

DAVIDSON ON ACTIONS: REASON, RATIONALITY, AND IRRATIONALITY

By

© 2014

Huei-Rong Li

Submitted to the graduate degree program in Philosophy
and the Graduate Faculty of the University of Kansas

in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chairperson John Bricke, Ph.D.

Ann Cudd, Ph.D.

Erin Frykholm, Ph.D.

Thomas Tuozzo, Ph.D.

Geraldo Sousa, Ph.D.

Date Defended: December 8, 2014

The Dissertation Committee for Huei-Rong Li
certifies that this is the approved version of the following dissertation:

DAVIDSON ON ACTIONS: REASON, RATIONALITY, AND IRRATIONALITY

Chairperson John Bricke, Ph.D.

Date approved: December 8, 2014

ABSTRACT

Donald Davidson's analysis of weakness of will explains the possibility of weakness of will and the cause of weakness of will. Davidson shows the compatibility between the existence of incontinent actions and the principle that an agent always acts on what he judges to be better through describing an incontinent agent as an agent who fails to detach an all-out, unconditional judgment from his all-things-considered judgment, but infers an all-out, unconditional judgment from another competing *prima facie*, conditional judgment with an insufficient reason. Davidson identifies the strong desire causing an incontinent agent to act as the cause of his incontinence. Such a mental cause bears a non-logical causal relation with its effect and brings about an inner inconsistency within an agent. The theory of partitioned mind explains how inner inconsistency is possible under the assumption that no propositional attitude can exist without causally and logically connecting to other propositional attitudes.

My dissertation seeks to explain and defend Davidson's analysis of weakness of will. Chapter I introduces Davidson's causal theory of action. Chapter II explains how weakness of will is compatible with principles or assumptions accepted by the causal theory of action. Davidson describes incontinence as a failure in reasoning, and in Chapter III I argue that incontinence as a failure in reasoning is possible because there is no implication between an all-things-considered judgment and an all-out, unconditional judgment, that the mental cause which is responsible for such a failure could be multiple, and that probabilistic akrasia as a case of failure in reasoning parallel to incontinent action is possible. Chapter IV investigates Davidson's taking incontinence as an inner inconsistency. I argue that the explanatory force of the theory of partitioned mind lies in the separation of two inconsistent propositional attitudes, that the theory

of partitioned mind is necessary to account for inner inconsistency, and that inner consistency can be a standard of rationality. Chapter V is a conclusion of this dissertation.

TABLE OF CONTENTS

I. BASIC FRAMEWORK: THE CAUSAL THEORY OF ACTION	1
1. The anti-causalist argument	3
2. Rationalization as a species of causal explanation	7
3. Anomalous monism	10
4. The truth of sigular causal statements	15
5. Conclusion	18
II. DAVIDSON ON WEAKNESS OF WILL	21
1. Plato and Aristotle on weakness of will	25
2. The nature of practical reasoning	33
3. The concepts of rationality and irrationality	41
4. The theory of partitioned mind	49
5. Conclusion	53
III. INCONTINENCE AS A FAILURE IN REASONING.....	56
1. Incontinence as a failure in reasoning	57
2. Two kinds of evaluative judgments	60
3. The motivational force of a prima facie judgment	79
4. Probabilistic akrasia	96
5. Conclusion	104
IV. IRRATIONALITY AS AN INNER INCONSISTENCY.....	107
1. A review of Davidson's theory of partitioned mind	108
2. The explanatory force of the theory of partitioned mind	113
3. The necessity of the theory of partitioned mind	124
4. Consistency, rationality, and irrationality	130
5. Conclusion	144
V. CONCLUSION.....	148
REFERENCES	157

I. BASIC FRAMEWORK: THE CAUSAL THEORY OF ACTION

The concept of weakness of will is defined as follows by Donald Davidson:

An agent's will is weak if he acts, and acts intentionally, counter to his own best judgment; in such cases we sometimes say he lacks the willpower to do what he knows, or at any rate believes, would, everything considered, be better.¹

An action which fits the description is called an incontinent action, and the agent an incontinent agent. The phenomenon of weakness of will is puzzling because its occurrence seems contrary to a widely accepted doctrine that when an agent acts intentionally, he acts in the light of what he imagines or judges to be better. The doctrine is a statement of internalism the advocates of which believe that there is a necessary connection between evaluation and action. At least a kind of evaluative judgment is capable of moving an agent to act. Unless one denies internalism and holds an externalistic thesis that evaluative judgment plays no distinctive role in action and practical thought, the principle must be true to him in a sense. Then, given the principle, how is weakness of will possible? Suppose such a phenomenon is acknowledged to be existent, how does an incontinent agent intentionally act contrary to his best judgment?

In the history of philosophy, many have tried to solve the problem of weakness of will. Some deny the possibility of weakness of will in order to preserve the truth of the internalist doctrine. Some provide accounts reconciling both the internalist doctrine and the phenomenon of weakness of will. Davidson also gives a solution to the problem which defends the truth of a version of the internalist doctrine and allows the possibility of weakness of will. The significance

¹ Donald Davidson, 'How is Weakness of the Will Possible?' (1969), reprinted in *Essays on Actions and Events* (2001), p. 21.

of Davidson's discussion on weakness of will, however, lies not only in the solution itself but also in Davidson's attempt to show how his causal theory of action accommodates the phenomenon of weakness of will. In order to see how Davidson's causal theory of action does not rule out the possibility of weakness of will, this chapter aims to represent relevant theses in Davidson's causal theory of action as a background.

Davidson's causal theory of action mainly argues for the view that the reason of an action is the cause of the action, which is a reaction against those who hold that a reason that explains an action and the action cannot be related as cause and effect. According to Davidson, a principal argument the anti-causalists hold is that 'causal relations are essentially nomological and based on induction while our knowledge that an agent has acted on certain reasons is not usually dependent on induction or knowledge of serious laws.'² Anti-causalists take being law-like, inductive, and predictable as the standard to identify causal relations and thereby deny that reasons and actions are in causal relations. In this chapter, I explain several theses in Davidson's causal theory of action in terms of Davidson's defense against this principal argument. In the following, section 1 introduces the anti-causalist argument from Wittgenstein. Section 2 illustrates Davidson's views on the nature of action and the nature of action explanation. It explains Davidson's theses that an action is picked out by its causal history and that rationalization is a species of causal explanation. Davidson develops his thought about mental and physical events as the anomalous monism, and section 3 explicates the thesis that mental events are physical events, but there are no psychological or psychophysical laws. Section 4 shows that Davidson's rebuttal of anti-causalist arguments is mainly based on the truth of singular causal statements. Section 5 is a concluding remark of this chapter.

² Donald Davidson, 'Introduction' of *Actions, Reasons, and Causes* (2001), xvi.

1. THE ANTI-CAUSALIST ARGUMENT

The view that reasons and actions can be related as cause and effect is rejected by Wittgenstein and his followers. Of the various anti-causalist arguments, Davidson points out that the main argument anti-causalists hold is that causal relations are essentially nomological, based on induction, and backed up by laws, but the relations between reasons and actions are not. Thus, reasons and actions are not in causal relations. The first section introduces the anti-causalist argument against which Davidson develops his causal theory of action as a reaction.

The investigation of the nature of action is to discover what distinguishes actions from purely bodily movements. Wittgenstein believes that what identifies an action is its voluntariness. An action must be a voluntary movement. Wittgenstein claims that,

What is voluntary is certain movements with their normal surrounding of intention, learning, trying, acting. Movements of which it makes sense to say that they are sometimes voluntary and sometimes involuntary are movements in a special surrounding.³

When a person acts voluntarily, his voluntary action requires a behavioral context, a normal surrounding consisting of intention, motivation, other mental states, and the behavior. A person has no doubt that he acts voluntarily when he acts voluntarily because he acts on his intention, motive, and other mental states, and usually he is conscious of the fact that he acts in such a normal surrounding. Thus, actions like drinking, eating, reading, and walking are voluntary actions as they occur in normal surroundings. Those that occur in abnormal surroundings such as in hypnosis or in narcosis are merely what happen to an agent. An agent does not know how to explain these bodily movements for there are no corresponding intentions, motivations, or

³ Wittgenstein, Zettel (1967), §577.

expectations with respect to them. To an observer, similarly, he can tell whether a bodily movement is an action or not when he knows the context in which the bodily movement takes place. In a context of goal-directed activity, if a bodily movement fits in with the context and appears to be reasonable, the bodily movement is an action. By placing the action in a larger context consisting of mental states such as motivation, intention, and the behavior, an observer understands why an agent does something by grasping the reason in the context that explains the action. The context provides him with a background of the action and makes the action intelligible in terms of its reason.

The reason abstracted from the context in which an action occurs functions as a justification of the action rather than as its cause. Relevant to the principal argument that Davidson mentions, Wittgenstein contrasts the differences between reason and cause to support the view that the reason of an action is not the cause of the action. First, an agent has the first-person authority to give the reason of his action, but he may not know the cause of his action. When an agent is asked why he does something, he needs not to observe anything before giving the reason, but can tell us what he remembers about the context in which the action occurs. What he sincerely claims to be the reason of his action is indeed the reason of the action. Even if he mistakenly believes something and acts on the wrong belief, when he gives the belief as a reason of his action without causing any obvious contradiction, the belief that well explains the action still functions as a justification. Wittgenstein says, 'The reason may be nothing more than just the one he gives when asked.'⁴

⁴ Wittgenstein, *Lectures on Philosophy 1932-33*, §4.

The cause of an action, however, is an event whose occurrence brings about an action as an effect and could be unknown to the agent. Wittgenstein gives an example:

Suppose an engine-driver is asked ‘Why do you stop here?’ and answers ‘Because the signal was at “Stop!”’ This answer is mistakenly regarded as stating a cause, when in fact it states a reason. (It may of course be that the sight of the red signal caused the man, through ‘conditioned reflexes’, to pull the brake lever, but he need not be aware of this cause; whereas he knows very well why he stopped here, this just shows that his answer states the reason for his action.)⁵

Causal explanations of this kind explain actions not from an agent’s point of view but in terms of physical, physiological, or other objective perspectives. Observations are required to discover the cause of an action, and any causal statements we make could be falsified by further observations. In this sense, what the cause of an action really is and whether the agent knows it are irrelevant to the rationalization of the action.

Secondly, to give the reason of an action is to specify the route the agent has taken, to describe a singular process, but to give a cause is to offer a hypothesis which assumes the principle of induction. Reasons have logical relations with what they justify. For example, in a calculation, Wittgenstein claims, ‘To give a reason is to go through a process of calculation, and to ask for a reason is to ask how one arrived at the result.’⁶ A person may give certain arithmetic laws that he appeals to in a calculation to explain how he reasons and reaches a conclusion. The arithmetic laws are what he believes to be true and holds to derive a reliable consequence. Whether the answer he reaches is correct or not, the arithmetic laws make his calculation intelligible. Causes and the phenomenon they explain, however, are not in logical relations. We do not specify a route or a process when we state a cause. ‘Stating a cause would be offering a

⁵ Waismann, ‘The Causal Interpretation of Language’ (1965), p. 121.

⁶ Wittgenstein, *Lectures on Philosophy 1932-33*, §4.

hypothesis.’⁷ A causal statement usually involves a hypothesis which can be expressed as a sentence with the form ‘If A, then B’ where A states the condition of being a cause and B that of being an effect. Two objects that are claimed to be cause and effect are believed to instantiate a hypothesis, rule, or general law to which there are no exceptions. Hypotheses, rules, or laws that back up ordinary causal statements are not known *a priori* but are discovered by experiments. They all assume the principle of induction which has the nature that, given a sufficient number of repeated associations between two objects, we can expect that the occurrence of a fresh association of the same kind is highly probable and that a general law about associations of this kind is highly probable or almost certainly to be true.⁸

Thirdly, ordinary causal statements are law-like, but statements using psychological terms are not. Reasons-explanations are statements with psychological terms such as ‘believe’ or ‘want’. They do not involve hypothetic elements, and so there are no laws on the basis of which we can form a precise prediction about what an agent is going to do or explain what an agent has done. On the contrary, that naming causal connections involves hypothetical elements allows that when we know a hypothesis, rule, or law, and the occurrence of something that is covered by it, we can deduce that another thing must follow to occur as an effect. Conversely, when we see something happen, we can infer backward that the occurrence of something else causes it. This causal relation relies on a hypothesis, rule, or law. Since causal explanations conform to this deductive-nomological model of explanation and reason-explanations do not, the reason of an action must be distinguished from the cause of an action.

⁷ Wittgenstein, *Lectures on Philosophy 1932-33*, §35.

⁸ Bertrand Russell, ‘On Induction’ (1959), pp. 66-8.

