the way a trustworthy

person actively tries to make sure another's dependency on her is not unwarranted than on

how the trustworthy person responds to the one who gives trust.

1. Introduction

So long as we are socially situated, we cannot avoid interacting with other people in some

ways even when these interactions are not of a deep and meaningful kind. The efficiency of our social

interactions and their smooth functioning hinge on our trusting others who interact with us. When we order coffee at a coffee house, we normally trust that the barista will not poison us. We do not demand proof that there is no poison in the coffee. A breakdown of trust in these scenarios would cause disruption and inconvenience.

Although trust is important, we normally do not trust everyone. We trust those whom we think to be trustworthy. If there are good reasons to think that the barista is looking particularly suspicious, we will cease to trust that she will not poison us. And there might be times we are not trustworthy and fail those who trust us in some ways. Trust and trustworthiness come in tandem, and it is difficult to separate the two. Still, some conceptual distinctions between trust and trustworthiness can be drawn. Each of these topics is worthy of investigation in its own right and will have implications for the other. In the existing literature, philosophers tend to focus more on trust, but recently we have begun to see more discussions that focus on trustworthiness.[1]

This essay focuses on trustworthiness and is primarily concerned with the psychology of one who is trustworthy. I will analyze the Confucian conception of *xin*, an attribute that broadly resembles what we would ordinarily call trustworthiness, and I will highlight what I call the *matching requirement*. My goal is not to argue against any of the existing accounts but to draw on Confucian insights to shed light on features of trustworthiness that are overlooked in current discussions. I hope to show that another appropriate way to conceive of trustworthiness is in terms of how the trustworthy person *actively* tries to make sure another's dependency is not unwarranted.

2. Confucian Conception of *Xin*

Confucian thinkers all agree that *xin* is an important ethical attribute that one should have. But Confucian thinkers in various periods of history have tended to have slightly different understandings of *xin*. Without trying to analyze how each Confucian thinker understands *xin*, I will assume that there is some common concern in their understandings of *xin*. In what follows, I will provide a general background discussion of the Confucian conception of *xin*.

In the early texts *xin* can mean "to believe," "what is truly so,"[2] and "straightforward." In the *Analects it* is variously translated as "trust,"[3] "truthfulness,"[4] "trustworthiness,"[5] "making good on one's word,"[6] and "being true to one's word."[7] The complexities of *xin* make translation of the term difficult. Translators often use a variety of English words to translate different usages of *xin* within the same text.

It is said that *xin* is one of the central teachings of Confucius (7.25).[8] Confucius once said that someone without *xin* is like a carriage without a yoke. Such a person will not be capable of getting on (2.22). When it is used as a noun or an adjective, *xin* is mainly used to refer to an attribute of a person. As an ethical attribute, *xin* is closely related to other attributes such as *zhong* (loyalty)[9] and *cheng* (sincerity/wholeheartedness).[10] *Xin* is also often discussed as an ethical attribute that governs interactions between friends (1.4, 1.7, 1.8, 5.26, 9.25). This essay is mainly concerned with *xin* as an attribute of a person.

The character *xin* comprises the radicals *ren* (human being/person) and *yan* (speech/words).

In some passages *xin* is discussed in connection with speech (e.g., 5.10, 15.5, 7.25). On the basis of these textual observations, for example, Cecilia Wee argues that *xin* is primarily concerned with verbal commitments.[11] There are also textual considerations that suggest that the notion of *xin* is more than merely keeping one's verbal commitments. One consideration is that *xin* in the early texts also appears with the term *zhi*, which has the literal meaning of "straight" and the connotations of being straightforward or forthright.

The term *xin* itself is sometimes used in a literal sense to mean straight without a bend. For example, there is a reference in the *Mencius* that one's finger is not *xin*, meaning that one's finger is not straight. This suggests that *xin* also has to do with honesty in a way that goes beyond saying what is true. One's words also need to be true in a straightforward way. There are situations where, even though one tells the truth, one is not being straightforward. For example, when your spouse asks you if you have anything special planned for your anniversary, and you reply, "It's a surprise!" when in fact you have not planned anything. Your answer is not false, but you are also not straightforwardly answering your spouse's question.

More evidence that *xin* is not merely keeping one's verbal commitments comes from *Analects* 19.10:

Only after being *xin* does the superior person work his people hard. If not *xin*, the people would regard him as being harsh on them. Only after being *xin* that the superior admonishes

his lord. If not *xin*, the lord would regard him as being slandered by the superior person.

Kwong-loi Shun notes that this passage is about how the ethically superior person has to represent working the people hard as something that is in the interest of the state. Such representation might take the form of verbal commitments, but it can also take non-verbal forms such as facial expressions and behaviors.[12] This again suggests that the scope of *xin* is not restricted to verbal commitments. One possible reading of *xin* here is that there has to be a correspondence between one's representation of things and facts. We can further infer that *xin* requires one to present herself the way she in fact is.

