

Weakness of Will, the Background, and Chinese Thought

Chris Fraser and Kai-Yee Wong

Department of Philosophy

Chinese University of Hong Kong

October 2005

## Weakness of Will, the Background, and Chinese Thought

Chris Fraser and Kai-Yee Wong

### Abstract

This essay applies John Searle's account of weakness of will to explore the classical Chinese problem of weak-willed action. Searle's discussion focuses on the shortcomings of the Western classical model of rationality in explaining weakness of will, so he naturally says little about the practical ethical problem of overcoming weak-willed action, the focus of the relevant Chinese texts. Yet his theory of action, specifically his notion of the Background, suggests a compelling approach to the practical issue, one that converges with a plausible account of the classical Chinese conception of agency. On this approach, the practical problem is due to weaknesses of the self in carrying out intentions. The key to overcoming the problem lies not in restructuring the agent's affective states, as suggested by prominent interpreters of Chinese thought such as David Nivison, but in strengthening the agent's Background capacities, much as we do when mastering new skills.

## Weakness of Will, the Background, and Chinese Thought

Chris Fraser and Kai-Yee Wong

### I. Introduction

The following is so common a kind of action failure that it is almost an everyday problem for many of us: Having formed a firm and unconditional intention to do  $y$ , after judging that it would be better on the whole to do so, a person ends up doing  $x$ , which is incompatible with  $y$ . Things like New Year's resolutions not kept and overdue papers yet to be written come to mind. Most of us will agree that sometimes we knowingly and intentionally perform an action counter to our best judgment, either moral or prudential. This phenomenon is commonly known as "weakness of will," or "akrasia," and such actions are commonly called "weak-willed," "incontinent," or "akratic." Weakness of will has traditionally been regarded as a kind of practical irrationality, and thus its nature and possibility have been considered a subject demanding philosophical explanation. However, John Searle and Nomy Arpaly have recently argued that weak-willed action is not characteristically irrational.<sup>1</sup> Searle's argument is very much simply a blunt denial based on the blatantly obvious: if we were to deem irrational every person who acts in weak-willed way, there would hardly be a practically rational person left in the land. On his diagnosis, the traditional idea that weakness of will is problematic for rational agency arises from a mistaken conception of the causal relation between reason and action. According to Arpaly, on the other hand, our belief about what would be rational for us to do—our best judgment—has a special normative force only when it is warranted by our

other beliefs and desires, insofar as the latter beliefs are responsive to evidence. So, to claim that acting against one's best judgment is always less rational than following it, one must rule out cases in which one's best judgment is not warranted by one's evidence—that is, cases in which one's supposedly best judgment is in fact irrational. But then the claim becomes the uninteresting one that acting against one's *rational*, best judgment is always less rational than following it. On Arpaly's diagnosis, the view that acting against one's judgment is necessarily irrational is an unfortunate result of the mistaken idea that deliberation is necessary for rationality. There clearly are, Arpaly argues, cases of rationality without deliberation.

Searle's critique, and indirectly Arpaly's, challenge the assumption that the possibility of weakness of will is something that requires special philosophical demonstration. This assumption is central to the contemporary Western conception of the problem of weakness of will. It has given rise to many ingenious arguments purporting to show that apparent cases of acting against one's unconditional intention are not genuine cases of weakness of will. In contrast, such a notion hardly has a place in the Chinese conception of weak-willed action. One of the aims of this paper is to develop this and other contrasts between recent Western philosophical approaches to weakness of will and the classical Chinese approach. In our view, some aspects of Searle's and Arpaly's critiques (especially the former) of common Western conceptions of the problem dovetail with the Chinese orientation. They thus may be useful in understanding the nature of the Chinese problem and the differences between it and familiar Western approaches.

In this paper, we will focus on Searle's discussion of weakness of will and bring it

to bear on central aspects of the traditional Chinese conception of weak-willed action. Searle's discussion concerns recent proposals concerning how to explain the phenomenon of weakness of will, and so he says little about the practical ethical problem of how an agent can *avoid* weak-willed actions. Yet his attack on the conventional conception of the problem can be taken as an invitation to shift emphasis from the *explanatory* to the *practical* issue, i.e., to an aspect of weakness of will that in recent Western thought has been overshadowed by the puzzle regarding how weakness of will is possible. In this sense, Searle's view indirectly encourages an investigation of weak-willed action along the lines of what we see as the Chinese approach, which focuses on the practical problem of overcoming weakness of will, rather than the theoretical problem of explaining it. As David Nivison's work insightfully suggests,<sup>2</sup> the practical problem of how to overcome weakness of will can fruitfully be treated as a central theme of classical Chinese moral psychology. However, Nivison's interpretation of early Chinese psychology tends to take for granted what Searle calls the Western classical model of rationality. He assumes that the problem with an agent who fails to act as she thinks she should lies in a motivational defect. Hence he sees the key to preventing weak-willed actions as lying in ethical cultivation aimed at restructuring the agent's affective states. One of our aims in the discussion below is to highlight the weaknesses in Nivison's account.

Searle's theory of action, specifically his notion of the Background, suggests a more compelling way of overcoming weakness of the will, one that converges in some respects with what we argue is a more defensible account of the classical Chinese conception of agency. In the final part of this paper, we suggest that a plausible way for

Searle to address the practical problem of weakness of will would be to show how weaknesses of the self in carrying out intentions can be ameliorated by strengthening capacities included in the Background, much as the development of such capacities enables us to master skills.

## II. The Classical Model of Rationality and the Gap

Let us briefly review Searle's analysis of the weakness of will by way of his criticism of the now-familiar Davidsonian account.<sup>3</sup> According to this account, we can characterize an akratic or weak-willed action thus:

- (A) Sometimes an agent makes a judgment that it would be better to do  $x$  than  $y$ , believes herself free to do either, and then intentionally does  $y$ .