2. RATIONALIZATION AS A SPECIES OF CAUSAL EXPLANATION

Of the points that Wittgenstein holds, Davidson shares Wittgenstein's idea that an action is intelligible when it is placed in a larger context which makes it rational, but he denies that merely knowing the context in which an action appears reasonable is enough to distinguish it from purely bodily movements and that grasping a reason from the context necessarily yields a correct reason explanation of the action. This is because an action could be shown as reasonable under different reasons and contexts. In such a situation, the real reason the agent holds for taking the action cannot be identified, and whether a given reason explanation is correct or not becomes unknown. To avoid this difficulty, Davidson suggests that the reason of an action and the action must be in a causal relation. The causality between reason and action is crucial to Davidson's analysis of the nature of action and of action explanation. According to Davidson, a bodily movement is an action if it is caused by a reason that the agent has for taking the action. A reason effectively explains an action only when the reason is the cause of the action. Thus, a reason explanation of an action must also be a causal explanation. This section explains these theses on the nature of action and action explanation.

Davidson explains the thought that rationalization is a species of causal explanation by illustrating the idea of primary reason. Davidson claims that, whenever someone does something for a reason, 'he can be characterized as (a) having some sort of pro attitude toward actions of a certain kind, and (b) believing (or knowing, perceiving, noticing, remembering) that his action is of that kind.'⁹ Pro attitudes are positively evaluative judgments of something, and can be desires, wantings, urges, promptings, moral views, social conventions, goals, values, and so on. All pro

⁹ Donald Davidson, 'Actions, Reasons, and Causes' (1963), reprinted in *Essays on Actions and Events* (2001), pp. 3-4.

attitudes, however, can be reduced to desires or wantings because an agent's having a pro attitude toward something entails that he thinks it desirable or wants it. Thus, desire or wanting can be treated as a genus including all pro attitudes as a species. When an agent acts for a reason, the reason could be a complex of various propositional attitudes, but a pair of belief and desire can always be found or refined from the complex to explain the action. The pair of belief and desire that the agent holds for taking the action is the primary reason of the action. The role a primary reason plays in an action explanation is to rationalize an action by leading us to see a goal the agent thinks attractive and the means he thinks possible to reach the goal. Whenever the primary reason of an action is given, it is informative enough for us to understand the action. For a reason to be a primary reason that correctly explains an action, it must satisfy two conditions:¹⁰

C1. *R* is a primary reason why an agent performed the action *A* under the description *d* only if *R* consists of a pro attitude of the agent towards actions with a certain property, and a belief of the agent that *A*, under the description *d*, has that property.

C2. A primary reason for an action is its cause.

C1 explicates what reason explanations are, and C2 identifies reason explanations with causal explanations. The two conditions together support Davidson's position that rationalization is a species of causal explanation.

According to C1, a pair of belief and desire explains an action only when belief, desire, and action are described in a proper way which reveals the logical relations between them. Davidson thinks that an action can have various descriptions of it. To describe an action in a way is to place the action in a circumstance to understand, and so re-describing an action enables us

¹⁰ Ibid., p. 5 & p. 12.

to understand the action from different aspects. Among the descriptions of an action, some describe the action in terms of its reason, a pair of belief and desire. A description of this kind exhibits a logical relation with the description of the action. It helps us to understand the action by rationalizing it. Davidson gives an example: 'I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home.' In this example, only one action occurs, but four descriptions are given. My wanting to turn on the light rationalizes my flipping the switch because my flipping the switch is a reasonable consequence of my desire to turn on the light and my belief that flipping the switch is a way to turn on the light. But it does not rationalize my alerting a prowler that I am home for it does not make sense to say that I alerted a prowler that I am home because I wanted to turn on the light. A pair of belief and desire whose descriptions exhibit a reason relation with the description of an action is essential to the rationalization of the action. Sometimes an agent merely offers a belief or a desire to explain his action, and this single mental state seems sufficient to rationalize his action. However, we know that the agent also has a paired belief or desire which is not explicitly indicated but from which and the paired belief or desire the agent forms a conclusion about what to do. Descriptions that rationalize an action entail that under these descriptions the action is intentionally done by the agent. Descriptions that fail to show an agent's intention to take an action do not rationalize the action. Intentional action can then be defined as one done for a reason.¹¹

What needs to be explained in C2 is why reason explanations are causal explanations. Davidson's positive reason for C2 is that a primary reason truly explains an action only when it is also the cause of the action. An action can be rationalized by various pairs of belief and desire

¹¹ Ibid., p. 6.

when the description of a pair of belief and desire and the description of an action are in a logical relation. However, since it is possible that an action can be rationalized simultaneously by various pairs of belief and desire, a rationalization may not truly explain an action unless the pair of belief and desire cited in the rationalization is indeed the reason on which the agent acts. The causal relation between a reason and an action it causes guarantees that the agent takes the action for the reason. Thus, the causal history between a reason and a bodily movement helps to pick out an intentional action on the one hand, and it makes sure that a successful reason explanation of the intentional action can be yielded on the other hand. It is only when reason and action are related as cause and effect, the word ‘because’ which an agent uses to explain his action is meaningful. As Davidson remarks,

Central to the relation between a reason and an action it explains is the idea that the agent performed the action because he had the reason. Of course, we can include this idea too in justification; but then the notion of justification becomes as dark as the notion of reason until we can account for the force of that ‘because’.¹²

When an agent explains his eating by saying that ‘I eat the apple because the unrefrigerated apple will soon rot’, the word ‘because’ connects the agent’s reason for eating and the act of eating as a cause and an effect. The reason that he gives to explain his action truly explains the action because it is the reason that causes the action.

3. ANOMALOUS MONISM

C2 is controversial. The thesis that reason explanation is a species of causal explanation is what anti-causalists argue against. However, Davidson also thinks that there are no

¹² Ibid., p. 9.

psychological or psychophysical laws on the basis of which mental events can be predicted and explained, so he would assent to Wittgenstein's assertion that the relation between reason and action is neither predictable nor nomological. This view is elucidated in Davidson's anomalous monism: 'monism, because it holds that psychological events are physical events; anomalous, because it insists that events do not fall under strict laws when described in psychological terms.'¹³ This section first explicates Davidson's argument for the anomalism of the mental, and then explains the argument for the physical monism.

A strict law, according to Davidson, 'may have provisos limiting its application,' but it 'allows us to determine in advance whether or not the conditions of application are satisfied.'¹⁴ It promises that whenever the conditions are satisfied, a consequence of a kind must follow to occur. In this sense, a law must support counterfactual claims about what something would be like if it were described as satisfying the conditions, and must be confirmed by instances properly described. Such a precise, explicit, and exceptionless law is possible only if it draws its concepts from a comprehensive closed theory.¹⁵ A domain is closed if nothing outside the domain will causally affect events within the domain. A theory is comprehensive if every event or phenomenon in the domain can be described by vocabularies of the theory and be explained and predicted under that description.¹⁶ It is only in a comprehensive closed theory that there are strict laws on the basis of which we can explain or predict events or phenomena in the domain.

What Davidson has in mind here is physics. Causal closure is a constitutive feature of the physical. 'It is a feature of physical reality that physical change can be explained by laws that

¹³ Donald Davidson, 'Psychology as Philosophy' (1974), reprinted in *Essays on Actions and Events* (2001), p. 231.

¹⁴ *Ibid.*, 233.

¹⁵ Donald Davidson, 'Mental Events' (1970), reprinted in *Essays on Actions and Events* (2001), p. 219.

¹⁶ Jaegwon Kim, 'Philosophy of Mind and Psychology' (2003), p. 116.

connect it with other changes and conditions physically described.’¹⁷ As a cause or effect of other physical events, every physical event has a physical explanation which is backed up by a physical law. The domain of mental phenomena, on the contrary, is not closed. Davidson believes that ‘at least some mental events interact causally with physical events.’¹⁸ For example, someone presses the button ‘on’ of a remote control and the air conditioner is on. His motive, intention, decision, and other mental events play a causal role in the operation of the air conditioner. Waiting for a while, he perceives a more comfortable temperature in the room, which causes him to believe that the air conditioner is operating properly. The lack of causal closure of the mental is a reason that Davidson holds to deny psychological laws.

There are no psychophysical laws either. Laws connect properties and kinds, not individual events. If there are psychophysical laws, the laws will connect psychological properties or kinds to physical properties or kinds. Because psychophysical laws connect the mental to the physical, there are constraints imposed by the physical side on the mental realm.¹⁹ That is, the psychophysical laws must be statements with physical predicates, predicates using quantitative and purely physical concepts to describe relevant properties or kinds. However, there are no such laws. Mental properties, such as meanings or thoughts, cannot be reduced to properties dealt with in physics. The cognitive field has a holistic character. Beliefs, desires, and an agent’s other propositional attitudes are logically connected, so they assume rationality, consistency, and coherence in a degree. These normative characters, rationality, consistency, and coherence, are constitutive and unique to the mental. There is no way to reduce mental properties

¹⁷ Davidson, ‘Mental Events’, p. 222.

¹⁸ Ibid., p. 208.

¹⁹ Kim, ‘Philosophy of Mind and Psychology’, p. 121.

to physical properties because the normative characters 'have no echo in physical theory'.²⁰ If there is any way to reduce the mental to the physical, the reduction will make the mental lose the normative characters and then deprives the mental of its nature. The mental is nomologically irreducible. Causal statements connecting the mental to the physical are not law-like statements, and there are no psychophysical laws supporting causal statements of this kind. The most that we can look for are rough correlations between psychological and physical phenomena.

The lack of necessary and sufficient conditions for acting on a reason is another reason which explains why there are no serious laws connecting reasons and actions and why action predictions cannot be successful. As Davidson indicates, a pair of belief and desire of the right sort does not guarantee a correct reason explanation. For a man may have a reason to do something and also take a corresponding action, but the reason is not the one with which he acts. The causality between reason and action is required as a necessary condition so that a reason that is efficacious in producing an action indeed explains the action. The causal connection between reason and action, however, is not sufficient to ensure that what is done is done with the appropriate intention or with any intention because deviant causal chains exist. For example, 'Someone might want tasty stew and believe sage would do the trick and put in sage thinking it was parsley; or put in sage because his hand was joggled'.²¹ In the latter case, the action is not intentional at all. In the former case, the action has a description in which it is intentional but is not intentional in the expected sense.

If there were necessary and sufficient conditions for acting on a reason, laws connecting reasons and actions could be given. Davidson remarks, '...suppose we had the sufficient

²⁰ Davidson, 'Psychology as Philosophy', p. 231.

²¹ Donald Davidson, 'Intending' (1978), reprinted in *Essays on Actions and Events* (2001), p. 87.

conditions. Then we could say: whenever a man has such-and-such beliefs and desires, and such-and-such further conditions are satisfied, he will act in such-and-such a way. There are no serious laws of this kind.²² A serious law must support counterfactuals, subjunctives, and must be confirmed by instances, but we can easily find that a man has a pair of belief and desire of certain kind but does not take a corresponding action. An action is caused by a pair of belief and desire, and it is a result of a process of reasoning in which the agent weighs competing reasons to form various practical arguments. Thus, to predict actions on the basis of beliefs and desires, knowing a single pair of belief and desire is not enough. A law for action prediction must be a quantitative calculus that brings all relevant beliefs and desires into the picture. It must find a way to evaluate the relative force of various desires and beliefs in the matrix of decision.²³ Nevertheless, even if constructing such a law is possible, the fact that beliefs and desires change over time would prevent us from getting accurate predictions.

To Davidson, there are no psychological or psychophysical laws but only physical laws, and all psychological events must also be physical events. The physical monism is a consequence of the following argument:

If psychological events cause and are caused by physical events (and surely this is the case) and if causal relations between events entail the existence of laws connecting those events, and these laws are, ..., physical, then it must follow that psychological events simply *are* (in the sense of *are identical with*) physical events.²⁴

There are causal interactions between the mental and the physical. For example, a man throwing a pebble into a lake causes a ripple on the surface of the lake, and the beautiful ripple makes him

²² Davidson, 'Psychology as Philosophy', p. 233.

²³ Davidson, 'Actions, Reasons, and Causes', p. 16.

²⁴ Donald Davidson, 'The Material Mind' (1973), reprinted in *Essays on Actions and Events* (2001), p. 248.

pleased. For any events related as cause and effect, they must instantiate a strict law. A mental event and a physical event that are causally connected must also instantiate a strict law. Davidson has shown that there are no psychological or psychophysical laws, all strict laws are physical laws. Since a mental event and a physical event that are in a causal relation instantiate a physical law, the mental event falls under a physical kind and has a physical description. Because physical events are picked out by physical descriptions,²⁵ the mental event which causally interacts with a physical event, instantiates a physical law, and has a physical description, is a physical event.

4. THE TRUTH OF SINGULAR CAUSAL STATEMENTS

Although there are no psychological or psychophysical laws which back up causal statements involving mental events and are able to predict actions from reasons, it does not follow that causal statements involving mental events cannot be true. This section explains why singular causal statements in which particular mental events are causally connected with particular mental or physical events can be true, and thereby represents Davidson's response to the anti-causalist argument.

The fact that there are no psychological or psychophysical laws merely rules out the truth of statements relating a type of reasons and a type of actions as cause and effect. Laws connect properties and kinds, not individual events. Since there are no psychological or psychophysical laws, statements claiming that an action of a kind must follow to occur whenever a mental event of a kind occurs cannot be true.