It is possible that the textual evidence for *xin* in the *Analects* can be interpreted otherwise. Lai, for example, takes *xin* in the *Analects* to mean reliability.[13] Still, the idea that *xin* requires a match between self-presentation and the way one is in fact is clearly emphasized and later developed by Confucians.[14] *Mencius* 7B:25 says that:

> The desirable is called "good". To have it within oneself is called [*xin*]. To possess it fully in
> oneself (*shi*) is called "beautiful" but to shine forth with this full possession is called "great."
> (trans., Lau)

This passage shows a link between *xin* and the possession of the relevant attribute in full, or having the relevant attribute in one's core (*shi*). We may infer that one who is *xin* has within oneself the

relevant attribute that it takes to substantiate the way she presents herself in her interactions with others. In the Neo-Confucian period, it is evident that this way of understanding *xin* that focuses on a match between self-presentation and the way one in fact is has become mainstream. Zhu Xi, for example, explains the above passage from the *Mencius* as saying that a *xin* person is someone who truly has goodness.[15] In light of the above textual observations, I assume a characteristically Confucian feature of *xin* is its emphasis on a match between how one presents oneself and the way one in fact is.[16] We may call this the *matching requirement*.

In the following, I will focus on the psychology that is entailed by the matching requirement. As far as I am aware, even though there are discussions of *xin* that gesture towards the matching requirement,[17] there has not been any detailed analysis of the kind of psychology that meets the matching requirement. I concede that the questions that I raise and the answers that I supply below might not be what the early Confucians themselves were concerned with. My methodology is to construct a philosophical account of being trustworthy that is characteristically Confucian. As will be made vivid, the focus of *xin* is different from the focus of our contemporary understanding of trustworthiness. To avoid confusion, I will use the term *xin* throughout this essay. I will also use some everyday examples in the following analysis. These examples are only meant to illustrate a certain point that I make. Our verdict on real-life cases will depend on the context.

3. The Scope of *Xin*

The matching requirement involves four necessary conditions for *xin*: (1) there must be a side that one presents to others, which we may call the "presenter"; the presentation of the presenter we may call "self-presentation." (2) There must be some person or group capable of receiving the self-presentation in a certain way, which we may call the "potential audience." (3) The presenter is in a position to form reasonable expectations of how this self-presentation is to be received. (4) The presenter is in a position to reasonably expect that this self-presentation will be received positively. I will elaborate on each of the four conditions in turn.

(1) Self-presentation occurs when one makes available to others an expression of oneself. If I watch the sunset with no one around and I say "That's beautiful," I am not presenting myself. I am merely expressing myself. If I watch the sunset with my friend but know that she cannot hear me and say "That's beautiful," I am also not presenting myself. If I am watching the sunset with my friend and know that she can hear me and say "That's beautiful," then I am presenting myself. When I self-present, I am not necessarily aware of the effect my comment has on her. She might come to believe that the sunset is beautiful, believe that I believe that the sunset is beautiful, believe that I like watching sunset, and so on. As a presenter, I only need to be aware that what I said is available to her. It counts as self-presentation as long as the presenter is aware that her words or actions are made available to others. Defined in this way, self-presentation is unavoidable whenever one comes into interaction with others. Hence, *xin* does not apply to people who are not in a position to present themselves regardless what attributes they have, such as a solitary individual in the desert. In this

essay, I use "not-*xin*" to mean that one is in a situation where *xin* applies but fails to be *xin*.

Self-presentation can take the form of speech acts, actions, facial expression, and other forms of bodily comportment, such as posture, looks, gestures, demeanor, tone of voice, and so on.[18] If I keep frowning and saying "tsk, tsk" when my friend tells me how unfairly she is being treated at work, I present myself as agreeing with my friend that she is being treated unfairly at work.

When one presents oneself, one does not necessarily present oneself as having certain attributes. One may also present oneself in a way that suggests one lacks certain attributes.[19] A bilingual teenager, for example, might feel embarrassed by the fact she can speak a certain language. When she is with her monolingual friends, she can present herself as lacking the ability to speak that language. Hence, we may draw a distinction between not presenting oneself as *x* and presenting oneself as not-*x*. One who is *x* but does not present oneself as *x* is not in the scope of *xin*.[20] One who is *x* but presents oneself as not-*x* satisfies condition (1).

Condition (2) is related to condition (1). If one is in a position to self-present, then one must think that there is someone who is in position to receive the presentation. This does not mean that the presenter's self-presentation is actually taken up. Suppose that I announce at dinner with my new colleagues that I am a vegetarian, but no one hears me. Still, I think that my new colleagues are in a position to receive the presentation. Condition (2) only requires the possibility of the presentation being received from the perspective of the presenter.

In cases where a presentation is received, not everyone receives it in the same way. The way in

which it is received depends on a range of factors, from the context to the receiver's expectations. When I say that I am a vegetarian, the reception could go beyond mere reception of the belief that I am a vegetarian. Some of my colleagues might come to regard me as someone who cares about animal welfare, while others might get the impression that I am religious. When we talk about the reception of self-presentation, it is more than just an uptake of the presentation that p is the case. Hence, the presenter has to have some reasonable expectations of how her self-presentation is to be received. This brings us to condition (3).

(3) The presenter is a competent communicator and is able to form reasonable expectations of how her self-presentation is to be received. The expectation has to be reasonable but can be generic, such as expecting others to receive saying "thank you" as saying something positive. This condition rules out cases where my presentation is received in a way that is drastically different from what I can reasonably expect. For example, I do not expect that it will make a difference whether I use black or blue pen to write a memo to an acquaintance. But suppose I use a blue pen to write a memo to my acquaintance who is in the same cultural context as me, but my acquaintance comes to regard my writing a memo in blue ink as saying that I see her as my best friend. She cannot fault me as not-*xin* when she realizes I do not regard her as my best friend, because the way in which she receives my presentation is beyond reasonable expectation. Hence, *xin* is irrelevant for one who is unable to form reasonable expectations, such as a young child, or someone who finds herself with a group of beings who have completely different psychology and/or culture that she cannot comprehend or form

expectations about.