The problem of weakness of will is that (A) seems to be incompatible with familiar and plausible principles concerning practical reasoning.<sup>4</sup>

If (A) is considered to describe *pure* cases of weak-willed actions, then a common approach to the problem of weakness of will in twentieth-century analytic philosophy, of which Davidson's account is an example, has been to deny that such pure cases exist. Davidson contends that actions that we think of as exhibiting weakness of will are not really instances of (A). Instead, they involve a *conditional*, rather than an unconditional, judgment by the agent that it would be best for her to do a certain action. So, to correctly describe the sorts of weak-willed actions that in fact commonly occur, the word 'judgment' in (A) should be qualified by "conditional" or "prima facie." So qualified, (A) no longer contradicts any familiar principle concerning practical reasoning.

On Searle's analysis, Davidson's solution to the problem is a consequence of a mistaken conception of the relation between reasons and actions. In Searle's view, Davidson assumes a central tenet of what Searle calls the "classical model" of rationality and practical reasoning (2001, pp. 5ff.), according to which intentional states such as beliefs and desires are causally sufficient to determine rationally motivated actions. Constrained by this idea, Davidson's account inevitably must claim that, when an agent acts contrary to her intention or best judgment, she did not really hold the unconditional intention or value judgment in question. In Searle's view, the Davidsonian approach to weakness of will is a kind of *reductio ad absurdum* of the classical model.<sup>5</sup> This model—and this approach to weakness of will—entails that there cannot be a *gap* between the psychological antecedents of an action and its intentional performance (or the intentional attempt to perform it). Yet the existence of such a gap, Searle argues, is absolutely essential to rational decision making and action.

The traditional puzzle about weakness of will is thus a result of a mistaken conception of intentional causation, according to which the process of deliberation is causally sufficient for the formation of an intention to act and such an intention is causally sufficient for the actual undertaking of the intended action.<sup>6</sup> By contrast, according to Searle, in voluntary, non-compulsive action, because of an inescapable causal gap between different stages of deliberation and action, the mere presence of reasons or an intention to act does not compel the agent to act (2001, p. 50). Rather, the agent must *act on* the reason or intention, making it *effective* as a cause of the action.<sup>7</sup> This "acting on a reason" corresponds to actually trying to do what we have made up our mind to do while

exercising freedom of the will.<sup>8</sup>

Weakness of will, therefore, presents no puzzle, since it is but a symptom of freedom of the self. Actions are not just events that follow from certain psychological antecedents by causal necessity, but *performances* by agents presented with an indefinite range of different reasons or options for action. Weakness of will arises as a natural consequence of the gap between the psychological antecedents and the performance of an action. This brings us to Searle's three-part conception of the gap.

### III. The Gap

Discussions of the problem of weakness of will focus almost invariably on cases where an agent is weak-willed in executing an intention to act according to his own best judgment. But reflection on our experience as agents seems to show that in the structure of deliberation and action there are also other junctures at which weakness of will may arise. One of the strengths of Searle's account of intentional action is how it contributes to articulating these other sorts of cases of weakness of will. In Searle's view, we do not normally experience the process of deliberating, making up our minds, and acting as a sequence of stages each of which is causally sufficient for the next to occur (2001, p. 50). Rather, we experience a continuous causal "gap," manifested in different ways at different stages of the process, by which each stage could fail to lead on to the next. Searle suggests that the overall experience of the gap divides fairly naturally into three distinctly identifiable gaps:

- (1) The first gap obtains between an agent's deliberation on her reasons for deciding



to act one way or another and her resulting decision—the formation of an intention to act, or what Searle calls a “prior intention.”

- (2) The second obtains between the prior intention and the intention-in-action, between deciding to do something and actually trying to do it.
- (3) The third holds between the initiation of an action and its continuation to completion, and thus it is present at every stage in the execution of a temporally extended intention-in-action. More precisely, as Searle puts it, this is the gap between “the causes in the form of the prior intention to perform the action and the intention-in-action on the one hand, and the actual carrying out of the complex activity to its completion, on the other” (2001, p. 63).

The problem of weakness of will, as traditionally conceived, relates mainly to the second gap, the failure to execute prior intentions. But Searle’s idea that there are at least three gaps, if correct, suggests that weakness of will need not be confined to this failure. An agent could also manifest weakness of will by failing to form the prior intention most justified by the reasons she considers in making up her mind, or she could manifest it in carrying a complex action to completion. The focus of the traditional problem of action failure as arising from the second gap<sup>9</sup> is but a result of a widespread tendency toward internalism about reasons. Searle’s notion of the gap thus provides several useful distinctions to employ in characterizing different aspects of—or, more precisely, different locations of—weakness of will. Traditional accounts have paid little attention to the location problem.

Searle’s three-part conception of the gap also opens up the interesting possibility

that the importance or plausibility of an account of weakness of will may depend on which gap one is talking about. Since, in Searle's view, the gap is closely tied to the experience of freedom, one might suggest that the three kinds of weakness of will can be only as different as the ways an agent may exercise freedom with respect to the three kinds of gaps, and this difference cannot amount to much, since the freedom in each case is essentially one and the same phenomenon. But this does not mean that very different solutions might not be applicable to the ethical and prudential issue of how one is to overcome weakness of will in the three sorts of cases. How one deals with a failure to carry out a complex activity might be very different from how one deals with a failure to form the best-justified prior intention as a result of a process of deliberation or practical reasoning.

Searle's critique of the classical model and his thesis of the gap will prove useful both in understanding the nature of the Chinese problem of weakness of will and the differences between it and the traditional Western problem. We will now briefly sketch the Chinese problem and then relate it to Searle's views.

#### IV. Weakness of Will: The Chinese Problem

Among the central issues of classical or early Chinese philosophy—the thought of the pre-Qin or Warring States era<sup>10</sup>—is a problem of moral weakness related to but distinct from the traditional Western problem of weakness of will. There are two major differences between the traditional Western philosophical problem and the Chinese problem.