²⁵ Davidson, 'Mental Events', p. 211.

However, this does not entail that causal statements involving particular mental events cannot be true. Events are unrepeatable, dated individuals, and can have different properties and fall under different kinds. Davidson claims that ‘The truth of a causal statement depends on *what* events are described; its status as analytic or synthetic depends on *how* the events are described.’²⁶ In a causal statement, if two events causally related are properly described by vocabularies employed by a law as having certain properties or being a certain kind that is covered by the law, the statement is shown as an instance of the law. But the truth of a singular causal statement is not determined by whether it can be described in a way as an instance of a law. A causal statement is true if and only if its referents are indeed in a causal relation. The causal interaction between the mental and the physical ‘deals with events in extension and therefore is blind to the mental-physical dichotomy.’²⁷ A reason explanation that explains an action by giving a reason as the cause of the action aims to give descriptions in which the reason and the action are logically related. Although a causal statement with descriptions of this kind does not instantiate a law, it is a true singular causal statement if the reason and the action that it refers to are truly in a causal relation.

A true causal statement with descriptions of a reason and of an action does not instantiate a law, but it entails a law. Davidson believes that ‘... where there is causality, there must be a law: events related as cause and effect fall under strict deterministic laws.’²⁸ For two events related as cause and effect, some law covering the events at hand exists. The events, when appropriately described in a causal statement, fit into the law. We usually do not know the law covering two events in a causal relation, but Davidson argues that ignorance of competent predictive laws does

²⁶ Davidson, ‘Actions, Reasons, and Causes’, p. 14.

²⁷ Davidson, ‘Mental Events’, p. 215.

²⁸ *Ibid.*, 208.

not inhibit valid causal explanations. In fact, we frequently make ordinary causal statements without knowing the laws behind them. We merely know a relevant generalization induced by similar cases, and we predict what is going to happen by this generalization. Generalizations are not laws. The function of generalizations is to provide evidence of the existence of a causal law covering the case at hand.²⁹ The truth of a generalization is confirmed by resembling single cases, and the truth of a law is confirmed by generalizations. For instance, ‘a paper clip, a book, or any object supported by nothing will fall to the ground’. This generalization, and other general statements, such as those about the revolution of the earth around the sun or the formation of tides, together confirm the law of attraction which is roughly understood as ‘objects with mass attract each other.’ In this sense, it is much easier to know the truth of a statement about a particular case than to know the truth of a generalization or a law. We can know that a singular statement is true while being ignorant of the law involved. There is a passage in Bertrand Russell’s ‘Induction’:

The probability of the general law is obviously less than the probability of the particular case, since if the general law is true, the particular case must also be true, whereas the particular case may be true without the general law being true.³⁰

Analogously, in a case in which reason and action are related as cause and effect, we do not know the involved law. We probably know that when an agent genuinely desires something and believes a certain kind of action is a way to satisfy his desire, generally he will take a corresponding action if nothing prevents him from doing so. But we can give a genuine causal explanation of an action even if we do not know the involved law.

²⁹ Davidson, ‘Actions, Reasons, and Causes’, p. 16.

³⁰ Russell, ‘Induction’, p. 67.

As to Wittgenstein's taking our different accesses to know reason and cause to distinguish the reason of an action from the cause of the action, Davidson replies that on some occasions an agent does not have first-person authority to give the reason of his action because he may be wrong about his reason, especially when he has two reasons for an action, one of which pleases him while the other does not.³¹ For example, a man killed a murderer and he thought that he did so because of justice, but in fact he was motivated by his desire to revenge his dead brother. Davidson agrees that in most cases, an agent can give his reason for an action without observation and evidence which are thought necessary for giving a cause. But it is possible that the evidence available to others may overthrow self-judgments.³² Self-attributions are not incorrigible, we do not always have indubitable or certain knowledge of our own attitudes, and so the privileged access to our own propositional attitudes cannot be used to distinguish reasons from causes. First-person authority does not lie in the fact that an agent usually has privileged access to know his own propositional attitudes and others do not since there are exceptional cases. It comes from the fact that in interpretations, 'speakers, but not their interpreters, are not wrong about what their words mean.'³³ Whether or not a speaker uses words with their standard meaning or misuses words, what his words mean is what he intends them to mean. In an interpretation, the interpreter needs to grasp the meaning that the speaker intends his words to have. Because the speaker usually knows what his words mean, he knows what he believes while the interpreter may not.

5. CONCLUSION

³¹ Ibid., p. 18.

³² Donald Davidson, 'First Person Authority' (1984), reprinted in *Subjective, Intersubjective, Objective*, (2001), p. 4.

³³ Ibid., p. 12.

In order to show how Davidson's causal theory of action accommodates his account of the phenomenon of weakness of will, this chapter represented relevant theses of Davidson's causal theory of action as a background. These theses are essential to understand how Davidson thinks about action, reason, cause, action explanation, and causal explanation. And I have explained these theses in terms of Davidson's response to the anti-causalist argument.

Anti-causalists deny that reason and action can be related as cause and effect because of the following contrasts between reason and cause. An agent has the first-person authority to give the reason of his action, but the cause of the action may be unknown to him. To give a reason is to specify a reasoning process one has taken, but to give a cause is to offer a hypothesis, rule, or law which assumes the principle of induction. The relation between reason and action is not law-like, but events in a causal relation instantiate a law and so their occurrence can be explained and predicted by laws.

Although Davidson is a causalist, he agrees with anti-causalists that there are no psychological or psychophysical laws on the basis of which we can explain or predict mental events and their connections to other mental or physical events. Physical events with mental properties are mental events, and events of this kind are anomalous. The lack of psychological and psychophysical laws is due to these facts: mental events are not causally closed, mental properties are not reducible to physical properties which are described by quantitative and purely physical concepts, and there are no necessary and sufficient conditions for intentional actions.

The relations between mental and physical events are neither nomological nor backed up by strict laws, but it does not follow that causal statements with psychological terms cannot be true. Whether or not events instantiate laws depends on whether they are described by concepts

employed by laws. Because the purpose of reason explanations is to explain why an agent does something by giving a reason that rationalizes his action, and the psychological predicates in these statements use phrases that are not employed by strict laws, reason explanations do not instantiate laws. When events in causal relations are described as reasons and actions, there are no laws instantiated by these mental descriptions. This fact rejects the view that there are laws connecting a type of mental events to another type of mental or physical events, but it does not contradict the claim that particular events can be in causal relations covered by strict laws.

Davidson believes that events in causal relations must involve strict laws. Statements relating events as causes and effects are true if and only if referents of the statements are truly in causal relations. Thus, although events described as reasons and actions do not instantiate laws, a reason and an action are in a causal relation when the former indeed causes the latter. Singular causal statements connecting reasons and actions are true when reasons and actions which the statements denote are in causal relations.

Singular causal statements which connect reasons and actions explain actions by giving reasons, rational causes. They are both reason explanations and causal explanations. Davidson points out that a reason that causes an action basically consists of a pair of belief and desire. It rationalizes the action when the descriptions of the belief, desire and action exhibit a logical relation which makes the action intelligible. Only when an action is caused by its reason, can we avoid the difficulty that the real reason an agent has for his action is not correctly identified when there are various contexts in which the action is shown as reasonable simultaneously. The reason of an action must also be the cause of the action so that whenever we give a reason to explain an action, the reason truly explains the action. Reason explanation is a species of causal explanation.

II. DAVIDSON ON WEAKNESS OF WILL

According to Donald Davidson, an action which shows weakness of will or incontinence can be characterized as follows:

D. In doing x an agent acts incontinently if and only if: (a) the agent does x intentionally; (b) the agent believes there is an alternative action y open to him; and (c) the agent judges that, all things considered, it would be better to do y than to do x .¹

If an agent once held an all-things-considered judgment that doing y is better than doing x , but he no longer holds the judgment when he does x , or if he did x and later judges that it would have been better for him to do y than do to x , his doing x is not an incontinent action. Incontinence requires that an incontinent agent simultaneously holds two conflicting evaluative judgments about what to do when he acts.

The existence of incontinent actions is problematic because it challenges a widely accepted doctrine that when an agent acts intentionally, he acts in the light of what he imagines or judges to be better. The doctrine is a statement of internalism which assumes a special connection between evaluative judgment and intentional action. Davidson believes that the doctrine can be spelled out by two principles:²

P1. If an agent wants to do x more than he wants to do y and he believes himself free to do either x or y , then he will intentionally do x if he does either x or y intentionally.

P2. If an agent judges that it would be better to do x than to do y , then he wants to do x more than he wants to do y .

¹ Donald Davidson, 'How is Weakness of the Will Possible?' (1969), reprinted in *Essays on Actions and Events* (2001), p. 22.

² *Ibid.*, pp. 22-3.

P2 states a connection between value judgment and wanting, and P1 a connection between wanting and intention. The two principles together entail that if an agent judges that it would be better for him to do x than to do y , and he believes that he is free to do either x or y , then he will intentionally do x if he does either x or y intentionally. The entailment seems to suggest that there is an internal or necessary connection between evaluative judgment, motivation, and action in the sense that whenever an agent makes an evaluative judgment, the judgment will guide him to take a corresponding action. However, if this is the case, an agent will never act against his evaluative judgment, and the phenomenon of weakness of will becomes impossible.

Davidson takes P1 and P2 to be true. He claims that ‘P1 and P2 derive their force from a very persuasive view of the nature of intentional action and practical reasoning.’³ The view is originally from Aristotle, and is modified and developed by Davidson as a part of his causal theory of action. Aristotle says, ‘The origin of action- its efficient, not its final cause- is choice, and that of choice is desire and reasoning with a view to an end.’⁴ An intentional action is an overt consequence of a practical reasoning in which the agent weighs reasons for and against certain courses of action, and finally forms an argument about what to do. The conclusion of the argument is an evaluative judgment about what is better for him to do. When it causes the agent to act, the agent acts on what he judges to be better.

Davidson, on the one hand, defends the truth of P1 and P2. On the other hand, he also holds P3 to be true:

P3. There are incontinent actions.

³ Ibid., p. 31.

⁴ Aristotle, *Nicomachean Ethics*, 1139a 31-32.

Whether the phenomenon of weakness of will exists is not a problem to Davidson. Instead, Davidson takes the existence for granted. However, there is an apparent conflict between the three principles. The truth of P1 and P2 seems to rule out the truth of P3. Eliminating the seeming conflict between the three principles is then crucial to the explication of the possibility of weakness of will. Furthermore, since P1 and P2 are essential elements in Davidson's causal theory of action, Davidson must show how his causal theory of action can accommodate the phenomenon of weakness of will. As Davidson indicates, 'if reasons are causes, it is natural to suppose that the strongest reasons are the strongest causes.'⁵ When an agent holds two conflicting evaluative judgments, the agent is supposed to be caused to act by the one with a stronger reason. Thus, Davidson's account of the phenomenon of weakness of will is also an attempt to show how weakness of will is possible given such a causalist thesis.

Another relevant problem that Davidson raises is a problem about the origin of weakness of will: Why does an incontinent agent act against his own best judgment? Although an incontinent agent acts against what he judges to be better, the action that he performs is caused by a reason and is intentionally done by him. The reason explains why he takes the action and reveals the rational aspect of the action. However, given that a reason for an action has the function of rationalization, the incontinent agent does not have a reason for his acting against his own best judgment. The reason that rationalizes the action he performs does not rationalize his acting incontinently. Davidson clarifies that,

⁵ Davidson, 'Introduction' of *Actions, Reasons, and Causes* (2001), xvi.

... although the agent has a reason for doing what he did, he had better reason, by his own reckoning, for acting otherwise. What needs explaining is not why the agent acted as he did, but why he *didn't* act otherwise, given his judgment that all things considered it would be better.⁶

In response to this second problem of weakness of will, Davidson looks not for what causes an incontinent agent to take the action but what causes him not to take the other available action which he judges to be better. That is, he looks for the cause of weakness of will.

This chapter aims to represent Davidson's solutions to the two problems of weakness of will, solutions which constitute a complete picture of Davidson's view on weakness of will. Section 1 begins with Plato's and Aristotle's views on the possibilities of weakness of will, for both Plato's establishing an internal and necessary connection between evaluation and action and Aristotle's introducing the form of a practical reasoning are influential conceptions to Davidson. Section 2 shows how Davidson interprets the three principles in a way which renders the three principles, P1, P2, and P3, consistent with one another. It emphasizes that what makes Davidson's interpretation achieve its purpose is the contrast between prima facie, conditional judgments and all-out, unconditional judgments. Through such a distinction, Davidson construes P2 as a mild form of internalism, and assigns prima facie judgment as the form of the evaluative judgment serving as the major premise of a practical syllogism. P3 is then shown as compatible with P1 and P2, and P1 and P2 are shown to be necessary conditions of P3. Incontinent actions are possible because incontinent agents, under Davidson's interpretation, do not hold logically contradictory beliefs when they act.