(4) The reasonable expectation of the presenter is one that her self-presentation will be received positively by the audience. By "positive reception," I mean that the presenter thinks that her audience will, as a result of her self-presentation, come to regard or relate to her in a positive way. It could be one in which the audience comes to depend on the presenter in some way. For example, when I apply for a job, I put "Fluent in Mandarin" on my CV. This leads my colleagues to regard me as a Mandarin speaker. Regarding me in this way might never amount to anything in practice. Perhaps there never is a situation at work that requires someone to speak Mandarin. However, should there ever be a situation that needs a Mandarin speaker, they can count on me for being one. The positive interaction can also be one in which the audience makes itself dependable for the presenter. For example, I present myself as a cat lover to a new colleague. On that basis, she becomes fond of me and is always keen to help me out with tasks at work. The positive way in which she relates to me is in part, even if it is a very small part, the result of her seeing me as someone who loves cats. What counts as positive reception does not have to be the audience coming to depend on the one who is *xin*. It counts as positive reception as long as the audience relates to the presenter in a way that she considers good, and such relation is in part the result of the presentation.

What is considered as a positive reception is relative to what would have been the case for how the audience relates to the presenter in a given domain had there been no presentation. Suppose the job does not require fluency in Mandarin, and I know that even if I do not mention that I can

speak Mandarin, the employer will not think negatively of me. They will only have a neutral attitude towards my language ability. However, if I put down Mandarin as my second language, by my lights, they will come to count on me in a way that they would not have had if I had not indicated that I am fluent in Mandarin. And I consider this way of counting on me as something positive. I might in fact be wrong. Perhaps it in fact makes no difference to the audience's reception between reading a CV that indicates that I am fluent in Mandarin and reading a CV that does not indicate that I am fluent in Mandarin. What matters for *xin* is that, from the perspective of the presenter, she expects that her presentation of herself as *x* will be received in a way that will lead the audience to relate to her in some positive way that they would not have if the presentation were absent.

If a situation does not meet any of the above four conditions, then it is a situation that falls outside the scope of *xin* and is a non-*xin* issue. There might be cases where one presents oneself in a way that matches the way one is but does not expect one's audience to receive one's self-presentation positively. Someone who holds racist views might openly present herself as holding those racist views. She does not expect her audience to receive her self-presentation positively but she still presents herself the way she is. Since condition (4) does not obtain, she cannot be considered *xin*, nor can she be faulted for not-*xin*.

Confucius sometimes seems to be talking about *xin* in the more specific sense to mean that one actually has the ethical qualities that one appears to have. Sometimes he talks about *xin* in the more generic sense to mean that one has the qualities that one appears to have. What qualities are

considered as favorable depend on the particular domain one is in. One could be *xin* to, say, a group

of audience that holds problematic moral beliefs. If the person who openly holds racist views expects

that an audience will like the fact that she has racist views, then she is *xin* to her audience. Hence,

being *xin* in a particular domain is not always ethically desirable. This is perhaps why Confucius

cautions us that *xin* could be dangerous sometimes (*Analects* 17.8).[21]

From here on, I use "not-*xin*" to refer to situations that fall outside the scope of *xin*. I will use

the expression "one who is *xin*" to refer to one who is *xin* in a particular domain. This is different

from saying that one is a "*xin* person," which is more generic and cuts across many domains. At this

point, I am only concerned with whether one is *xin* in a particular domain to make an analysis of *xin*

more manageable.

Note that conditions (1) to (4) are only necessary conditions for *xin*. It is unclear to me at this

point if they are jointly sufficient, nor whether they exhaust all necessary conditions. If any of these

conditions does not obtain, then one is not-*xin*. But even if all four conditions obtain, it can still be

unclear if one is *xin* or not-*xin*. Nonetheless, these four necessary conditions should give us

something to work with for now in our attempt to understand the psychology of one who is *xin*.

4. *Xin* and not-*xin*

Having laid down some necessary conditions for *xin*, let us return to the matching

requirement. If one is *xin*, then conditions (1) to (4) obtain and there is a match between one's

self-presentation and the way she is. If conditions (1) to (4) obtain but there is a mismatch between one's self-presentation and the way she is, then one is not-*xin*. Because of conditions (1) to (4), the match that is required by the matching requirement is not just an accidental match. Rather, the match has to be the result of some level of self-awareness and interpersonal awareness.

By "the way one is x," I only mean that one has the corresponding dispositions or states that she presents herself as having. This is not to suggest that there is a stable character, or some essential or real self. It only means that if one presents oneself as having x, one must have x. This is compatible with non-factualist view of self-presentation. It does not have to be the case that there are some preexisting dispositions or states for one to become aware of and then present. It leaves open the possibility that there are some dispositions that cannot be considered independently of one's presentation of those dispositions. Perhaps when I sincerely utter "I believe p," for example, I can bring into the existence the mental state of believing p. When I sincerely utter "I promise you I will go," I bring into the existence my intention of going.[22] The direction of how the relevant state is brought into existence does not concern us here. What matters is that the relevant state or disposition has to be present when one presents oneself as having that state or disposition.