First, where the Western philosophical problem mainly concerns *explaining* weakness of will—that is, understanding how it is possible—Chinese philosophers were

mainly concerned with *overcoming* weakness of will—that is, preventing it from occurring. The Western philosophical problem is one of theory; the Chinese problem one of practice.<sup>11</sup> Moreover, in the Western formulation of the problem, weakness of will occurs when an agent fails to act according to his best judgment. The failure is sometimes regarded, on the classical model, as one of rationality. The Chinese problem is how an agent can overcome moral weakness or character weakness so that he adheres more reliably to the right ethical path. Thus it seems more appropriate to label the Chinese version a problem of *moral weakness* or *character weakness*, involving a failure of reliability in performance, rather than weakness of will, involving a failure of rationality.

Relabeling the topic “character weakness” helps bring out the second way the Chinese problem of action failure is different from the Western problem. The unit of action at stake in Chinese discussions is the ethical path, not the single action, and the problem is a failure of character, not a failure to carry out an intention. Western philosophers have tended to focus on reasoning about individual actions and to treat weakness of will as a problem concerning the logical relation between practical reasoning and the agent’s actions. The challenge is to resist temptation or failure on an act-by-act basis. Weakness of will is seen as revealing a failure to execute an intention, rather than a failure to consistently adhere to the right path in pursuing ethical goals.

By contrast, the central issue in pre-Qin discussions of action failure concerns agents who know the right ethical path and desire to follow it but fail from a lack of resolve, nerve, or moral fiber. It is epitomized by the following exchange in the *Analects*, a collection of sayings attributed to and anecdotes about Confucius.

Ran Qiu said, “It’s not that I don’t delight in the master’s Way (*dào*, path), it’s that my strength is insufficient.” Confucius said, “Those whose strength is insufficient collapse along the way (*dao*). Now you draw a line.”

We will note just two points about this passage. First, here, as throughout pre-Qin thought, the unit of action at issue is the practice of an ethical path or “way” (*dào*)—a set of practices, habits, and styles of conduct—rather than the execution of a discrete individual action. Second, the problem in question is Ran Qiu’s failure to practice the Way because of his belief that he lacks the needed ability or “strength.”<sup>12</sup> For this, Confucius rebukes him for giving up in advance, using rhetoric that would sound familiar coming from a sports coach: Trying and failing is one thing, but placing limits on one’s abilities before even giving the task an honest try shows a shameful lack of gumption.

As this passage illustrates, moral weakness, in the Chinese context, is a failure to stick to the path—to consistently and reliably do the sorts of things one knows one should do. Pre-Qin thinkers generally agreed that this sort of failure is to be rectified by training and habituation. So when discussing the practical problem of moral weakness, they tended to talk about how agents can develop virtues leading them to reliably act correctly, or more precisely, to adhere to the right path. Following a path is matter of developing certain abilities and habitual patterns of behavior, which add up to developing a certain sort of virtuous character. Thus Chinese thinkers were concerned mainly with the performance of skills and habits and the issue of practical training aimed at cultivating skills, habits, and virtues.

Searle’s thesis of the gap is useful in characterizing the difference between the

Chinese and Western problems of weakness of character or will. We noted that the traditional Western problem has concerned the second gap, the failure to act on one's prior intention. By contrast, Chinese thinkers' main concern is with something like the third gap, that between the agent's intentions and the actual performance of a complex, temporally extended activity. We can think of early Chinese thinkers as concerned with two interlaced practical problems, that of maintaining the concrete practice of the Way over time, and that of continuing the process of character development and training that is both part of and strengthens our ability to practice the Way.<sup>13</sup>

Searle's account of the gap should also help to underscore the importance of the Chinese version of the problem of action failure as compared with the Western version. Recall that in Searle's view, instead of something strange and puzzling, weakness of will is a natural consequence of the gap, whose existence is essential for the operation of free, rational decision making and voluntary action. To the extent one is not skeptical about freedom of the will, the possibility of weakness of will presents no particular explanatory problem. Moreover, as Searle reminds us repeatedly, a natural consequence of existence of the gap and human freedom is that no matter how you structure the antecedents of your action, weakness of will is always possible. What then calls for a solution, one might say, is not the *explanatory* problem but the *practical* problem of how to cope with or overcome weakness of will. In understanding the problem of moral weakness as fundamentally a practical problem, early Chinese thinkers may have rightly drawn our attention to where it should be.

## V. Nivison on Weakness of Will in Early Chinese Thought

In several essays collected in *The Ways of Confucianism*,<sup>14</sup> David Nivison usefully suggests that the *Analects*, *Mencius*, *Mozi*, and other early Chinese texts can fruitfully be approached as addressing the problem of how to overcome weakness of will. Interestingly, Nivison proposes that the Chinese problem is not *akrasia* so much as *acedia*, a sort of apathy or lack of motivation to do what one knows is right (1996, p. 93). This proposal already hints that his interpretation of the Chinese version of the problem presupposes the classical model of rationality.<sup>15</sup> For he identifies *acedia* as a condition in which the agent does not “care enough about it to act” (p. 92) on her practical judgment, and he ties moral weakness to “the problem of my ability to feel the way I would have to be genuinely moved to do” what a normative argument tells me I should (p. 96). As Nivison frames it in one place, the problem is that “the motivational component is not something that just automatically is there when one sees the connection” between the premises and conclusion of a practical argument (p. 145). So he sees the Chinese paradigm of moral weakness, or action failure, as a case in which an agent correctly judges what she should do but fails to do it because she either has insufficient motivation or has not applied her existing motivation correctly. On his view, the Chinese solution, exemplified by his interpretation of *Mencius*, is for the agent to refocus and restructure her emotions or desires so that they coincide with her normative judgments (see, e.g., pp. 99, 102, 144). This is achieved through a process of self-cultivation in which the agent “extends” natural affective responses, such as the love we feel for kin, to new objects, such as distant strangers. Then the agent will “care enough” and action will ensue.