Although an incontinent agent commits no logical blunder when he acts, he is wrong in that he intentionally acts against what he judges to be better. An incontinent agent as a thinking

⁶ Donald Davidson, 'Paradoxes of Irrationality' (1982), reprinted in *Problems of Rationality* (2004), pp. 176-7.

creature is determined to own fundamental norms of rationality, including the principle of continence which requires him to perform the action judged best on the basis of all available relevant reasons.⁷ Because of the built-in principle of continence, when an incontinent agent acts, he simultaneously holds two inconsistent beliefs which make him irrational: the judgment to take the action, and the judgment enjoined by the principle of continence. Section 3 focuses on Davidson's characterization of the concepts of rationality and irrationality. Davidson takes rationality as constitutive of creatures having propositional attitudes and intentional actions, and takes the attribution of rationality to agents to be a basic assumption for understanding their utterances and actions. Incontinent actions, irrational phenomena defined by agents' holding inconsistent propositional attitudes, make sense only when they are seen against a background of rationality. Section 4 introduces a theory of partitioned mind which is developed to explain the mental cause of weakness of will, a mental cause that is not a reason for what it causes. The theory has a more general purpose as well for it explains how an agent can hold inconsistent beliefs without putting them together. Section 5 is a conclusion of the chapter.

1. PLATO AND ARISTOTLE ON WEAKNESS OF WILL

Davidson interprets the three principles in a way that shows them to be consistent, and what makes Davidson's interpretation work is his distinguishing two kinds of evaluative judgments. The distinction allows Davidson to interpret P2 as a mild form of internalism, which is different from a strong version of internalism held by Plato, and to assign prima facie judgment as the form of the major premise of a practical argument, which is different from Aristotle's treating the major premise of a practical syllogism as a universally quantified conditional.

⁷ Davidson, 'How is Weakness of the Will Possible?', p. 41.

Davidson's strategy, in a sense, is a modification of Plato's internalism and Aristotle's concept of practical reasoning. The aim of this section is to give some context to understand Davidson's modification.

One account of weakness of will is that those who act incontinently are overcome by pleasure. In Plato's *Protagoras*, Socrates denies such an account and claims it to be absurd.⁸ Socrates argues that we judge something to be good or bad not according to the pleasure or pain it results in immediately but according to the overall pleasure or pain it brings about in the long term. Something painful is 'good, whenever it relieves pains greater than the ones it contains or brings about greater pleasures than pains', and something enjoyable is 'bad whenever it deprives us of greater pleasures than it itself provides, or brings about greater pains than the very pleasures inherent in it.'⁹ The criterion of being good and bad entails that what is pleasant is good and that what is painful is bad. 'Pleasant' and 'good' are co-referential, and so are 'painful' and 'bad'. The absurdity of the 'overcome by pleasure' account is obvious when 'good' substitutes for 'pleasure'. To say that an agent is incontinent because he is overcome by pleasure is to say that an agent is incontinent because he is overcome by the good. The agent then does not make any mistake. For when he is overcome by pleasure, the good, to act, he acts on what is better. The 'overcome by pleasure' account fails to show why the incontinent agent is wrong, and it even contradicts the description of the phenomenon of weakness of will.

Although Socrates denies the 'overcome by pleasure' account of weakness of will, he does not give an account of it. Instead, Socrates rejects the existence of weakness of will. Knowing how to weigh pleasure against pain and weigh the good against the bad is an art of

⁸ Plato, *Protagoras*, 355a-b.

⁹ *Ibid*, 354d-e.

measurement. The art of measurement is knowledge, which is supposed to rule in our behaviors and not be dragged around by other things.¹⁰ It guides us to make right choices about pleasure and pain, good and bad. Socrates assumes that when we act, we ‘pursue pleasure as being good, avoid pain as bad’¹¹. Given such a hedonistic assumption, no one goes willingly toward the bad or what he believes to be bad. Weak-willed action is not possible. With the knowledge of measurement, an agent who is assumed to hold the hedonistic principle will always act rationally and wisely for he knows what is better and acts on his better judgment.

As to cases described as weakness of will, Socrates explains that in such cases the mistaken act is caused by ignorance. The agent lacks the knowledge of measurement, misestimates the quantities of pleasure and pain, and acts on his false judgment. But even so, the agent still acts according to the hedonistic principle. His action is against the actual value ranking of alternatives available to him, but is consistent with his own evaluation and with the hedonistic principle. The agent just makes a mistaken application of the principle.¹² If he had the knowledge of measurement, he would not take the mistaken act.

In *Nicomachean Ethics*, Book VII, Aristotle claims that ‘That the man who behaves incontinently does not, before he gets into this state, *think* he ought to act so, is evident.’¹³ This assertion suggests that at least before he takes the action, the incontinent agent had a judgment that it would be better for him to take another course of action available to him. Although Aristotle disagrees with Socrates’ denial of the possibility of weakness of will, he seems not to completely reject Socrates’ thought about ignorance in incontinent actions. Aristotle asks, ‘if he

¹⁰ Ibid., 352c.

¹¹ Ibid., 354c.

¹² Gerasimos Santas, ‘Plato’s *Protagoras* and Explanations of Weakness’ (1966), p. 23.

¹³ Aristotle, *Nicomachean Ethics*, 1145b30-31.

acts by reason of ignorance, what is the manner of his ignorance?’¹⁴ Whether the incontinent man acts knowingly or ignorantly is a significant problem. Only when the incontinent man acts knowingly, does what he does fit the definition of incontinent action. Therefore, Aristotle needs to explain in what sense the incontinent man acts knowingly. Furthermore, because Aristotle argues for the possibility of weakness of will, he needs to explain how the incontinent man acts against his best judgment, and what role knowledge plays in an incontinent act.

Aristotle makes several distinctions to clarify these problems. He first distinguishes two senses of the word ‘know’. One of these is having but not exercising the knowledge, and the other is having and exercising the knowledge. The former is achieved when a person has acquired the knowledge, and the latter is achieved when a person uses the knowledge. For example, a man knows that walking is good for a man. If he does not walk when he also considers that he is a man, he has but is not exercising the knowledge. If he walks when he considers that he is a man, he both has and is exercising the knowledge.

Aristotle maintains that there are two kinds of premise, universal premise and particular premise, which respectively serve as the major premise and the minor premise of a syllogism. A universal premise is like a rule or a principle. It is a proposition in the form ‘All S are P’ or ‘No S are P’. A particular premise states a particular case which falls into the type of case with which a universal premise is concerned. It is a proposition in the form ‘This S is P’ or ‘This S is not P’. A syllogism is a piece of reasoning typically consisting of two premises and a conclusion. In a syllogism, the conclusion logically follows from the two premises, which is a result of necessity. For example:

¹⁴ Ibid., 1145b29-30.

Major premise: All human beings are mortal.

Minor premise: Socrates is a human being.

Conclusion: Socrates is mortal.

Syllogism of this kind is theoretical. A theoretical syllogism is an inference about immovable objects.

A practical syllogism, a different kind of syllogism, is a reasoning about what to do. A practical syllogism usually consists of a major premise which is a universal proposition evaluating a certain type of action, a minor premise which is a proposition about a perceptibly particular action of the type with which the universal proposition is concerned, and a conclusion resulting from the two premises. A structure comprising propositions that ‘everything sweet ought to be tasted’, that ‘this is sweet’, and that ‘this ought to be tasted’ is an example of a practical syllogism. Aristotle thinks that theoretical syllogism and practical syllogism are similar in forms, but differ in outcomes. The conclusion of a theoretical syllogism is affirmed by the soul, while the conclusion of a practical syllogism is ‘the starting point of action.’¹⁵ Aristotle says:

The one opinion is universal, the other is concerned with the particular facts, and here we come to something within the sphere of perception; when a single opinion results from the two, the soul must in one type of case affirm the conclusion, while in the case of opinions concerned with production it must immediately act.¹⁶

An agent who forms a practical syllogism must take a corresponding action immediately if nothing prevents him from doing so. In the example of tasting something sweet, the agent who forms the practical syllogism must eat the particular sweet thing if nothing stands in the way.

¹⁵ Aristotle, *De Anima*, III. 10. 433a17.

¹⁶ Aristotle, *Nicomachean Ethics*, 1147a25-30.

The presence of the minor premise of a practical syllogism is a requisite step determining what an agent is going to do. Merely having a universal premise, the scientific knowledge, is not enough to move an agent to act. A relevant minor premise, the perceptual knowledge, must be present so that an agent can form a conclusion causing him to act when he considers the two premises together. For instance, an agent who holds that dry food is good for every man and that he is a man will not eat bread, as a sort of dry food, until he knows that this particular, sensible object, bread, is a sort of dry food. The minor premise that bread is dry food determines that the agent is going to eat a piece of bread.

After distinguishing the two senses of the word ‘know’ and the two kinds of premise, Aristotle explains the possibility of weakness of will in this way: a person who has both premises may use only the universal premise and not the particular, which makes his acting against his knowledge possible. Appetite, passion, or sentiment is the cause of a person’s having but not exercising the particular premise, his perceptual knowledge. Here is Aristotle’s explanation in *Nicomachean Ethics*:

When, then, the universal opinion is present in us forbidding us to taste, and there is also the opinion that ‘everything sweet is pleasant’, and that ‘this is sweet’ (now this is the opinion that is active), and when appetite happens to be present in us, the one opinion bids us avoid the object, but appetite leads us towards it (for it can move each of our bodily parts); so that it turns out that a man behaves incontinently under the influence (in a sense) of a rule and an opinion, and of one not contrary in itself, but only incidentally- for the appetite is contrary, not the opinion- to the right rule.¹⁷

This passage shows that before the incontinent man acts, he had two competing syllogisms in his mind. The syllogism forbidding him to taste may look like this:

¹⁷ Aristotle, *Nicomachean Ethics*, 1147a31-1147b3.

Universal premise: Tasting sweet is not good for human health.

Minor premise: This is sweet. I am a human being.

Conclusion: This is not good for health. I should not taste this.

The syllogism encouraging the man to taste may be as follows:

Universal premise: Everything sweet is pleasant. Human beings pursue what are pleasant.

Minor premise: This is sweet. I am a human being.

Conclusion: This is pleasant. I ought to taste this.

The conclusion of one of the syllogisms ruled out that of the other. The man judged that the former syllogism is better than the latter, and thought that he ought to act on the better judgment. However, he acts against his better judgment. The minor premise of the former syllogism which is in accordance with his reason¹⁸ is inactive when he acts because of the influence of appetite, passion, or sentiment, so no conclusion can be deduced, and no action is caused. The minor premise of the other syllogism in accordance with his appetite, on the contrary, is active when he acts. The conclusion of the syllogism is effective and causes the man to act.¹⁹ In other words, the incontinent man acts on the conclusion of the bad syllogism because the force of his appetite prevails over the force of his reason and leads him towards the worse action. Appetite makes the incontinent man temporarily unable to exercise the minor premise of the good syllogism and so he fails to take a rational action.

¹⁸ In *Nicomachean Ethics* VII.3.1147a30-1147b5, Aristotle does not explicitly say that the syllogism forbidding a man to taste sweet things is in accordance with his reason and that the incontinent man encountered the conflict between his reason and appetite before he acts. However, in *De Anima* III.10, Aristotle clearly describes the picture of conflicting desires occurring whenever reason and the appetite are opposed, which is compatible with what Aristotle says in VII.3.1147a30-1147b5: 'The desire, however, also come to be opposed to one another, which occurs whenever reason and the appetite are opposed; and this comes about in beings that have perception of time, since intellect bids us resist on account of the consequences, while appetite is on account of the immediate (for the immediate appears pleasant, and simply pleasant, and simply good, on account of not seeing the consequences).' (See Aristotle's *De Anima*, III.10.433b1-10.)

¹⁹ Similar view appears in *De Anima*, III.9.433a1-4: 'Besides, it also happens that, when intellect orders and thought says to pursue or avoid something, one is not moved but acts according to appetite, as in those lacking self-command.' (See Aristotle's *De Anima*, III.9.433a1-4.)

The incontinent man's having but not exercising the minor premise is, according to Aristotle, like a man asleep, mad, or drunk who utters sentences that come from knowledge without knowing the meaning and says things in the manner of reciting. Aristotle denies Socrates' claim that there is no such thing as incontinence, but he also asks, 'if he (the incontinent man) acts by reason of ignorance, what is the manner of his ignorance?' The answer given by Aristotle is that the incontinent man is ignorant in the sense that he either does not have or is not exercising the minor, particular premise, the perceptual knowledge, when he acts. The major, universal premise of the good syllogism is still there, what is temporarily ineffective is the minor, particular premise. In other words, the scientific knowledge is not dragged around by anything. What is dragged around by appetite, passion, or sentiment is perceptual knowledge.