It should also be clarified that there is a distinction between one being x and one's believing that she is x. *Xin* only requires a match between the way one presents oneself and the way one is. One can still be *xin* if one is x even though one does not know that she is *xin*. But if one is not-x but presents herself as x (assuming that something like self-deception is possible), even if she sincerely

believes that she is *x*, she still is not-*xin*. She is presenting herself as *x*, not as she believes that she is *x*. But she in fact is not-*x*. Hence, there is mismatch between the way she presents herself and the way she is. There might also be cases where one is *x* and presents herself as *x*, but believes that she is not-*x*. This is a hard case and I am not sure this person is actually *xin*. It is certain that this person is not trying to be *xin*. If she believes that she is not-*x* and is committed to *xin*, she will present herself as not-*x*. If she believes that she is not-*x* and yet still presents herself as *x*, then she cannot be thinking from her own perspective that she is trying to be *xin*. But such kind of cases are possible.

Imagine a case where a police officer goes undercover to infiltrate a gang. The undercover officer over time, in fact, develops deep loyalty to the gang and presents herself as loyal to the gang; but she believes that she is not loyal to the gang. Let us assume that her belief that she is not loyal is not self-verifying such that in believing she is not loyal she actually becomes disloyal. In this case, it is unclear if she is *xin* to the gang. As mentioned before, I have only said that conditions (1) to (4) and fulfilling the matching requirement are necessary for *xin*. It is unclear if the undercover police office fails some other necessary conditions for *xin*. This will be an interesting case for future investigation and is a test case for whether there are odd cases where one is committed to not-*xin* when in fact she is *xin* in a specific domain.

Here is an example of one who is *xin*. Imagine a grocery shop owner X who supports selling ethically-sourced food. She puts up posters about selling ethically-sourced food in her shop. Her customers gather from these posters that the owner supports selling ethically-sourced food. X is *xin*

because she presents herself as supporting ethically sourced food and she in fact supports ethically sourced food. A not-*xin* grocery shop owner Y, by contrast, puts up posters about supporting selling ethically-sourced food when she in fact does not support selling ethically-sourced food. Even if no customer were aware of the posters that Y puts up, her customers could have easily seen the posters and come to regard Y as supporting selling ethically-sourced food.

There are situations where one is *xin* to one audience but not to another. Suppose that Alex and Ben are best friends, and they both fall in love with Cassi at the same time. Ben knows that Alex loves Cassi. Alex does not know that Ben also loves Cassi. Alex continues to confide in Ben how he feels about Cassi. Since Ben values his friendship with Alex, he decides that he will not pursue Cassi and behaves coldly towards Cassi. In this case, Ben cannot be faulted for being not-*xin* to Cassi because condition (4) does not obtain. Ben expects that his self-presentation will be received negatively by Cassi. Whether or not Cassi actually receives Ben's self-presentation negatively does not affect whether he is *xin*.

Ben, however, is not-*xin* to Alex. Ben reasonably expects that his self-presentation will be received positively by Alex even though he in fact is not the way he presents himself to be. This does not mean that Ben expects that Alex likes the particular way in which he interacts with Cassi. Perhaps he knows that Alex wishes that Ben were friendlier to Cassi. Ben expects that his self-presentation will be received by Alex in a way that keeps Alex relating to him in the intimate way that he is. Recall that "positive" is relative to how Alex would relate to Ben had Ben told Alex that he also loves Cassi. By

Ben's lights, Alex's unreserved sharing of his feelings is in some part the result of Alex's not being

aware that Ben also loves Cassi. If he tells Alex that he also loves Cassi, Alex will no longer

unreservedly confide in him. Ben is not-*xin* to Alex because conditions (1) to (4) obtain, and there is

not a match between the way he presents himself to Alex and the way he is.

To recap, *xin* is not affected by how the audience actually receives the presentation. It

depends on whether there is a match between the way one presents oneself and the way one is, and

whether the presenter is in a position to reasonably expect that her presentation will be received

positively. This suggests that there are two areas that matter to *xin*: one's presentation in relation to

others' expectations and one's presentation in relation to the way one is. One's mind's eye, so to

speak, has to look both inwardly and outwardly. In the following, I will address two potential

worries that may be raised with regard to these two areas.

5. Two Worries

Two potential worries about *xin* may be raised. One worry is that one who is *xin* is overly

concerned with how others view oneself. [23] A second worry is that one who is *xin* is overly

self-conscious. In the following, I will try to assuage these two worries. I agree that *xin* requires

interpersonal awareness and self-awareness, but they do not in any problematic way hinder how we

normally function as social agents.

With regard to the first worry, we need to be clearer about what is meant by "how others

view oneself." If what is meant is whether others like oneself or how others would judge oneself, then one who is *xin* is not concerned with that at all. Self-presentation is not the same as self-promotion. The attention of one who is *xin* is not on how to market herself or present herself in positive light, but on whether she is in fact the way she presents herself to be. One who is not-*xin*, by contrast, is often concerned with others' favorable view of herself rather than whether she is in fact the way she presents herself to be.

Suppose that I know that the job that I apply to prefers candidates who are competent in Korean, but I am not. If I am *xin*, I will not present myself as competent in Korean, even though I know that by presenting myself as competent in Korean I can get the search committee to have a more favorable view of me. But if I am not-*xin*, I will present myself as competent in Korean (assuming that I think I will get away with it) to get the search committee to have a more favorable view of me.