As this solution suggests, Nivison assumes that motivation, or at least morally worthy motivation, can arise only from desires or emotions, and this assumption structures his way of framing the Chinese problem and his Mencian solution.<sup>16</sup>

Thus Nivison assumes motivational internalism and something like the classical belief-desire model of action, in that he assumes motivational force must come from feelings or desires. He sees the problem of moral weakness as in effect a failure to carry out the result of a piece of practical reasoning, though in the Chinese context the reasoning in question is analogical, not syllogistic (pp. 96ff.). The agent sees the case at hand as analogous to a case in which a certain action is performed, but lacks the motivation to perform the analogous action. Nivison's solution is to strengthen the agent's affective response to the situation, so that the agent indeed responds to it in a fashion appropriately analogous to some paradigm case, such as the treatment of beloved family members. The agent shifts the needed motivation into place by analogically extending virtuous reactions from paradigmatic cases to other, relevantly similar cases.

Nivison conceives of two sorts of situations in which the agent might modify her motivations in this way. We can think of these as synchronic and diachronic. In synchronic cases, we know what to do and have a disposition to be moved to do it, but are unable or unwilling to apply that disposition (p. 89). Here the solution is to learn to perform an "inner act of thought" (p. 85) that brings the disposition into play.<sup>17</sup> For instance, for Nivison's Mencius, we have natural pro-attitudes toward the good and the right, and the key to overcoming moral weakness is to focus thought on these properties to bring the motivational power of our affective response to them into play (pp. 85, 113). Diachronic

cases receive most of Nivison's attention, because they are central to his understanding of the development of virtue.<sup>18</sup> These are cases in which the project of bringing our motivations into line with our normative knowledge requires an extended, indeed lifelong, process of self-cultivation (p. 85). Gradual, unselfconscious cultivation is needed, because just forcing oneself to act would not solve the problem, since the only reliable solution to moral weakness is for action to be driven by a spontaneous, "lively, animated" affective response.

Nivison takes Mohist consequentialism—and, astonishingly, any ethical view that appeals to normative principles to guide action—to face a serious problem of moral motivation, because the Mohists do not place the cultivation of affective dispositions at the center of their moral psychology and moral reform program. Instead, they believe that normative arguments, social encouragement, and a system of material incentives and disincentives will be enough to motivate most people to abide by moral norms and eventually to become good. Nivison takes the exchange between Mencius and a wayward Mohist named Yí Zhī (*Mencius* 3A:5), who failed to observe Mohist doctrines by giving his parents lavish funerals instead of modest ones, to illustrate the pitfalls of a purported Mohist doctrine of moral cultivation. Nivison imputes to Yí Zhī the view—not found anywhere in the *Mozi*—that agents can be led to practice the Mohist doctrine of all-inclusive care by redirecting and expanding their natural concern for kin so that it becomes universal and impartial (pp. 102ff.).<sup>19</sup> But, as Nivison sees it, the Mohist approach is condemned to failure, because it pushes agents to perform, for normative ethical reasons, actions that their affective dispositions in themselves would not otherwise lead them to



perform. The purportedly superior Mencian solution is an internalist one, on which actions spring from suitably cultivated affective responses, for on Nivison's interpretation this is the only reliable way to prevent action failure due to moral weakness.

Searle's attack on the classical model and his account of weakness of will help highlight the flaws in Nivison's account in two ways. They remind us that ancient Chinese thinkers could have been applying a theoretical framework other than the classical model, and they cast doubt on the plausibility of Nivison's Mencian solution to moral weakness. For on Searle's account, no matter what one does to strengthen the affective component of one's reasons for action, weakness of will remains an ever-present possibility. A plausible practical solution will not come from modifying the premises of the practical reasoning that leads to the intention to perform an action. The solution must come from whatever it is in the agent that translates intentions into actions. And that is the Background.

## VI. Resolving Moral Weakness

We suggest that Nivison misconstrues the nature of the problem of moral weakness for Chinese thinkers. They do not focus on individual acts, and they do not think the solution to the problem is to cultivate and refocus one's feelings and desires, bringing them into harmony with one's judgments. Rather, they emphasize that everyone has the ability or capacity to develop a virtuous character, just as every normal human being has the capacity to learn to read, speak a foreign language, swim, or ride a horse. Having this capacity means that we already possess sufficient motivational resources to enable us to become virtuous. But to develop and apply our abilities reliably, we need habituation and training.

Initially, the process of habituation and training may help us to recognize the right ethical Way, if we have not already. But generally it functions simultaneously as a component of our practice of the Way and as a means of strengthening that practice.

Having set out and made some progress in following the Way, we may experience two sorts of moral weakness: isolated instances of failure to act, which are a familiar, common form of weakness, and a comprehensive collapse in our practice of the Way, a rare but severe form of failure. Isolated instances of weakness are akin to errors in the performance of skills, whether the failure of a novice still acquiring a skill or an expert having a bad day. In this sort of case, we have what Searle calls “a failure in ‘how to do things’,” due fundamentally not to a problem with the agent’s beliefs and desires, but to “a breakdown in the functioning of the pre-intentional capacities that underlie the intentional states in question” (1983, p. 155). The solution to such isolated breakdowns or instances of weakness would seem to lie in continued training to develop and perfect one’s moral dispositions and skills. Comprehensive weakness is akin to abandoning a training program, course of study, avocation, or career path. Here again the failure seems not to lie in the content of the agent’s intentional states, nor in the rational relation between them and the initial intention in action that got the agent started on the path. What is missing is rather a kind of perseverance, commitment, or determination—that is, whatever quality or qualities of character move us across the gap from the intentional, psychological antecedents of action to its performance. (We hope this suggestion will strike a chord with anyone who has ever had difficulty quitting smoking, losing weight, or sticking with a rigorous training program.) If so, then the solution to this sort of weakness may also lie in a sort of

habituation and training to strengthen the relevant aspects of the agent's character.

Thus if we have correctly sketched some of the structural elements of early Chinese philosophy of action, the Chinese solution to the problem of moral weakness lies in continued training and practice aimed at developing both particular virtues, such as kindness, and broader character traits, such as perseverance or resolve, which lend coherence to the self and enable us to carry out temporally extended intentions. The immediate aim of such training is not to modify one's affective responses to things, but simply to develop reliable habits and dispositions. (Very likely, these will involve certain sorts of affective responses, but such responses are not the crucial feature.) The success of the training process depends less on the status of the agent's conative or affective states at any one stage than on the agent's raw determination to stick with the program.