Aristotle concludes:

And because the last term is not universal nor equally an object of scientific knowledge with the universal term, the position that Socrates sought to establish actually seems to result; for it is not in the presence of what is thought to be knowledge proper that the affection of incontinence arises (nor is it this that is dragged about' as a result of the state of passion), but in that of perceptual knowledge.²⁰

Aristotle wants to defend a thesis contrary to Socrates', but his account of weakness of will, especially the conclusions that the incontinent man is in the state of ignorance and that scientific knowledge is not dragged around by anything, makes many interpreters think that Aristotle is concessive to Socrates' account of weakness of will. For example, R. M. Hare remarks that Aristotle's solution is 'very Socratic',²¹ and A. W. Price also comments that Aristotle's

²⁰ Aristotle, *Nicomachean Ethics*, 1147b13-18.

²¹ R. M. Hare, 'Weakness of Will' (1992), p. 1789.

disagreement with Socrates' denial of the possibility of weakness of will is merely verbal.²²

Terence Irwin notes that Aristotle's account is different from Socrates' account in that Aristotle thinks that the incontinent man previously formed the correct decision, including the correct particular prescription against the incontinent action, but 'He retains the Socratic claim, which many regard as counterintuitive, that incontinents do not sincerely judge at the time they act incontinently that they ought not to do what they do.'²³ Aristotle shares with Socrates the claim that the incontinent do not act against their knowledge²⁴ when they act.

2. THE NATURE OF PRACTICAL REASONING

It seems that neither Plato's Socrates nor Aristotle gives a satisfactory account of incontinent actions. Socrates directly rejects their existence. Incontinent actions are impossible because agents are always guided to act by the hedonistic principle and their knowledge about what is better to do. Such a view gives evaluative judgment and action an internal and necessary connection. It is a strong version of internalism. Aristotle gives an explanation according to which the incontinent agent is not aware of both of his conflicting judgments when he acts, and so the explanation does not explain incontinent actions as they have been defined.

What Davidson is more concerned with is Aristotle's conception of practical reasoning. As we have seen, Aristotle suggests that a practical syllogism usually consists of three elements: (1) A major premise, a universal premise which is a rule or a principle about the desirability of a type of action, (2) a minor premise, a particular premise which expresses perceptual knowledge

²² A. W. Price, 'Acrasia and Self-control' (2006), p. 236, p. 243.

²³ Terence Irwin, 'Aristotle reads the *Protagoras*'(2008), p. 41.

²⁴ The knowledge here refers to the universal premise, the scientific knowledge. Irwin called it 'strict knowledge', 'full knowledge', or 'universal knowledge' which is distinguished from the perceptual knowledge (pp. 39-40).

that an action would be a case of the type of action with which the universal premise is concerned, and (3) a conclusion, which logically follows from the two premises, that the action is desirable. An agent will take a corresponding action immediately if nothing prevents him from doing so. In this sense, the conclusion of a practical syllogism is identical with an action. Aristotle's identifying drawing the conclusion with acting is not appropriate for it makes Aristotle's position inconsistent. Aristotle believes that there are incontinent actions, but this identification entails that an agent always acts on his better judgment. There is no room for cases in which an agent intentionally acts against what he judges to be better.

Furthermore, since Aristotle treats the major premise as a rule or a principle which has the form of a law, whenever an agent has a belief that an action would be a case covered by the rule, he will deduce a conclusion that the action is worth performing and perform the action. The problem is that an action, when it is considered from different aspects, could be both desirable and undesirable. If the propositional expression of a desire or wanting is a universally quantified conditional, that is, a statement that any action with a certain attribute is desirable, an agent would deduce two contradictory conclusions about the same action when he considers both the desirable and undesirable attributes of the action. For example, an object may be both sweet and poisonous.²⁵ An agent desires to eat something sweet and to avoid eating something poisonous, knows that an object is both sweet and poisonous, will conclude that the object is both desirable and undesirable. If we let '*a*' name the edible object, '*Sx*' mean that *x* is sweet, '*Px*' mean that *x* is poisonous, and '*Dx*' mean that *x* is desirable, the agent's reasoning can be expressed as follows in symbols:

²⁵ Donald Davidson, 'Intending' (1978), reprinted in *Essays on Actions and Events* (2001), PP. 97-8.

$$\forall x (Sx \rightarrow Dx)$$
$$Sa$$
$$\therefore Da$$
$$\forall x (Px \rightarrow \neg Dx)$$
$$Sa$$
$$\therefore \neg Da$$

The two conclusions are deduced from principles held true by the agent, but they constitute a contradiction ($Da \& \neg Da$) which is an obvious mistake.

To avoid the contradiction, conclusions about what to do should not be allowed to be detached directly from principles with the form of universalized conditionals. Davidson suggests that the propositional expression of a desire which provides the major premise of a practical syllogism should be a prima facie judgment, that is, a judgment that an action is desirable in so far as it has a certain characteristic.²⁶ A prima facie, evaluative judgment is a conditional judgment formed in the light of a certain property an action has, and it has the form $pf(a \text{ is better than } b, e)$. The concept of 'prima facie' here is used as an operator on pairs of sentences related as an evaluative judgment ' a is better than b ' and its ground ' e '. What is drawn from such a judgment is not a conclusion about what is desirable to do, but a prima facie judgment that doing something is desirable in relation to the reasons that an agent has considered. Accordingly, the above two syllogisms can be reconstructed in this way:

$$pf(Dx, Sx)$$
$$Sa$$
$$\therefore pf(Da, pf(Dx, Sx) \& Sa)$$

²⁶ Davidson, 'How is Weakness of the Will Possible?', pp. 37-8.

$pf(\neg Dx, Px)$

Pa

$\therefore pf(\neg Da, pf(\neg Dx, Px) \& Pa)$

When the two conclusions are put together, they do not constitute a contradiction because the two conflicting evaluative judgments are related with different sets of reasons, and the conjunction of the two conflicting prima facie judgments means that there is something to be said for doing so, and there is something to be said against doing so. Nothing contradictory is derived from desires held by an agent when the expression of a desire is a prima facie judgment.

The conditionalization that prevents conclusions drawn from prima facie, conditional judgments from forming a contradiction also insulates the conclusions from actions.²⁷ Prima facie judgments do not directly associate with actions for it is not reasonable to perform an action merely because it has a certain desirable characteristic. Before an agent acts, he usually evaluates competing reasons for different kinds of actions available to him. Because there is more than one desirable characteristic that he evaluates, multiple prima-facie evaluative judgments are involved in his consideration. These prima-facie evaluative judgments, when relevant beliefs are given, yield prima-facie evaluative judgments as conclusions. None of these conclusions can directly cause an agent to act. They merely mean that the agent, at this stage, judges that some actions available to him are desirable in terms of certain reasons so far he has considered.

For the agent to arrive at a conclusion leading to an action, he must continue to reason. The fact that an intentional action is performed represents that the agent judges that other considerations do not outweigh a desirable characteristic that the action of a type has, such a

²⁷ Davidson, 'How is Weakness of the Will Possible?', p. 39.

desirable characteristic is enough to act on.²⁸ In other words, of the prima-facie evaluative judgments that he previously considered, one of them was thought to be sufficient for acting by the agent. The agent takes this prima-facie evaluative judgment, together with a relevant belief, to form a judgment that the present action is desirable. Corresponding to the description that an action is caused by a pair of belief and desire that rationalize the action, there is a practical argument showing how the agent infers from his reason to the desirability of the action he performs. It is usually constituted by a minimum of three elements:²⁹

$$\begin{array}{l} pf(a \text{ is better than } b, r) \\ r \\ \therefore a \text{ is better than } b \end{array}$$

The agent's belief and desire provide him with premises of an argument. The major premise is a prima facie, conditional judgment. It is an evaluative proposition concerning the desirability of a type of actions classified by a certain characteristic. The minor premise comes from a belief which indicates that the present action could be a case of the type of actions in question. Its expression in an argument is the content of the belief. What is inferred, non-deductively, from the two premises is an all-out, unconditional judgment. An all-out, unconditional judgment is also an evaluative judgment, but it differs from the major premise in form- the *pf* operator is dropped out, and the judgment is not in relation to certain aspect of the action. The conclusion is an evaluation of the desirability of an actual action as a whole. It is not a conditional but has the form 'This action is desirable' or 'It would be better to do *a* than to do *b*'. It causes an action immediately if nothing stands in the way.

²⁸ Davidson, 'Intending', p. 98.

²⁹ Davidson, 'How is Weakness of the Will Possible?', p. 39.

The following case would serve to exemplify Davidson's conception of practical reasoning. Suppose a man wants to relieve the fatigues of the day, believes that the cool breeze along the bank would relieve the fatigues, and therefore goes to the riverside. The practical reasoning the man holds for acting can be so constructed to fit the suggested form:

Any act of mine would be desirable in so far as it is an act of relieving the fatigues of the day.

Going to the riverside would relieve the fatigues of the day.

Therefore, it is desirable to go to the riverside.

The minor premise is the content of the belief. If the expression of the major premise were the description of the desire or the content of the desire such as 'I want to relieve the fatigues of the day' or 'I relieve the fatigue of the day', then nothing would be inferred from these premises.

Only when the major premise is expressed by a prima-facie evaluative proposition, can the man use it to arrive at a conclusion leading to an action.³⁰

A practical argument like this is involved in a case in which an agent acts with an intention, and a case of this kind is covered by P1 and P2. P1 and P2 entail that if an agent judges that it would be better for him to do x than to do y , and he believes that he is free to do either x or y , then he will intentionally do x if he does either x or y intentionally. An agent's acting with an intention implies that there is a process of reasoning in which the agent reasons from a prima-facie, conditional judgment and a belief to an all-out, unconditional judgment. The all-out evaluative judgment, a judgment about the desirability of a present action, causes the agent to take the action. Thus, whenever an agent acts with an intention, he acts on what he judges to be better.

³⁰ Davidson, 'Intending', p. 86.

An incontinent action is done for a reason, so it is an intentional action covered by P1 and P2. An incontinent action is a consequence of a practical reasoning which ends with a conclusion about what to do. The conclusion is an all-out, unconditional judgment, say, that *a* is better than *b*. When an incontinent acts on the conclusion, he acts on what he judges to be better.

What makes the agent incontinent is his acting against the all-things-considered judgment that he holds at the same time. An all-things-considered judgment is a prima facie, evaluative judgment which is formed on the basis of ‘things known, believed, or held by the agent, the sum of his relevant principles, opinions, attitudes, and desires’.³¹ Its expression is with the form ‘*pf(a is better than b, e)*’, where *e* includes all relevant reasons that the agent has considered. Suppose *r*’ represents the agent’s all relevant considerations, including the reason *r*. The agent is incontinent when he fails to infer an all-out, unconditional judgment that *b* is better than *a* from his all-things-considered judgment that *pf(b is better than a, r)* and the reason *r*’, but forms an all-out, unconditional judgment that *a* is better than *b* from another competing prima facie judgment that *pf(a is better than b, r)* and the reason *r* instead.³² In other words, he is weak-willed when he fails to form an intention, an all-out, unconditional judgment, consistent with his all-things-considered judgment in a practical reasoning. When the agent acts on the all-out, unconditional judgment that *a* is better than *b*, he acts on his value judgment, which is covered by P1 and P2. However, he acts against his best judgment *pf(b is better than a, r)*. He acts against the principle of continence which requires him to perform the action judged best on the basis of all available relevant reasons and so is incontinent.

³¹ Davidson, ‘How is Weakness of the Will Possible?’, p. 40.

³² Donald Davidson, ‘Replies to Essays I-IX’, *Essays on Davidson: Actions and Events* (1985), p. 204.

In this account, incontinence is a failure in reasoning, but it is this failure in reasoning that prevents the incontinent agent from holding two contradictory all-out, unconditional judgments. In addition, because the incontinent agent's all-things-considered judgment and all-out, unconditional judgment have different forms, they do not contradict with each other either. The phenomenon that an agent simultaneously holds two conflicting evaluative judgments and acts on the one with a relatively weaker reason is possible as he does not hold logically contradictory judgments.

Davidson's analysis of the nature of evaluative judgment and of practical reasoning avoids problems that Plato and Aristotle encounter. Davidson distinguishes evaluative judgments into two kinds, prima-facie conditional judgments and all-out unconditional judgments. The former serve as major premises of practical arguments and do not directly associate with actions. The latter are conclusions of practical arguments and can cause actions. This analysis gives P2 a mild form of internalism. Plato thinks that there is a necessary and internal connection between evaluative judgment and action, and this strong version of internalism rules out the possibility of weakness of will. To Davidson, not all evaluative judgments but only all-out, unconditional judgments are associated with actions. Unlike Aristotle who claims that a conclusion about what to do is deduced from a desire expressed as a universally quantified conditional in a syllogism, Davidson believes that a conclusion about what to do is not deduced from a desire expressed as a prima-facie conditional judgment, but is inferred from one of the competing prima-facie conditional judgments in a further step that goes with decision and action. Because an additional step is required for an agent to arrive at an action, this step gives an agent the room to act incontinently.