It is right to say that one who is *xin* has to have some reasonable expectation about how her audience will receive the self-presentation and hence be sensitive to others' view. To borrow Erving Goffman's distinction, one who is *xin* is not only aware of the expression that she "gives," that these expressions convey certain information. She is also aware of the expression that "she gives off," that others will take these expressions to be revealing or indicating something about her.[24] However, such sensitivity to others' view does not amount to anything like mind-reading or manipulation. It only requires a level of sensitivity to others' view that does not exceed what is required of competent

communicators in everyday social interactions.

A competent communicator who is capable of communication to others, verbally or non-verbally, should have some reasonable expectations about how her words or actions are going to be taken by the audience. When I say, "I am feeling cold," I should reasonably expect that my audience, who are also competent communicators, will not take me to be saying that I am feeling warm or I am hungry. Some kinds of interaction, like those that involve conversational implicatures, require a higher level of competency. For example, when a friend invites me to watch a film together and I reply, "I would love to see this film. Someone told me that it is really good." Assuming I am a competent communicator, I should reasonably expect that my friend, who is also a competent communicator, will take my reply to imply that I have not seen the film yet. The level of sensitivity required in this case is higher than the case where I am asserting p and expecting audience to take me as asserting that p, but being sensitive to others' view in this way is not problematic.

As Bernard Williams points out, we normally do not demand from a speaker to merely tell the truth. If I tell my colleague "Someone's been opening your mail" when in fact it is me who opens her mail, I should reasonably expect that what I said will mislead my colleague to think that I am not the one who opens her mail.[25] As a normal social agent, I will need to be sensitive to the fact that others are expecting me in "informative communication" not just say sentences that are not false. If we do not think this level of sensitivity is problematic, then what is required of one who is *xin* is just this kind of sensitivity. It requires us to be sensitive to how our self-presentation communicates

informatively.

Even if one tries her best to take into account others' perspectives and present herself as *x*, there can still be evidential underdetermination. One who wants to adopt a child might do her best to present herself as a reliable parent to the adoption agency staff but the staff still think that the evidence is underdetermined. But as mentioned before, whether the audience actually receives the presentation positively does not affect whether the presenter is *xin*. She only needs to present herself in a way that matches with the way she is. So, there can be cases where one is *xin* but is not trusted by the audience.

As for the second worry that one who is *xin* is overly self-conscious, if "self-conscious" means something like one is uncomfortable or nervous with the way she appears to others, then one who is *xin* is not self-conscious in this sense. As discussed above, one who is *xin* is not concerned with whether others like her. If "self-conscious" means higher-order self-monitoring, *xin* also does not require one to be self-conscious when she presents herself. Self-presentation is not the same as self-conscious self-presentation. One can bow to show that she is humbled without being self-conscious of the fact she is bowing to show that she is humbled. There might be situations where one has to self-consciously present herself as *x* even though she is in fact *x*. For example, someone in one cultural context where nodding one's head down means "Yes" might move to a different cultural context where shaking her head side-to-side means "Yes." Here she has to self-consciously shake her head to mean "Yes." Once she has learnt and gotten used to communicating to others that she agrees

with something by shaking her head, she will not need to constantly monitor her presentation at this

higher-order level.[26]

Hence, a competent communicator will not need to be always self-monitoring when she

presents herself. And since *xin* requires one to be a competent communicator, one who is *xin*

probably does not constantly self-monitor her presentation.[27]

That said, this does not mean that one who is committed to being *xin* does not self-monitor

at all. One needs to monitor the way she interacts with others and if she detects any discrepancies

between her self-presentation and the way she is, she should promptly rectify the situation. Suppose

that I was asked if I want to sign an open letter on my way to teach a class and I signed it in a hurry.

After teaching, I reflect on the matter more thoroughly and decide that I do not support the position

of the letter. I will then need to check if there is any way for the organizer to remove my name from

the letter.[28]

One who is committed to *xin* will also need to be self-aware in order to obtain the relevant

self-knowledge. As mentioned in section three, one might lack self-knowledge but still be *xin* as long

as she is in fact *x*. But if she is not *x*, even if she believes that she is *x*, she is still not-*xin*. Assuming that

human beings are fallible and there is a limit to self-awareness, there might be situations that we

sincerely believe that we are *x* and present ourselves as *x*, but we are not *x*, such as someone who

sincerely believe that she is a devout religious practitioner when in fact she is not. To avoid any

disconnect between one's self-presentation and the way one is, one who is committed to *xin* would

need some self-knowledge, at least in the relevant domain. And one crucial way in which the self-knowledge may be obtained is through self-awareness.[29] There might be other ways one can obtain self-knowledge. We might accept what our therapist told us about ourselves. Assuming what the therapist told us is true and we present ourselves as being that way, this will not affect us being *xin*. Nonetheless, there are many states that we are only able to tell others in the first person. For example, when my friend and I walk out of the theater and she asks if I like the play, I will not have the time to consult third-personal evidence as to whether I like the play. I need to be able to tell her, somehow, that whether I like the play.[30]

Hence, in order to be *xin*, we will need some self-monitoring. However, such self-monitoring is not overly directed to self-presentation, but more to one's own dispositions, intentions, abilities, and so on. This kind of self-monitoring does not seem problematic. As a social agent who needs to communicate to others something about ourselves, it is inevitable that we ourselves often have to monitor whether we are *x*.