If this interpretive hypothesis about early Chinese thought is correct, then we should expect to find the texts frequently expressing the view that everyone has the ability to be virtuous, if only we simply "get out there and do it," and that though we will inevitably fail sometimes, when we err we must simply get back on track and continue practicing until the Way becomes second nature to us. We suggest—though for brevity we will not marshal texts to support this claim—that Confucius, Mozi, and Mencius are frequently depicted as expressing, directly or indirectly, roughly just this view, and that in the dialogue with King Xuan (*Mencius* 1A:7), Mencius is portrayed as stating the first half of it quite explicitly.

## VII. Modifying the Background and a Strengthened Self

We suggest that Searle's theory of action, in particular his thesis of the "Background," has the resources to provide a complementary solution. The thesis of the Background is Searle's account of the role of non-intentional or pre-intentional capacities in intentionality; the "Background" refers to the various non-intentional capacities, abilities, and know-how that enable intentional states to function. (Searle capitalizes the word to indicate that he uses it as a technical term.) We can summarize the thesis as the claim that all intentional phenomena—meaning, understanding, belief, desire, experience, action, and so forth—function only within a set of non-intentional capacities that play an indispensable role in determining their intentional content, and thus their status as intentional phenomena.<sup>20</sup> In and of itself, for instance, an utterance is merely a pattern of sound waves, meaning nothing in particular. Utterances have meaning only in some context of use. But any such context will involve a range of non-intentional, causal capacities of the speaker and audience, from brute perceptual capacities to the know-how or abilities that enable us to participate in complex social practices. These capacities are aspects of the Background that enable particular utterances, or for that matter, intentional states, to have the meaning or content they do.

We suggest that the complementary solution Searle's theory can provide to the practical problem of moral or character weakness, and thus weakness of will, is an account of how habituation and training can modify and strengthen the agent's Background, a significant portion of which can be considered the basis for, and perhaps even partly the locus of, the agent's character or self. This training would seek to develop virtuous

dispositions and habits that function automatically, without requiring intentional control or guidance, once the agent has commenced an intentional course of action. Though the patterns of activity issuing from these dispositions and habits can to some extent be monitored and controlled by conscious reasoning and judgment, generally they are not. Instead, virtuous dispositions and habits are allowed to operate immediately and spontaneously. These dispositions will include the sort of rule-responsive dispositions that allow us to act in accordance with the rules of social institutions without actually employing representations of the rules consciously or unconsciously to guide our action.<sup>21</sup> Such Background dispositions constitute a central part of the underlying action-guidance mechanism of an agent with a virtuous character.

On the picture of moral agency we have in mind, a well-trained agent follows the Way consistently and reliably. For him, commitment to his ongoing practice of the Way—and thus his continuing ethical training and cultivation—is so much a part of his character that turning off the path is not a real option. Paradoxically, the more advanced the agent's training becomes, the more restricted his sense of the possibilities or options open to him, yet the greater the degree of his control over his actions, and thus the greater his moral freedom. One might wonder how this restriction of “real” options could avoid diminishing the agent's freedom, because to be free is usually understood as a matter of being presented with a wide range of possible courses of action. But consider the example of a skilled professional basketball player who spots a chance for a shot. His response will be so automatic that he can experience only a very restricted range of options as open to him. Obviously, this does not reduce his shot to an involuntary, robotic action, however.

These observations suggest that modifying and strengthening the Background may be an instructive way to understand the character training needed to overcome moral weakness, for the Background is simply the underlying set of pre-intentional and non-intentional capacities that enable action and other intentional phenomena to function. A virtuous act is an intentional action that flows from the agent's virtuous character. We can think of such acts as morally appropriate, "skilled" responses to particular situations that issue from reliable dispositions developed as a result of a long-term ethical "training program." We suggest that the dispositions in question reliably generate morally skilled responses precisely because they are part of a pre-intentional or non-intentional Background that functions largely spontaneously, without intentional guidance or control.

The notion of the Background also helps to clarify how an agent can experience an apparent restriction of options without a reduction in freedom. As free agents, we are accustomed to thinking of ourselves as presented with a wide range of possibilities for action. Nevertheless, this range is always limited by various factors, including our physical environment, biological capacities, and Background (2001, p. 25). Like our biological capacities, the Background sets genuine limits on what we can freely do. But far from reducing our freedom, the Background is what enables us to act in the first place. It functions to open up, not close off, possibilities for action, and indeed it is partly constitutive of our character and our very capacity for agency. The training an amateur basketball player receives that transforms him into a professional does not turn him into a robot.<sup>22</sup> Just the opposite: It enhances the reliability and skillfulness of his Background capacities, giving him the ability to perform new sorts of actions and forging his identity as

a basketball player. Similarly, the training and habituation a moral agent might undertake to modifying her Background abilities and strengthen her character may narrow her range of real options for action, but it does not make her robotic. Rather, it is a process of strengthening the self, and the agent is likely to experience the concomitant restriction of “live” options not as a limitation but as strength of character.

The example of the basketball player also helps to bring out a related point about the role of training in cultivating moral fortitude. Just as an expert player needs to maintain his training (if only to retain his already perfected skills), part of being a virtuous agent is maintaining a commitment to continuing moral “practice.” One may have already traveled far along the Way, but the gap, as an inescapable feature of genuine agency, is always present. Sticking to the path requires training; continuing to stick to the path requires sticking to further training. Moral weakness is displayed not only by straying from the Way or by failing to develop our native ability to do what is right into a virtuous character. It is displayed also by a lack of commitment to the ongoing training needed to keep oneself on the path. Developing our abilities is the key to overcoming isolated instances of action failure; the continuing commitment to training and practice is the key to preventing us from abandoning the Way outright.

These considerations point toward a solution to moral weakness that focuses not on intentional states, reasoning, and deliberation, but on training the agent’s ability to perceive and respond to moral situations in the right way—the way that an agent with a virtuous character perceives and responds to them.