The existence of incontinent actions does not threaten the causal theory of action either. An intentional action is caused by an agent's all-out, unconditional judgment which is in turn caused by a prima facie, conditional judgment and the content of a belief. Whether it is continent or not, however, depends on whether the agent follows the principle of continence to reason. Davidson's analysis shows that 'A reason that is causally strongest need not be a reason deemed by the actor to provide the strongest (best) grounds for acting.'³³ That is to say, although an all-things-considered judgment is a prima facie, conditional judgment supported by an agent's best reason and, given the causal theory of action, is supposed to have the strongest motivational force, because an all-out, unconditional judgment consistent with it cannot be deduced from it directly, it is not necessary for the agent to infer an all-out, unconditional judgment from it. An agent could make a mistake in the reasoning and be incontinent. Davidson's causal theory of action does not claim that an agent always acts on the judgment with the best reason, but merely says that whenever an agent acts, he acts on what he judges to be better. Thus, incontinent actions are not ruled out but are accommodated by P1 and P2. The three principles are a consistent triad.

3. THE CONCEPTS OF RATIONALITY AND IRRATIONALITY

As an intentional action, an incontinent action is covered by P1 and P2, which shows that it has a rational aspect which can be understood through rationalization. An incontinent action is caused by a pair of belief and desire. The pair of belief and desire provides the agent with premises of a practical argument in which the agent reasons that his action is desirable from the desire expressed as a prima facie, conditional judgment and the content of the belief. The logical

³³ Davidson, 'Introduction' of *Actions, Reasons, and Causes* (2001), xvi.

relation between the belief-desire pair and the action shows the rational aspect of the incontinent action. For any action done for a reason, there is a reason explanation of the action and a corresponding practical reasoning which exhibits rationality of the action. In the reason explanation, we attribute a pair of belief and desire to the agent to rationalize his action, and thereby attribute rationality to the agent. Thus, even though an incontinent action is an irrational action, it has a rational element at the core. Davidson claims that ‘the irrational is not merely the non-rational, which lies outside the ambit of the rational; irrationality is a failure within the house of reason.’³⁴ The purpose of the present section is to explicate Davidson’s thought that an incontinent action makes sense only when it is seen against a background of rationality. This thesis implies that P1 and P2 are necessary conditions for the existence of incontinent actions.

Davidson captures the idea of irrationality by the concept of inner inconsistency within a set of beliefs. No single belief by itself is irrational. Davidson claims that ‘It is only when beliefs are inconsistent with other beliefs according to principles held by the agent himself- in other words, only when there is an inner inconsistency- that there is a clear case of irrationality.’³⁵ Accordingly, an agent is irrational not because he fails to hold a belief or attitude that others deem reasonable but because he fails, within himself, to present a coherent or consistent pattern of beliefs, attitudes, emotions, intentions, and actions.

As an irrational action, an incontinent action exhibits an inner inconsistency within the agent. In an incontinent action, although the agent holds an all-out, unconditional judgment not in accordance with his all-things-considered judgment, the two value judgments do not contradict with each other because they are different in form. Only when the agent also holds the

³⁴ Davidson, ‘Paradoxes of Irrationality’, p. 169.

³⁵ Donald Davidson, ‘Incoherence and Irrationality’ (1985), reprinted in *Problems of Rationality* (2004), p. 192.

principle of continence which requires him to act on his best judgment, does the agent hold inconsistent beliefs. The principle of continence is a second-order principle, a belief that 'I ought to act on my own best judgment'.³⁶ Because of holding the principle of continence, an incontinent agent's all-out, unconditional judgment is then literally inconsistent with his all-things-considered judgment enjoined by the principle of continence.³⁷ Thus, the occurrence of irrationality is an inconsistency within three beliefs which are active simultaneously: a judgment that the present action is desirable, a judgment that all things considered, another available course of action is better, and the principle of continence.

Davidson takes the principle of continence to be a standard to evaluate whether a man is rational or not. A man who follows the principle of continence to form an intention in accordance with his best judgment has consistent beliefs and is therefore a rational agent. Otherwise, he is irrational. Such a principle is inherent in an agent and so he either holds consistent or inconsistent beliefs when he reasons to act, and can be seen as rational or irrational. But should we accept that the principle of continence is inherent in an agent? Davidson considers this question from an interpreter's perspective in understanding an action.

According to Davidson, to have a language is to be able to speak and to interpret utterances from other speakers. Merely being able to speak is not sufficient to have a language. A speaker who is unable to interpret utterances from other speakers cannot construct a theory of meaning through which he grasps the meaning of a sentence by knowing its truth condition. A speaker who does not understand the meaning of utterances from other speakers cannot

³⁶ Davidson, 'Paradoxes of Irrationality', p. 177.

³⁷ Davidson, 'Incoherence and Irrationality', p. 194.

communicate with others and is not qualified to be a speaker of a language. Thus, a speaker must himself be an interpreter of others.³⁸

When we try to understand an action by attributing a pair of belief and desire that rationalize the action to the agent, we give his belief and desire an interpretation. In the interpretation, we know the meaning of the belief and desire by knowing their truth conditions. Beliefs are cognitive states whose propositional expressions are true, false, or neither. Desires, conative states, are not truth-evaluable when they are represented as wants with descriptive content, such as ‘wanting to go jogging’. Desires, however, are truth-evaluable when they are represented by sentences with explicitly evaluative terms, such as ‘thinking it desirable to go jogging’. When an agent gives a desire as his reason for taking an action, or when we attribute a desire to an agent to understand his action, what is used is an explicitly evaluative sentence which is truth-evaluable.³⁹ Because ‘An utterance has certain truth conditions only if the speaker intends it to be interpreted as having those truth conditions,’⁴⁰ in order to know what the agent intends, believes, and desires, we probably need to construct a theory of the agent’s thoughts,

³⁸ Donald Davidson, ‘Thought and Talk’ (1975), reprinted in *Inquiries into Truth and Interpretation* (2009), p. 157.

³⁹ Desires, conative states, are generally thought to be non-truth-evaluable, but through Davidson’s suggestions that the natural expression of a desire is evaluative in form (See Davidson’s ‘Intending’, p. 86) and that value judgments are objective and have truth values (See Davidson’s ‘The Objectivity of Values’, p. 39), desires are also expressible in evaluative sentences which can be true or false. The conative way of representing evaluation is more basic than the cognitive way of representing evaluations in that an interpreter first needs to identify a speaker’s desire in an interpretation. The shift from representing evaluations in the conative way to the cognitive way, however, is a necessary step in the interpretation. John Brice elaborates this point in his paper ‘Hume and Davidson: Passion, Evaluation, and Truth’ (2011): ‘The radical interpreter’s move from basic to non-basic ways of representing evaluations is ineluctable, however, given the necessary expressibility, in truth-evaluable evaluative sentences, of states characterized as desires and emotions. While eschewing, at the start, the interpretation of sentences that prove (upon interpretation) to be evaluative ones, the interpreter must eventually turn to the evaluative sentences needed if the subject is to give expression to her reasons for acting and feeling as she does. To do that is to pair desires and emotions with the holding true of the evaluative sentences that express them. And to do THAT is to introduce the cognitive way of representing evaluations that are already in place’ (p. 13). When a speaker’s desire serves as a reason to explain his action or a premise of a practical argument, his desire is represented in an expression with an explicitly evaluative term.

⁴⁰ Davidson, ‘The Content of the Concept of Truth’ (2005), p. 50.

emotions, and language. Due to the multiplicity of mental factors that produce behavior and speech, the prospect of constructing such a theory is unlikely.

Davidson suggests that what makes an interpretation work is simply the assumption that ‘the person to be understood is much like ourselves.’⁴¹ We, the interpreters, assume that the agent has beliefs and values similar to us, that the agent, like us, supposes those he wants to understand live in a world having people with minds and motives, and that the agent also shares with us the desire to find warmth, love, security, success, and the desire to avoid pain and distress. We take most beliefs the agent has, like our beliefs, to be true. We also believe that the agent, like ourselves, reasons according to logical rules and acts on a judgment with the best reason. Davidson claims that ‘unless we can interpret others as sharing a vast amount of what makes up our common sense we will not be able to identify any of their beliefs and desires and intentions, any of their propositional attitudes.’⁴² Once we assume that the agent’s mind and ours have many things in common, there is a shared background of rationality which allows successful communications and interpretations.⁴³

⁴¹ Davidson, ‘Paradoxes of Irrationality’, p. 183.

⁴² Ibid.

⁴³ Here I merely talk about the assumption of rationality in order to emphasize the significance of an agent’s built-in rationality in interpreting his utterances and understanding his action. Davidson’s theory of meaning is much more complex than what is mentioned in this section. The theory has two constraints, formal and empirical. On the one hand, the theory is purely extensional. The theory consists of axioms for reference conditions on names and terms, satisfaction conditions for predicates and relations, combination rules, and theorems with the form ‘*s* is true in L if and only if *p*’, where *s* is replaced by a sentence in an object language L and *p* is replaced by a meta-language sentence which translates *s* and describes a state of affairs obtaining which renders *s* true. On the other hand, these theorems ‘should optimally fit evidence about sentences held true by native speakers’ (Davidson, ‘Radical Interpretation’, p. 139). A radical interpreter does not know the meaning of any of the speaker’s terms and of the speaker’s propositional attitudes, and he determines whether a theorem is correct or not by observing whether a native speaker takes an utterance to be true. When he observes that native speakers assent to the truth of an utterance at different time and locations, the evidence confirms the correctness of a theorem. All theorems constituting the theory are required to be empirically justified in similar situations. Because ‘to give truth condition is a way of giving the meaning of a sentence’ (Davidson, ‘Truth and Meaning’, p. 24), when the two constraints are satisfied by a meaning theory, a person who knows the theory can interpret sentences.

The principle of continence is a principle of rationality shared by both the speaker and the interpreter in an interpretation. When we assume that an agent who entertains propositional attitudes and performs intentional actions is in general rational, i.e., he holds and follows such a principle on most occasions, an intentional action can be understood as part of a rational pattern comprising various attitudes logically connected with one another, and an action that is incontinent can be regarded as an aberration against the background of rationality. If we do not believe that an agent exhibits rationality to some degree whenever he reasons or acts, none of his utterances or behaviors could be intelligible. As Davidson puts it, 'The essential point is that the more flamboyant the irrationality we ascribe to an agent, the less clear it is how to describe any of his attitudes, whether deviant or not, and that the more basic we take a norm to be, the less it is an empirical question whether the agent's thought and behavior are in accord with it.'⁴⁴ We are in a better position to interpret an agent's reason for his action when we assume that he holds the principle of continence.

Another common principle of rationality relevant to the issue in question is the principle of logic which governs the relations among propositional attitudes. The principle of logic gives rules according to which propositional attitudes can be identified by their causal relations to

An interpreter tries to understand why an agent does something by knowing the belief and desire that cause him to do so. In a radical interpretation, the interpreter is assumed not to know the meaning of the agent's belief and desire. The only thing open to observation is the agent's ordinal preferences or choices among options, i.e., the agent's preference that one sentence representing his desire rather than another is true. The interpreter who knows the agent's preference with respect to the truth of sentences knows what the agent takes the sentences to mean. Once he knows the meaning of the desire causing the agent to act, he knows the meaning of the paired belief for the fact that the agent chooses to take the action entails that the agent prefers a possible consequence of an action and believes that the possible consequence will be achieved by performing the action. The interpreter attributes a pair of belief and desire expressed in a cognitive way to the agent to explain his action, and the assumption of rationality plays a role in the attribution: '... the relations between beliefs play a decisive constitutive role; we cannot accept great or obvious deviations from rationality without threatening the intelligibility of our attributions. If we are going to understand the speech or actions of another person, we must suppose that their beliefs are incorporated in a pattern that is in essential respects like the pattern of our own beliefs' (Davidson, 'A Unified Theory of Thought, Meaning, and Action', pp. 156-7).

⁴⁴ Davidson, 'Incoherence and Irrationality', p. 196.

events and objects in the world and by their logical relations to one another.⁴⁵ A man who perceives the maple leaves gradually turn into yellow, red, and brown believes that autumn approaches. A belief that the room he stays in is on fire would cause the man who has the belief to desire to escape from the room. Although the various sorts of thought may not be reduced to one or a few sorts, belief is central to all kinds of thought. Having a certain kind of thought always requires having relevant beliefs as its background. For example, someone who knows that a big rabbit called German giant weighs 22 pounds and measures over 3 feet must have beliefs about the nature of rabbits, about the comparative concepts such as heavy and tall, and many others. The person may not have a particular list of beliefs to support the thought, but there must be endless interlocked beliefs. If an agent, for instance, knows that 'P and Q' is true, then he knows that 'P' is true as well. Or if he accepts 'P' to be true, then he will not accept '-P' to be true. We identify an agent's belief 'by locating it in a logical and epistemic space',⁴⁶ and all his other thoughts must be identified in the same way. An agent cannot have a thought which is single, isolated, not logically connected with other thoughts, and escapes from the government of the principle of logic for 'the identity of a given thought depends in part on its relations to other thoughts.'⁴⁷ An agent whose action is a consequence of a reasoning among his propositional attitudes must, like us, hold the principle of logic.