6. Active Respect for Others

I have identified and analyzed features of the Confucian conception of *xin*. In this section, I will highlight an implication of the Confucian conception of *xin* that can contribute to our contemporary understanding of trustworthiness. I assume that there are some broad similarities between the Confucian conception of *xin* and our contemporary conception of trustworthiness.

Broadly construed, being *xin* is about one's having the relevant attributes such that one is reliable, dependable, or worthy of trust in the relevant domain. Given the affinity between *xin* and contemporary understanding of trustworthiness, an analysis of the psychology of one who is *xin* should shed light on our understanding of what underlies trustworthiness.

Contemporary discussions of trustworthiness tend to center around the trusted person's responsiveness to the trustor's dependency. Russell Hardin, for example, argues that the trustworthy person is someone who does what the trustor wants them to do. Stephen Wright argues that the trustworthy person acknowledges the value of trusting relationships and rationally decides how to act. Faulkner argues that the trustworthy is someone who sees the trustor's depending on her φ-ing as a reason to φ and is moved to φ for this reason.[31] Karen Jones argues that the richly trustworthy person has to take the trustor's counting on her to be a compelling reason for acting as counted on. Common across these accounts of trustworthiness is the view that trustworthiness is primarily a response to another's dependency. The attention of the trustworthy person is on what the trustor wants and on whether she can act in ways on which her trustor can depend.

Confucians might well agree with what these contemporary philosophers have to say about trustworthiness. One who is *xin* is also someone who would respond to the trustor's dependency. But for the Confucians, the focus of *xin* is not so much on whether or how the trustworthy *responds* to another's dependency. The Confucian focus is on the *active* respect one has for others, irrespective of whether they are trustors who trust her or depend on her.

One who is *xin* is sensitive to the fact that others have limited access to who she is, whether it is her mind, her abilities, or her history. Our self-presentation is a main source of information to others about who we are. It is primarily through what is publicly out there, our words, actions, and bodily comportment, that others gather evidence about us and come to form beliefs about us and decide how they want to relate to us. In a way, one who is *xin* is deferring to her audience how they want to relate to her by holding herself responsible for her audience's uptake. Such kind of deference is the reverse of manipulation. Instead of presenting oneself in a way that misleads the audience or takes advantage of a relationship, one who is *xin* tries to ensure that there is no gap between how the audience sees her and the way she is.[32] In taking others' perspectives into account and trying to present oneself as accurately as possible, she is also trying to give others reliable evidence for who she is so that they can make informed judgments and decisions on how they want to relate to her. In this way, one who is *xin* avoids taking advantage of others' limited epistemic access to shape, even sometimes inadvertently, others' view of her and steer how they relate to her in a way that they would not have wanted if her presentation matched the way she really is.

One who is not-*xin*, by contrast, withholds crucial information that she reasonably expects to be received in some negative way by her audience. In withholding that crucial information, one who is not-*xin* single-handedly steers in a direction that she herself welcomes. Hence, one who is *xin* actively preempts any unwarranted dependency by making sure, to the best of her abilities, that her self-presentation gives others reliable evidence about herself. It is not necessarily prominent to the

psychology of one who is *xin* that there is a specific "trustor" who has certain expectations and demands of her.

Sensitivity to others' limited epistemic access alone is not enough. One who is *xin* also has to have respect for others. Without respect, one can be sensitive to others' limited epistemic position and take advantage of it to benefit oneself. One who is hypocritical, for example, is also in conditions (1) to (4) and is sensitive to others' limited epistemic position.[33] But since the goal of the hypocritical person is to get her audience to receive her self-presentation positively, she presents herself in a way that manipulates the audience. One main difference between the one who is *xin* and the hypocritical is that the former has respect for others and does not take advantage of their limited epistemic access. Sometimes, being *xin* would require one to prevent others from thinking well of her. In presenting herself in a way that matches with the way she is, one who is *xin* puts her audience in a good epistemic position to make decisions on how they relate to her.

Another way in which one who is *xin* respects others is that in being *xin*, she does not let herself get off the hook too easily when there is a misunderstanding on the audience's part. For example, when I tell my colleagues that I am a vegetarian, I am also mindful that in the particular context, people normally assume that one is a vegetarian for ethical reasons and come to form more positive views of the vegetarian. If I were *xin*, I should quickly correct that possible positive reception and add that I am a vegetarian for health reasons. Suppose I never clarified that I am vegetarian for health reasons and my colleagues were later surprised to find out that I was a vegetarian for

non-ethical reasons. I could have said something like "I never said I am a vegetarian for ethical reasons" or "It never came up." By leaving things vague, I leave some ways to get off the hook. But if I am in a position to reasonably expect that others will assume that I am a vegetarian for ethical reasons, I have to clarify that I am not a vegetarian for ethical reasons to preempt others relating to me in a way they would not have had they learnt that I am not vegetarian for ethical reasons.