To avoid moral weakness, such perception and responsiveness must of course be

highly reliable. This brings us to two further functions of the Background that Searle identifies, which fill out our explanation of how habituation and training can strengthen character and overcome moral weakness. The Background facilitates certain kinds of readiness, preparing us for the sorts of situations we are likely to encounter, and it disposes us to respond to those situations in certain ways (1995, p. 136). Two aspects of this sort of readiness and response are particularly important here. First, a major function of the Background is to enable perceptual interpretation to take place; it is by bringing Background skills to bear on raw perceptual stimuli that we apply categories to things perceived (1995, p. 132). Clearly, what goes for perception will also go for the agent's ability to recognize types of moral situations. By modifying our Background abilities, we can enhance the immediacy and reliability of our moral perception, a key component of a virtuous character.

Second, in Searle's view, the Background can be "causally sensitive" to the rules of institutions "without actually containing any beliefs or desires or representations of those rules" (1995, p. 141). Consider again the example of the basketball player. A player learning how to play the game may initially need to learn a set of explicit rules and strategies and deliberately guide his play by them. But as he becomes skilled, he acquires a set of dispositions that enable him to respond smoothly and directly to the demands of the situation, so that he moves appropriately yet automatically, without actually thinking about the rules and strategies. Searle's theory suggests that it is a mischaracterization to say that at this point the player has learned to apply these rules and strategies more skillfully. Rather, he is no longer actually applying them at all. Instead, he has developed



Background skills and abilities that are “functionally equivalent to the system of rules, without actually containing any representations or internalization of those rules” (1995, p. 142). This role of Background causation in explaining rule-responsive behavior helps to explain one aspect of the virtuous agent’s moral strength—the agent’s capacity to reliably respond in a morally “skillful” way to the ethical demands of particular situations. As with the exercise of any skill, activity governed by ethical norms reaches its highest level of reliability and mastery when it is integrated into the Background and becomes the sort of activity we can perform automatically, without conscious thought. We suggest that what Searle contends goes for language, games, and social practices goes for morality as well.<sup>23</sup>

To sum up, the Background—and thus the agent’s strength of character and moral fortitude—can be modified through training and practice in such a way that the agent becomes increasingly disposed to reliably carry out his intentions and, in moral contexts, to act virtuously. Searle’s notion of the Background includes many of the dispositions, capacities, and other factors that constitute a virtuous character. Recognizing the function of the Background and the role of reliable dispositions within it highlights a central feature of the Chinese way of thinking about weak-willed action. Fundamentally, the weakness in question is not simply a failure of rationality or in the execution of intentions, so much as a failure of the Background mechanisms that normally enable the agent to consistently adhere to a chosen course of action, moral or prudential.

<sup>1</sup> John Searle, *Rationality in Action* (Cambridge, Ma.: MIT Press, 2001) and Nomy Arpaly, *Unprincipled Virtue* (Oxford: Oxford University Press, 2003).

<sup>2</sup> David S. Nivison (1996), *The Ways of Confucianism*, La Salle, Ill: Open Court.

<sup>3</sup> Davidson, arguing that that weakness of will is explicable and intelligible, disregards the moral aspects of the problem and treats it simply as an issue concerning practical reasoning.

But Searle considers the general idea behind Davidson's account to be the same as that behind accounts such as R. M. Hare's for resolving *akrasia* in moral contexts. Thus, Searle thinks, his criticisms of the shared general idea apply to both. See Searle (2001, pp. 221ff.)

<sup>4</sup> Davidson mentions two such principles: (i) If an agent wants to do  $x$  more than she wants to do  $y$  and she believes herself free to do either  $x$  or  $y$ , then she will intentionally do  $x$  if she does either  $x$  or  $y$  intentionally; and (ii) If an agent judges that it would be better to do  $x$  than to do  $y$ , then she wants to do  $x$  more than she wants to do  $y$ . See Donald Davidson, "How is Weakness of Will Possible?" in Davidson (2001), *Essays on Actions and Events*, Oxford: Oxford University Press, p. 23. Davidson's article was first published in 1969. See also Searle (2001, p. 221f.).

<sup>5</sup> Granting that Searle is right about the Hare-Davidson approach, there is still a question about just how far back the Hare-Davidson conception of the puzzle of weakness of will goes in the Western philosophical tradition. Aristotle, for example, does not seem to share this conception. Clearly, whether Searle's critique of the approach constitutes a (partial) refutation of the classical model of rationality and practical reasoning depends a great deal on how one answers this question. In this paper, we will not concern ourselves with this

historical and interpretive issue, however.

<sup>6</sup> According to Arpaly, deliberation is not a necessary condition for acting for reasons, either. See Arpaly (2003), Chapter 2.

<sup>7</sup> Here we wish to raise a question concerning the causal role of what Searle calls “effective reasons.” A central assumption of the classical model that Searle attacks is that rational actions are caused by beliefs and desires. A prominent idea in his critique is that since “cause” in this assumption has the sense of the common or Aristotelian “efficient cause” or “sufficient cause,” the assumption in effect fails to provide a model of rationality. For it is only in cases of compulsive, irrational actions, such as an addict’s use of drugs, that the psychological antecedents of an action are really causally sufficient to bring about the action. In the case of rational action, the existence of the gap entails that the agent’s beliefs and desires (for instance) are not casually sufficient to cause her action. Given the thesis of the gap, then, Searle must offer a new explanation of the basic causal connection between intentions and actions.

His approach is to replace the notion of a reason or an intention as a sufficient condition or an efficient cause with the notion of a reason “made effective” by the agent (2001, p. 85). But it is not completely clear exactly what kind of causation is in play when an effective reason functions as the cause of an agent’s action. For on the one hand, Searle distinguishes between the notion of “causing” and that of “performing” an action (p. 157), suggesting that the proper notion to apply in explaining action, at least at the level of intentional phenomena, is that the self *performs* actions—nothing “causes” them, in the

sense of a sufficient cause. Nothing “fills” the gap. (This of course leaves open the possibility of an explanation in terms of sufficient causes at the neurophysiological level.) On the other hand, he also suggests that an appeal to such a reason to explain intentional action is a case of non-sufficient causal explanation (p. 83). This seems to imply that a reason made effective by the agent thereby causes the agent’s action, through a form of intentional causation in which an intentional state causes the very state of affairs it represents (p. 41). And though Searle emphasizes that a reason cannot be an efficient cause or a sufficient condition, he characterizes intentional causation as a type of mental causation, which is a subcategory of efficient causation (p. 41). This leaves us unclear as to the precise causal role of effective reasons.