Besides the principle of continence and the principle of logic, there are other principles of rationality that are norms shared by all thinking creatures: Carnap's principle of total evidence which 'counsels an agent to accept the hypothesis supported by the totality of evidence he or she

⁴⁵ Ibid., pp. 195-6.

⁴⁶ Davidson, 'Thought and Talk', p. 157.

⁴⁷ Davidson, 'The Problem of Objectivity' (1995), reprinted in *Problems of Rationality* (2004), p. 11.

has,⁴⁸ Harman's inference to the best explanation, Quine's principle of conservation that 'other things being equal, change as few expectations as possible when accommodating recalcitrant appearances,'⁴⁹ and other general principles. Because of having these principles of rationality, an agent who reasons against any of these principles holds inconsistent beliefs and is irrational.

Some may disagree with Davidson's allegation that principles of rationality are norms inherent in thinking creatures. If these principles are not necessarily held by an agent, an agent is not in a state of inner inconsistency when he disobeys any principles. However, Davidson argues that rationality is a condition of having thoughts at all. A thinking creature does not decide whether he wants to accept principles of rationality or not but is determined to have these fundamental attributes of rationality. For once he starts considering anything, he needs to follow the principle of logic for any of this thought to make sense, and he needs to follow other principles of rationality to reason, theoretically or practically, so that his reasoning is intelligible. He is then displaying those attributes of rationality.⁵⁰

With the assumption that an agent necessarily holds principles of rationality and he reasons and acts largely in accord with these principles, we can attribute propositional attitudes to rationalize his thought and action, and understand his thought and action by placing them in a rational pattern. The principles of rationality also serve as criteria according to which we judge whether an agent is rational or not. Because of having the attributes of rationality, that an agent occasionally reasons or acts against principles of rationality and holds inconsistent beliefs can be understood as a departure from these standards of rationality. Davidson remarks that 'The

⁴⁸ Davidson, 'Incoherence and Irrationality', p. 190.

⁴⁹ Ibid., p. 192.

⁵⁰ Davidson, 'Incoherence and Irrationality', pp. 196-7.

possibility of inconsistency depends on nothing more than this, that an agent, a creature with propositional attitudes, must show much consistency in his thought and action, and in this sense have the fundamental values of rationality; yet he may depart from these, his own, norms.⁵¹ In other words, rationality is a prerequisite of irrationality. An incontinent action as a temporary aberration has a rational element at the core. Thus, incontinent actions are not only covered by P1 and P2 which entail rationality and apply to all intentional actions, but must have P1 and P2 as necessary conditions for their existence.

4. THE THEORY OF PARTITIONED MIND

An incontinent action, like all other intentional actions, can be described as an effect caused by a pair of belief and desire that rationalizes the action, but a reason explanation of this kind does not rationalize an incontinent agent's incontinence. The irrational aspect of an incontinent action is a part that cannot be figured out by placing the action in a rational pattern. A reason explanation reveals why the incontinent agent takes the action, but why the agent acts against his own principle of continence remains unknown. To explain why an incontinent agent does not act otherwise, why he fails to form an intention consistent with his best judgment, Davidson develops his theory of partitioned mind to identify the mental cause of an agent's incontinence.

Davidson uses the following example to illustrate the mental cause of incontinence:

A man walking in a park stumbles on a branch in the path. Thinking that the branch may endanger others, he picks it up and throws it in a hedge beside the path. On his way home it occurs to him

⁵¹ Ibid., p. 197.

that the branch may be projecting from the hedge and so still be a threat to unwary walkers. He gets off the tram he is on, returns to the park and restores the branch to its original position.⁵² ... It is easy to imagine that the man who returned to the park to restore the branch to its original position in the path realizes that his action is not sensible. He has a motive for moving the stick, namely, that it may endanger a passer-by. But he also has a motive for not returning, the latter consideration outweighs the former; yet he acts on the former. In short, he goes against his own best judgment.⁵³

The man has reasons for his throwing the branch away and for his returning to the park to restore the branch to its original position, and so corresponding reason explanations can be produced to rationalize both of his actions. When he returns to the park, however, he also has a judgment that, all things considered, it would be better not to return to the park. Thus, when he returns to the park, he goes against his own principle of continence. How do we explain his incontinence?

Davidson replies that,

... And there is no denying that he has a motive for ignoring his principle, namely that he wants, perhaps very strongly, to return the branch to its original position. Let us say this motive does explain the fact that he fails to act on his principle.⁵⁴

Accordingly, what causes his incontinence is the desire to replace the branch. In this case, the desire plays two roles. On the one hand, the desire serves as a reason for the agent's returning to the park to replace the branch. It is a propositional attitude which bears appropriate logical relations with other propositional attitudes and with those mental states or events causing it or caused by it. The desire, on the other hand, causes the agent to ignore the principle of continence but does not rationalize his ignoring it. The desire and the agent's ignoring his principle of continence are in a causal relation, but the logical relation between the two is either missing or

⁵² Davidson, 'Paradoxes of Irrationality', p. 172.

⁵³ Ibid., p. 174.

⁵⁴ Ibid., p. 178.

distorted. In other words, the desire is a mental cause that is not a reason for his being incontinent.

This account of the cause of incontinence raises a problem. Given the principle of logic and Davidson's holistic picture of the mental, a propositional attitude is supposed to retain logical relations with those that are causally related with it. How can the strong desire exist as a mental cause that bears no rational relation with its effect? The strong desire causes the incontinent agent to form an intention to perform the present action, which is inconsistent with his all-things-considered judgment and the principle of continence. How can the agent knowingly harbor inconsistent beliefs in his mind?

The existence of a mental cause that is not a reason for its effect and the state of inner inconsistency are possible when the mind is partitioned into parts. According to Davidson, whenever an agent holds beliefs that are inconsistent, his mind can be regarded as having two or more semi-autonomous subdivisions. The structure of each subdivision of the mind contains various propositional attitudes and mental states that are causally and logically connected with one another. The consequence of an interaction among propositional attitudes in a subdivision can be rationalized by propositional attitudes in the subdivision that causes it. When a propositional attitude cannot be rationalized by the propositional attitude that causes it, the two propositional attitudes are separated into two subdivisions each of which is supported by a set of interlocking propositional attitudes. Davidson explains, 'The breakdown of reason relations defines the boundary of a subdivision... The parts are defined in terms of function; ultimately, in terms of the concepts of reason and of cause.'⁵⁵ Propositional attitudes in a subdivision are in

⁵⁵ Ibid., p. 185.

causal and logical relations, but between two propositional attitudes that are separated into two subdivisions, there is a non-logical causal relation. Whenever there are inconsistent beliefs, such a boundary is established to separate beliefs into different parts so that they can simultaneously exist in the mind.

Davidson clarifies that boundaries of this kind are not to define permanent and separate territories but are ‘conceptual aids to the coherent description of genuine irrationalities.’⁵⁶ A boundary that separates inconsistent beliefs into different parts of the mind is not something that exists out there and can be discovered by introspection. It is merely an idea used to depict the image of a circumstance in which an agent’s holding inconsistent beliefs is possible. Since an agent is irrational when he holds inconsistent beliefs, Davidson remarks that ‘the irrational step is therefore the step that makes this possible, the drawing of the boundary that keeps the inconsistent beliefs apart.’⁵⁷ When one of two inconsistent beliefs is destroyed, the line between the two is then erased.

According to this idea of partitioned mind, we can imagine that the inconsistent judgments that an incontinent agent has are separated into different parts of the mind as if there is a boundary dividing them. On the side of all-things-considered judgment, there are interlocking beliefs, desires, expectations, assumptions, and attitudes supporting the sober judgment. As a belief, the principle of continence also belongs to this part. On the side of incontinent intent and action, there is a similar supporting structure of reasons. The principle of continence is supposed to govern the interactions between any set of propositional attitudes, but the incontinent agent’s desire to take the present action causes the principle of continence to be

⁵⁶ Donald Davidson, ‘Deception and Division’ (1986), reprinted in *Problems of Rationality* (2004), p. 211.

⁵⁷ Ibid.

temporarily isolated and ineffective in the course of forming the incontinent intention and performing the incontinent action, which causes the state of inner inconsistency within the agent. The compartmentalization of the mind explains the existence of the cause of incontinence, a mental cause that is not a reason for the mental state it causes, and explains the possibility of incontinence, a case of inner inconsistency.

5. CONCLUSION

We have seen how Davidson solves the two problems of weakness of will in this chapter. Through the solutions, several concepts crucial to Davidson's analysis of the phenomenon of weakness of will are revealed: the nature of practical reasoning, the nature of evaluative judgments, the interdependent relations among reason, rationality, and irrationality, a mental cause that is not a reason for its effect, and the compartmentalization of the mind.

Davidson solves the first problem, how incontinent actions can exist given the principle that an agent always acts on what he imagines or judges to be better, by giving a description in which the existence of incontinent actions is consistent with his causal theory of action which endorses the principle. The point is that what causes an intentional action, continent or not, is an all-out, unconditional judgment. An all-out, unconditional judgment is an evaluation of competing reasons for an action and is of the form '*a* is better than *b*'. When an agent acts on his all-out, unconditional judgment, he acts on what he judges to be better. In the practical reasoning, the agent also holds an all-things-considered judgment which is a prima facie, conditional judgment established on all reasons that the agent has considered. A prima facie, conditional judgment is not directly associated with an action, but can be a major premise of an argument.

Whether or not the agent acts continently is determined by whether the all-out, unconditional judgment is inferred from his all-things-considered judgment. A continent agent who follows his principle of continence to act detaches an all-out, unconditional judgment, say, that a is better than b , from his all-things-considered judgment that $pf(a$ is better than $b, e)$ in a practical argument. An incontinent agent, in contrast, fails to follow his principle of continence but detaches an all-out, unconditional judgment that b is better than a from another competing prima facie, conditional judgment that $pf(b$ is better than $a, r)$. Incontinence is a failure in reasoning. In this description, the incontinent agent acts against his best judgment, but his action is caused by his all-out, unconditional judgment and so is compatible with the principle that an agent always acts on what he imagines or judges to be better and with the causal theory of action.

An incontinent action, as a consequence of a practical reasoning, has a rational aspect which can be explained by giving a reason, but it is an irrational action because it is against the principle of continence, a principle of rationality inherent in an agent. An agent is assumed to frequently follow the principle of continence and other principles of rationality to reason. This assumption makes most of his actions intelligible to us, and allows us to understand his irrational actions which occasionally occur as deviations from the principle of continence. The principle of continence is a criterion to evaluate whether an agent is rational or not. Because the principle of continence is a built-in belief, an incontinent agent who holds an all-out, unconditional judgment that is inconsistent with his all-things-considered judgment and his principle of continence holds a set of inconsistent beliefs. Incontinence, in this sense, is a case of inner inconsistency.

To the second problem of weakness of will, why an incontinent agent didn't act otherwise, Davidson identifies the strong desire which causes an incontinent agent to form an intention to

take the present action as the cause of his incontinence. The strong desire causes the incontinent agent to ignore the principle of continence, but does not rationalize his ignoring it. It is a mental cause that is not a reason for the mental state that it causes. When the intention to perform the present action is formed, the incontinent agent holds an inconsistent set of beliefs. Given the assumption that no mental event or state can exist without logically connecting to other mental events or states, Davidson explains that a mental cause that is not a reason for its effect and the state of inner inconsistency occur when inconsistent propositional attitudes or mental states involved are separated into different parts of the mind. Each of the inconsistent propositional attitudes or mental states is supported by a set of interlocking mental events and states which constitute a part, but between them there is only a non-logical causal relation. In this theory of partitioned mind, that a part of the mind is a semi-autonomous subdivision preserves the truth of the assumption that a propositional attitude or mental state must be logically connected with other propositional attitudes and mental states, and the non-logical causal relation between two subdivisions allows the existence of an inconsistency between two mental events or states.

III. INCONTINENCE AS A FAILURE IN REASONING

Donald Davidson explains the phenomenon of weakness of will as a failure in reasoning. In a case of incontinent action, the incontinent agent considers competing reasons for or against different courses of actions. Although he holds the principle of continence which requires him to perform the action judged best on the basis of all available relevant reasons, he does not infer an all-out, unconditional judgment, an intention to do something, from his all-things-considered judgment which is supported by the best reason. Instead, he detaches an all-out, unconditional judgment from the other competing prima-facie, conditional judgment that is formed in relation to a relatively weaker reason. The incontinent agent fails to follow the principle of continence and draws a false conclusion about what to do in a practical reasoning. His incontinence is a failure in reasoning.