All these suggest that one who is *xin* assumes an active respectful stance towards others. In contrast to contemporary accounts, it is not that the audience has already invested their trust and then one who is *xin* sees it as a reason to act in a way that can be counted on. Before the audience even assume the role of "trustor," one tries to make sure that she is someone who warrants the trust. Part of it is to make sure she has the relevant disposition and mental state, and part of it is to make sure she presents herself as having the relevant disposition and mental state so that the audience can be in a reliable epistemic position to make up their own mind whether to place their trust in her. In this way, one who is *xin* also reduces the amount of interpersonal work her audience has to do when engaging with her.[34]

One may question the relevance of *xin* to contemporary societies. Being *xin* hinges on one's ability to form reasonable expectations about how one's presentation is going to be received by others, which is largely fixed by social and cultural norms.[35] In a society that has clear social and cultural norms, it is much easier for one to form expectations and make inferences. The early Confucians, for example, are in a society that is governed by a set of well-defined and stable ritual

practices, norms, and codes of behaviors. In that kind of society, even wearing a certain color could be a clear indication to others what state one is in. However, such clear norms and expectations are dissolving in contemporary societies. Wearing a maroon outfit to a wedding no longer signals anything significant. Wearing a ring on one's ring finger does not necessarily signal that one is married. Aside from the ambiguities of the more general cultural norms and expectations, we often encounter situations where the relevant norms in a specific domain are ambiguous. For example, I have been using the example that "I am a vegetarian" throughout this essay. It is not clear to me what the norm governing the practice of giving examples in first-person pronoun in academic papers is. Do readers usually take examples given in first-person pronoun as autobiographical or hypothetical? Should I clarify at some point of this essay that I am in fact not a vegetarian?

Contemporary societies indeed present various challenges to the cultivation of *xin*. In societies where norms and expectations are more ambiguous and unstable, it is more difficult for one to form expectations about how her presentation is going to be received. These ambiguities could dis-incentivize *xin*. One might have lower sensitivity to others' perspective or mistaken expectations about others' expectations. The ambiguous circumstances could also make it more tempting for one to take advantage of others' limited epistemic access and leave things vague so that she can get off the hook easier in the future, even when she is not trying to manipulate others. And in societies that value pluralism and diversity, it is quite likely that *xin* will lose some of its desirability as a trait, replaced instead by traits such as tolerance and acceptance.[36]

26

However, as long as we are social members who have to interact with others, *xin* will still be relevant. For example, *xin* will be relevant to the practice of giving testimonies. In the foreseeable future, we still have to depend on testimonial exchanges to learn about the world. There are a vast number of things with which we do not have direct acquaintance with. *Xin* is relevant as long as others' direct access to information about us is limited, and that we have to tell others things. *Xin* requires a speaker to tell her audience things in a way that respects and does not take advantage of the audience's epistemic dependence on her. Richard Moran, for example, argues that when a speaker tells her audience that *p*, the speaker assumes responsibility for what her audience believes.[37] The Confucians conceive being *xin* as one way in which the speaker assumes responsibility.

Even though we might have relatively fewer well-defined and stable norms and expectations now compared to Confucius's time, there still exist enabling social conditions for *xin*. Some implicatures and conventions will still be in place to enable and govern social interactions. One who is *xin* can also offset the ambiguities in some other ways, such as being more attentive to the particularities of her audience or relying more on explicit speech acts rather than behavioral signals in her self-presentation.

7. Conclusion

In this essay I have analyzed a distinctive feature of the Confucian conception of *xin*. If my analysis is correct, then one way in which the Confucian conception of one who is *xin* differs from

the contemporary understanding of the trustworthy person is its emphasis on how the trustworthy person shows respect to her audience by actively trying to make sure that their trust will not be unwarranted. This does not affect the point that one who is *xin* values trusting relationships in general, or values a particular trusting relationship, or is held responsible by some social mechanisms such as contracts and sanctions. But these features do not take the center stage of the Confucian conception of *xin*.

The key aspect that concerns early Confucians has to do with a match between the way one is and the way one presents oneself. This in turn requires us to be sensitive to our audience's epistemic position and show respect for others, by giving them reliable evidence about ourselves and then allowing them to decide, in an informed manner, how they want to relate to us.[38]

---

[1] Cf. Russell Hardin, "Trustworthiness," *Ethics* 107 (1996): 26-42; Nancy Nyquist Potter, *How Can I be Trusted? A Virtue Theory of Trustworthiness* (Lanham MD: Rowman & Littlefield, 2002); Stephen Wright, "Trust and Trustworthiness," *Philosophia* 38 (2010): 615-27; Karen Jones, "Trustworthiness," *Ethics* 123 (2012): 61-85.

[2] Kwong-loi Shun, "Wholeness in Confucian Thought: Zhu Xi on Cheng, Zhong, Xin, and Jing" in *The Imperative of Understanding: Chinese Philosophy, Comparative Philosophy, and Onto-Hermeneutics*, ed. On-cho Ng (New York NY: Global Scholarly Publication, 2008), p. 267.

[3] For example, D. C. Lau, trans., *Confucius: The Analects* (Hong Kong: The Chinese Univ. Press, 2002); Peimin Ni, *Understanding the* Analects *of Confucius: A New Translation of* Lunyu *with Annotations* (Albany NY: SUNY Press, 2017).

[4] For instance, James Legge, trans., *The Chinese Classics: with a Translation, Critical and Exegetical Notes, Prolegomena, and Copious Indexes,* Vol. 1: *Confucian Analects, The Great Learning, and The Doctrine of the Mean* (Taipei: Southern Materials Center Publishing Inc., 1985).