<sup>8</sup> In terms of the three gaps identified in §III below, we can say that the existence of the first gap rebuts principle (ii) in Davison’s argument (see note 3) and the existence of the second gap rebuts (i).

<sup>9</sup> Of course, as Searle sees it, the conventional conception of the problem arises not from the second gap, strictly speaking, but from a failure even to recognize the existence of the gap.

<sup>10</sup> “Classical,” “early,” “pre-Qin,” and “Warring States” are four labels for the same historical period, which can be conveniently demarcated as lasting from the death of Confucius in 479 B.C. to the founding of the Qin dynasty in 221 B.C.

<sup>11</sup> This is not to suggest that the practical problem has been entirely neglected in the

Western tradition. But it seems to have been addressed mainly under the rubric of religion, not philosophy.

<sup>12</sup> Notice that Ran Qiu is portrayed as “delighting” in the Way, an attitude that can plausibly be taken to imply that he sincerely believes this Way is ethically right and desires to practice it himself. He thus has the psychological antecedents normally taken on the Western classical model to be sufficient for him to act.

<sup>13</sup> The committed attitude required to succeed in this project is surely part of the Confucian virtue of *chéng*.

<sup>14</sup> See David Nivison (1996), *The Ways of Confucianism*, La Salle, Ill: Open Court, in particular the essays “Weakness of Will in Ancient Chinese Philosophy,” “Motivation and Moral Action in Mencius,” “Philosophical Voluntarism in Fourth-Century China,” and “Two Roots or One?”

<sup>15</sup> Specifically, Nivison’s approach is likely to have been influenced by his colleague Donald Davidson’s treatment of *akrasia*.

<sup>16</sup> The emphasis on affective responses seems to be something Nivison brings to the texts, motivated probably by his construal of internalism, for it is hard to find it in the texts themselves. Early Confucian texts do insist that the proper performance of rituals requires the right sort of attitude or feeling, and the *Mencius* links the capacity for moral goodness to the capacity to feel alarm at danger to others (2A:6) and to feel love for one’s parents (7A:15). But there is little evidence that feelings were considered essential to moral motivation or moral worth. Nivison (1996, p. 99) cites Mencius 4B:19, which distinguishes

between “acting from kindness and duty”—that is, acting from virtue—and “putting kindness and duty into practice”—acting according to the norms of kindness and duty, but without yet having fully developed the virtues oneself (this is a plausible interpretation, though the classical text is vague). But it is question-begging to suggest that acting from kindness and duty involves experiencing any particular emotions, rather than simply having a reliable disposition to act in the right way.

<sup>17</sup> One might object to Nivison’s use of the term ‘disposition’ here, since the term usually refers to a reliable response, not one that needs to be intentionally invoked. Perhaps ‘capacity’ would be a more appropriate label for what Nivison is getting at.

<sup>18</sup> This is due to a nuance in Nivison’s view arising from his version of internalism, which leads him to run together the problem of moral weakness with a particular conception of moral worth. He believes actions are morally worthy only when they issue from certain sorts of feelings, such as compassion for others; that these feelings are the product of “cultivating” and “extending” our natural feelings for those close to us; and that this process of extension normally takes time. Hence in his view merely pushing oneself to act in the right way cannot solve the problem of moral weakness, since the goal is action that is not merely right, but morally worthy. We believe that Nivison’s views do not explain the texts well—the *Mencius*, for example, says little or nothing about the diachronic cultivation of emotions and seems to adopt more or less just the views Nivison proposes are being criticized there—but the interpretive arguments needed to explore these issues thoroughly would take us beyond the scope of this paper.

<sup>19</sup> The Yí Zhī story is almost certainly unrelated to moral cultivation, since neither Yí Zhī's explanation nor Mencius's response allude to anything like a cultivation process.

(Moreover, it is odd to offer an interpretation of Mohist doctrine on the basis of few words from a failed Mohist reported in the text of one of the Mohists' opponents.) Yí Zhī tries to excuse his failure to conform to Mohist doctrine by saying that "care has no degrees, but practice begins with one's family." The point of this remark in the context is probably that—consistent with Mohist doctrine—he has an equal degree of moral concern for all, but that—again, consistent with Mohist doctrine—in practice he takes care of his family's needs first. The implication is that he would give everyone's parents a lavish funeral if he could. The excuse is mysterious, since it fails to explain his action, which violated the Mohist doctrines of frugality and modest funerals, not the doctrine of all-inclusive moral care. But Yí Zhī's remark has nothing to do with a cultivation or reform regimen. Mencius responds with an imagined tale about the origin of funerals in people's spontaneous disgust and horror at seeing insects and animals eat their dead parents. Again, nothing resembling a cultivation process is depicted.

<sup>20</sup> Searle's most careful, precise statement of the thesis is: "All conscious intentionality—all thought, perception, understanding, etc.—determines conditions of satisfaction only relative to a set of capacities that are not and could not be part of that very conscious state. The actual content by itself is insufficient to determine the conditions of satisfaction." See John R. Searle (1992), *The Rediscovery of the Mind*, Cambridge, Ma.: M.I.T. Press, p. 189). For an earlier version, see Searle (1983), *Intentionality*, Cambridge: Cambridge University

Press, p. 143. Searle's most informative accounts of the Background are at his (1992, pp. 175–96) and his (1995), *The Construction of Social Reality*, New York: The Free Press, pp. 129–47.

<sup>21</sup> We are referring here to the sorts of dispositions Searle describes in his account of rule-sensitive or rule-conforming behavior that issues directly from the Background, and thus is no longer dependent on intentional rule-following. Searle's thesis is that we can develop behavioral dispositions that enable us to act *in conformity to* rules without actually *following* them. See Searle (1995, pp. 137–47).

<sup>22</sup> One reason for this is that the relevant locus of freedom in extended patterns of activity, such as pursuing a training program, lies in the voluntary, intentional commitment to a complex, long-term task or goal. Since much of what we do that makes human life human lies in the pursuit of long-term projects, such commitment, and the apparent "restriction" of freedom that comes with it, is characteristic of human agency.

<sup>23</sup> This point is particularly germane in the context of ancient Confucian thought, in which the virtuoso, context-sensitive performance of an elaborate system of ritual actions was considered a central feature of the virtuous agent's activity.



Index of Chinese Characters

*Chéng* 誠

*dào* 道

*Yí Zhī* 夷之

## REPLY TO KAI-YEE WONG AND CHRIS FRASER

*John R. Searle*

I thought the paper by Kai-ye Wong and Chris Fraser was fascinating and insightful. Two things I especially appreciated are the clarity with which they summarize my views. I think they are quite fair and accurate. Second, I appreciate their suggestion that the way to deal with the practical problem of weakness of will has much to do with the role of the Background in shaping our actions. I think they are especially on the right track when they say that the improvement of Background skills may actually narrow the range of real options for action, (p. 21) nonetheless, they do not decrease freedom. As they say, "It is a process of strengthening the self, and the agent is likely to experience the concomitant restriction of 'live' options not as a limitation but as strength of character." (p. 21). That seems to me very much on the right track. What they are suggesting, and it is a powerful addition to my own writings, is that we should not just think of the Background as facilitating languages, games and social practices generally, but for morality as well (p. 23).

The special philosophical puzzle about weakness of will has to do with a specific theory of human action and a specific conception of rationality that goes back to Aristotle. As they are aware, worries about the very possibility of weakness of will are the preserve of professional philosophers. These philosophers are confronted with a paradox: the prevailing theory of action is called a "causal theory" because it says an intentional action is defined in terms of its causes. A typical version of the story goes: actions are caused by reasons, but reasons for actions are beliefs and desires, but then an intentional action is an event caused "in the right way" by beliefs and desires. If the beliefs and desires set sufficient causal conditions, how is it possible that there could be such a thing as an incontinent or akratic action? If by definition actions are caused by reasons, and if a person has the right reasons and recognizes them as the right reasons for the right action, then how can he fail to act on those reasons? How can weakness of will be possible? But it is quite obvious that weakness of the will is not only possible, it is actual. It is very common in real life. And the fact that the standard theories

make it seem impossible or at least bizarre shows a weakness of those theories. Weakness of the will is a natural consequence of the existence of what I call the “gap.” There is a causal gap between reflecting on the reasons for action and actually making up one’s mind on the basis of those reasons; there is a continuation of this gap between making up one’s mind and actually initiating the action that one has decided to do; and in the case of actions extended over time there is a continuing gap between one’s commitment to the course of action and one’s actually carrying it out through to the conclusion.

When we face this problem, we realize that there is a distinction between what worries professional philosophers and what worries educators. The point I wish to make now is that the real life educators in the West have very much the same problem that philosophers in the Chinese tradition have. The problem is not to explain the possibility of weakness of will but to enable us to overcome it. I think it may be misleading to characterize the distinction between Western and Chinese philosophy in dealing with the problem of weakness of will as a difference between the *theoretical* problem of how weakness of the will is possible and the *practical* problem of overcoming weakness of will, especially in moral situations. The reason I think this is misleading is that of course the problem of overcoming weakness of will is very much part of the traditional moral education in Western culture, as it is in other cultures. In my youth, what religious people and school educators worried about is very much what the Confucians worried about: Strength of Character, Learning, Fortitude, Self-Control. All of this is part of the stock in trade of Sunday school teachers and school principals. Perhaps the school principals gave up. But in the United States the standard aim of moral education was to produce self-control. We were always taught that you have to learn strength of character and self-control. The effort to induce self-control in the young seldom worked, but it was traditionally the objective of much moral education.

Now, it is true that there are some things missing in the Western tradition. One is anything at all similar to the Confucian rituals. Another thing missing is the idea that there is “the way.” But with those qualifications, I think the educational philosophy and the religious philosophy in Western countries, is much like the way the authors characterize the Chinese problem.

I think it is an implausible and indeed unrealistic feature of technical philosophy in English speaking countries that weakness of will is made

out to be something bizarre or unusual. The idea is that it is puzzling that such a thing could ever happen. The idea that weakness of will is some weird anomaly, as Wong and Fraser are right to see, results from a failure to see the importance of what I call the gap. I want to say, in opposition to the prevailing Western tradition, that if you think weakness of the will is a remarkable problem, you have a mistaken theory of action, because weakness of the will is very common in real life.

Weakness of will is as common as tea in China. I think most ordinary people have cases of weakness of will several times a day. Perhaps they have not followed the Confucian disciplinary training well enough, but all the same, weakness of the will is not uncommon. For example, one thinks, "I should not drink another glass of wine, but the Cabernet tastes very good, so I will have a little bit more." This sort of thing occurs to me often.

Davidson's solution to the problem of weakness of the will, seems to me, a solution by fiat. He says that if the causes of the action in the form of reasons are satisfactory, it is impossible that weakness of will should occur. But it obviously does occur. Therefore, he stipulates that in any case where it occurs, there must have been something wrong with the antecedents of the action. The formation of the intention wasn't an unconditional all-out intention, but only a *prima facie* conditional intention. But that is simply a solution by fiat—I want to say: whatever the form of the intention, however strong and unconditional you make the intention, you can still have weakness of will and still have akrasia. The only way to avoid this is to make it a tautology that unless the psychological states in the form of reasons caused the action, then there was something inadequate or conditional about the psychological states.

I very much appreciate their suggestion that the practical problem of overcoming the weakness of the will is in large part a matter of developing Background abilities, and as they are right to see, this has theoretical implications both for our philosophy of action and for our philosophy of morality.