[5] E.g., Lau 2002; Ni 2017; Edward Slingerland, *Analects: With Selections from Traditional Commentaries* (Indianapolis IN: Hackett, 2003).

[6] E.g., Roger T. Ames and Henry Rosemont Jr., *The Analects of Confucius: A Philosophical Translation* (New York NY: The Random House Publishing Group, 1998).

[7] E.g., Arthur Waley, *The Analects of Confucius* (New York NY: Book-of-the-Month Club, 1992).

[8] The other things are: culture (*wen*), moral conduct (*xing*), and doing one＇s best/loyalty (*zhong*).

[9] E.g., *Analects* 1.4, 1.8. 5.28, 7.25, 9.25, 12.10, 15.6. See Winnie Sung, ＂*Zhong* in the *Analects*: with Insights into Loyalty＂ in *Confucianism for a Changing World Cultural Order*, ed. Roger Ames and Peter Hershock (Honolulu HI: Univ, of Hawaii Press, 2017), pp. 175-96 for discussion of early Confucian conception of *zhong* (loyalty).

[10] E.g., *Mencius* 4A:12 and 5A:2 D. C. Lau, trans., *Mencius* (London UK: Penguin Books, 2003). The link between *xin* and *cheng* become more conspicuous in the *Liji* and even more so in Zhu Xi＇s commentaries on the early texts.

[11] Cecilia Wee, ＂*Xin*, Trust, and Confucius＇s Ethics,＂ *Philosophy East and West* 61 (2011): 516-33.

[12] Shun, p. 267.

[13] Karyn L. Lai, ＂Learning to be Reliable: Confucius＇s *Analects*＂ in *Cultivating a Good Life in Early Chinese and Ancient Greek Philosophy*, ed. Karyn Lai, Rick Benitez, Hyun Jin Kim (London UK: Bloomsbury, 2018).

[14] I am grateful to Kwong-loi Shun for helping me think through this point.

[15] See Zhu Xi's commentary on *Mencius* 7B:25 in *Mengzi jizhu*, *Sikuquanshu* edition.

[16] There could well be other aspects of *xin* that are also distinctive to Confucian thought such as reliability that I do not have the space to explore here.

[17] E.g., David Hall and Roger Ames, *Thinking Through Confucius* (Albany NY: SUNY Press, 1987); Shun 2008; Ni 2017.

[18] Hagop Sarkissian, "Situationism, Manipulation, and Objective Self-Awareness," *Ethical Theory and Moral Practice* 20 (2017): 489-503. Sarkissian makes a similar point about how "one's *presence* (including not just one's clothing but one's posture, tone, and volume of voice, countenance, demeanor, and other factors that are the focus of others' attention and part of others' perceptual field) is taken to be a source of influence on one's immediate environment, and therefore a site of scrutiny" (Sarkissian, pp. 496-7). In the article Sarkissian is making a more general point about how one's self-presentation can influence others and may be partly responsible for others' behavior. I agree with Sarkissian on this. Here, I am making a more specific point that one has to be mindful of whether there is a match between one's self-presentation and the way one is. Thanks to Ian James Kidd alerting me to this point.

[19] Thanks to Nicolas Bommarito for helpful comments on this and for suggesting a similar example.

[20] This does not mean that a person of this sort cannot be faulted on some other grounds.

[21] I assume here that there are cases where a failure to be *xin* does not mean that an agent fails to do what she ought to do. There could be cases that an agent is morally required to be not-*xin*.

[22] Cf. Jane Heal, "On First-Person Authority," *Proceedings of the Aristotelian Society* 102 (2001): 1-19 for a non-factualist account along this line.

[23] I am grateful to Sinyee Chan for raising this point.

[24] Erving Goffman, *The Presentation of Self in Everyday Life* (New York NY: Anchor, 1959), p. 15.

[25] Bernard Williams, *Truth and Truthfulness* (Princeton NJ: Princeton Univ. Press), pp. 96-107.

[26] See Amy Olberding, "Etiquette: A Confucian Contribution to Moral Philosophy," *Ethics* 126 (2016): 422-46 for further discussion on the role of etiquette in helping one develop habits that can do away with self-conscious monitoring but remain attentive to others.

[27] In this regard, a not-*xin* person might also present herself as *x* without self-consciously presenting herself as *x*.

[28] It is likely in this sense that the Confucians said that one has to reflect daily whether one has been *xin* in one's dealing with friends (*Analects* 1.4).

[29] Here, I will bracket the question as to how self-awareness and self-knowledge are possible.

[30] This also suggests a way in which self-awareness and self-knowledge are important to good social relations.

[31] Hardin 1996; Wright 2010; Paul Faulkner, *Knowledge on Trust* (Oxford UK: Oxford Univ. Press, 2011), p. 147; Jones 2012.

[32] In this respect the self-presentation of the *xin* person is different from the kind of control that Goffman discussed.

[33] See Winnie Sung, "*Xiang Yuan* (Village Worthies): The Appearance-only Hypocrite," *Dao: A Journal of Comparative Philosophy* 15 (2016): 175-92 for an analysis of early Confucian conception of hypocrisy.

[34] Ian James Kidd made this point clearer for me.

[35] I am grateful to Ian James Kidd and Graham Parkes for this point.

[36] I am grateful to Ian James Kidd for raising this point and helping me articulate it.