Davidson's account of the phenomenon of weakness of will is challenged by many commentators, and in this chapter I discuss several lines of argument against the account. These lines of argument include problems about the possibility of the failure to reason from an all-things-considered judgment to a consistent unconditional judgment, the cause of an incontinent agent's failure in reasoning, and the analogy between incontinence and probabilistic akrasia, another kind of failure in reasoning. I defend Davidson's account of incontinence against these criticisms. The defense focuses on the clarification of the role the principle of continence plays in practical reasoning and on the application of Davidson's anomalism thesis that there are no psychological or psychophysical laws on the basis of which mental events can be predicted and explained.

This chapter is structured as follows. Section 1 briefly revisits Davidson's account of the

phenomenon of weakness of will. Section 2 discusses the distinction between two kinds of evaluative judgment, all-things-considered judgment and all-out, unconditional judgment. I will give an analysis of the relations between modus ponens, the principle of continence, and the principle of total evidence to explain the non-implication relation between the two kinds of evaluative judgment. Section 3 investigates why the desire with an insufficient reason is motivationally stronger in a case of incontinent action. I will argue that no mental state of a specified kind can be identified to explain the stronger motivational strength of the desire, but an incontinent agent's emotion, personal character, custom, or physical impulse could interfere with his practical reasoning and cause him to draw a false conclusion about what to do. Section 4 explores the possibility of probabilistic akrasia. I will argue that probabilistic akrasia is possible because conflicting probabilistic judgments can both be true to an agent and it occurs when, say, strong faiths, intuitions, or prejudices intervene in an agent's inductive reasoning and cause him to draw a false conclusion about what is more likely to be true.

1. INCONTINENCE AS A FAILURE IN REASONING

Davidson's general view on intentional action can be sketched as follows. According to Davidson's causal theory of action, an intentional action is caused by a pair of belief and desire that rationalizes the action. As a cause, the pair of belief and desire ensures that the action caused by it is intentionally done by the agent. As a reason, the pair of belief and desire explains why the agent takes the action. Corresponding to such an action-explanation, a practical argument can be construed to show the logical relation between his reason and action. The pair of belief and desire provides the agent with premises of the argument. The premise provided by the desire is a *prima facie*, conditional judgment that an action of a kind is desirable in so far as it has a certain

characteristic. The premise provided by a belief is the content of the belief, a proposition that his action is of that kind. The conclusion inferred from the two premises is an all-out, unconditional judgment that taking the present action is desirable. Because the intentional action that the agent performs is a consequence of the practical argument, the logical relation between the intentional action and its reason shown in the argument reveals the rational aspect of the intentional action.

An incontinent action is an intentional action covered by the causal theory of action. An incontinent agent takes his action for a reason, and so an aspect of the action can be explained by giving the reason that causes the action. When an incontinent agent deliberates about what to do, he considers reasons for and against certain kinds of action. The competing considerations are *prima facie*, conditional judgments with the form '*pf(a is better than b, r)*', where *r* is a reason in relation to which an action of a kind is judged to be better than an action of the other kind. When the incontinent agent takes one of the competing *prima facie*, conditional judgments as a sufficient reason to act, he detaches an all-out, unconditional judgment from it. An all-out, unconditional judgment is with the form '*a is better than b*'. When the incontinent agent acts on his all-out, unconditional judgment, he acts on his evaluative judgment. The *prima facie*, conditional judgment from which the incontinent agent detaches the all-out, unconditional judgment, the all-out, unconditional judgment, and the action are in causal and logical relations. Hence, they show the rational aspect of the incontinent action and the compatibility between the incontinent action and the causal theory of action.

What makes the incontinent agent wrong is that, in the reasoning process, he reasons against the principle of continence which requires him to perform the action judged best on the basis of all available relevant reasons, and fails to detach an all-out, unconditional judgment from

his all-things-considered judgment, i.e., a prima facie, conditional judgment formed in relation to all of his relevant considerations. Davidson denies that a prima facie, conditional judgment logically implies an unconditional judgment. An all-out, unconditional judgment, an intention to do something, cannot be deduced from a prima facie, conditional judgment because contradictory conclusions could be formed from conflicting prima facie, conditional judgments concerning different aspects of an action. Only when the agent takes one of the competing prima facie, conditional judgments to be a sufficient reason for acting, can he detach an all-out, unconditional judgment from it. Since there is no logical implication between the two kinds of evaluative judgments, it is not necessary for an agent to detach an all-out, unconditional judgment from his all-things-considered judgment. Incontinence occurs when an agent fails to infer from his all-things-considered judgment that $pf(b \text{ is better than } a, e)$ and all the relevant considerations e to an all-out, unconditional judgment that b is better than a , but detaches an all-out, unconditional judgment that a is better than b from the other competing prima facie, conditional judgment that $pf(a \text{ is better than } b, r)$ and the reason r . In other words, incontinence is a failure in reasoning, a failure to draw a conclusion about what to do from an agent's best judgment.

Davidson believes that the similarity between weakness of will and probabilistic akrasia would help understand incontinence as a failure in reasoning. In an inductive reasoning, an agent considers competing probabilistic judgments which are established in relation to evidence. A conclusion about what is almost certainly to occur cannot be detached directly from any probabilistic judgment because a contradiction could be yielded from two conflicting probabilistic judgments supported by equal evidence. It is in a further step that an agent detaches a conclusion about what is almost certainly to occur from one of the probabilistic judgments that

he considers. An agent who holds the principle of total evidence knows that he should give his credence to the hypothesis supported by all available relevant evidence.¹ An irrational agent, however, reasons against his principle of total evidence and fails to detach a conclusion about what is almost certainly to occur from the probabilistic judgment supported by the strongest evidence. Instead, he detaches a conclusion about what is almost certainly to occur from a probabilistic judgment supported by less evidence. Accordingly, incontinence is parallel to probabilistic akrasia in that to each kind of reasoning, the major premise is a relational judgment, that the major premise and the minor premise together may not yield a corresponding conclusion, and that the agent fails to follow a principle of rationality and draws a false conclusion.

2. TWO KINDS OF EVALUATIVE JUDGMENT

Whether an all-things-considered judgment can be separated from a consistent all-out, unconditional judgment in practical reasoning is controversial. Paul Grice and Judith Baker, Christopher Peacocke, and William Charlton do not deny that there is a distinction between prima facie, conditional judgment and all-out, unconditional judgment. However, they seem to think that the existence of an all-things-considered judgment logically implies the existence of an all-out, unconditional judgment in accordance with it. An agent who holds an all-things-considered judgment in favor of *a* can not fail but is compelled to form an unconditional judgment in favor of *a*. Since the occurrence of an all-things-considered judgment necessarily involves that of an unconditional judgment consistent with it, explaining incontinence as a failure in reasoning cannot be successful.

¹ Donald Davidson, 'How is Weakness of the Will Possible?' (1969), reprinted in *Essays on Actions and Events* (2001), p. 41.

In this section, I consider arguments against the non-implication relation between all-things-considered judgment and unconditional judgment, and defend the thesis that it is possible for an agent to hold an all-things-considered judgment without holding an all-out, unconditional judgment consistent with it. I clarify that because the inference from a reason to an intention is not an automatic transition, an entailment, and practical arguments arriving at unconditional judgments are not deductive arguments, an agent's failure in reasoning is not impossible. What is crucial to this possibility is the role the principle of continence plays in practical reasoning. Some commentators claim that the principle of continence is analogous to modus ponens and necessarily compels an agent who holds it to reason consistently. I investigate the suggested analogy and conclude that whether or not there is an analogy between the principle of continence and modus ponens, incontinence as a failure in reasoning is possible.

2.1 Arguments against the non-implication relation between two kinds of evaluative judgment

According to Grice and Baker, the most plausible interpretation of all-things-considered judgment is to consider it as one consisting of an 'all things before me' judgment together with an 'optimality' judgment about it, i.e., 'given the fact that on my evidence I should do *A*, and that my judgment that this is so is optimal, I should judge that I should do *A*'.² Given the normative force of logic, however, this interpretation entails that an all-things-considered judgment that one should do *a* is not distinct from an unconditional judgment that one should do *a*. Grice and Baker argue that

² Paul Grice and Judith Baker, 'Davidson on 'Weakness of Will'' (1985), p. 38, p. 46.

... extreme logical confusion apart, one judges on what seems to one an adequate basis that one should judge that one should do *a* if and only if one judges on the same basis that one should do *a*. That is to say, any case of believing, on what seems to one an adequate basis, that one should do *a* is, logical confusion apart, to be regarded as also a case of believing that one's belief that one should do *a* is a belief which one should hold: and any case of believing, on what seems to one an adequate basis, that one should believe that one should do *a* is, logical confusion apart, to be regarded as a case of believing that one should do *a*.³

In other words, an agent who holds an all-things-considered judgment favoring *a* necessarily involves an unconditional judgment consistent with it. This interpretation of all-things-considered judgment excludes the possibility that an agent can hold an all-things-considered judgment without detaching an unconditional judgment from it.

Davidson believes that an agent who acts incontinently commits no logical blunder because his all-things-considered judgment and unconditional judgment differ in form. Although they are inconsistent, they are not contradictory, and so an incontinent agent is allowed to hold them at the same time.

Grice and Baker, on the contrary, claim that an agent who is fully conscious of his all-things-considered judgment favoring *a* but makes an unconditional judgment favoring *b* is holding a contradiction. Given the above interpretation of all-things-considered judgment, unless an agent disobeys rules of inference, he cannot avoid concluding an unconditional judgment favoring *a* when he holds an all-things-considered judgment that favors *a*. In this sense, an all-things-considered judgment favoring *a* is indistinct from an unconditional judgment favoring *a*. Since to attribute an all-things-considered judgment in favor of *a* to an agent is to attribute him an unconditional judgment in favor of *a*, an incontinent agent who holds both (1) an all-things-

³ Ibid., 47.

considered judgment that $pf(a$ is better than $b, e)$, which is indistinct from an unconditional judgment that a is better than b , and (2) an unconditional judgment that b is better than a , in fact holds two contradictory unconditional judgments. Thus, Davidson's characterization of incontinence is not successful for no one can hold an obvious contradiction.

Grice and Baker propose an analogy between modus ponens and the principle of continence to explain why an all-things-considered judgment is indistinct from an unconditional judgment consistent with it. They indicate that it is modus ponens that makes ' q ' necessarily follow from 'given the fact that p and that if p, q '. Analogously, the principle of continence which requires an agent to perform the action judged best on the basis of all available relevant reasons must make an agent who holds an all-things-considered judgment that $pf(a$ is better than $b, e)$ and e form an unconditional judgment that a is better than b . An agent who refuses to move from an all-things-considered judgment that $pf(a$ is better than $b, e)$ and e to an unconditional judgment that a is better than b is analogous to an agent who holds 'if p , then q ' and ' p ', but refuses to conclude ' q '. The irrationality involved in the two cases is the same in kind.⁴

Charlton holds a similar view that an agent who holds an all-things-considered judgment favoring ϕ is required to hold an unconditional judgment favoring ϕ :

If I am to think it best, all things considered, to ϕ I must think not merely that all the facts known to me on balance favor ϕ ing but that they collectively warrant ϕ ing or even demand it. What does this further condition involve? I must think either that the facts make it best to judge, without further enquiry, that it is best sans phrase to ϕ , or simply that they make it best to ϕ without further enquiry. In either case it seems I must judge that it is best sans phrase to ϕ .⁵

⁴ Ibid., 40.

⁵ William Charlton, 'Davidson and his Critics' (1988), p. 124.

Accordingly, since an all-things-considered judgment in favor of ϕ ing is conditioned by all relevant reasons an agent can think of, there is no further inquiry concerning ϕ ing that he needs to make. The reasons that the agent has will force him to judge unconditionally that it would be better to ϕ . The reasons that the agent has makes it the case that there is an implication between an all-things-considered judgment favoring ϕ ing and an unconditional judgment favoring ϕ ing.

Peacocke seems to think that an all-things-considered judgment that $pf(a$ is better than $b, e)$ can exist without an unconditional judgment that a is better than b only when the reason e is not at present available to the agent and so the agent fails to detach an unconditional judgment that a is better than b .⁶ Describing an incontinent agent as someone who holds an all-things-considered judgment $pf(a$ is better than $b, e)$ and e but fails to detach an unconditional judgment a is better than b is not acceptable.

A reason that I think would support the view that there is an implication between an all-things-considered judgment favoring a and an unconditional judgment favoring a is that Davidson takes the operator ' pf ' as a sentential connective and treats all-things-considered judgment as a special case of prima facie, conditional judgments. Davidson also points out that a prima facie judgment can be paraphrased as 'if so and so is (or were) the case...'.⁷ Given these assertions, a prima facie judgment that $pf(a$ is better than $b, r)$ can be understood as a proposition with the form 'If ... then'. The if-clause states a reason that, if true, supports the main clause, a judgment about the desirability of taking an action of a kind. If an all-things-considered judgment that $pf(a$ is better than $b, e)$ is understood as 'if e , then a is better than b ', then when an agent also holds ' e ', it is natural to think that he will infer that ' a is better than b '. In this view,

⁶ Christopher Peacocke, 'Intention and Akrasia' (1985), p. 61.

⁷ Donald Davidson, 'Replies to Essays I-IX', *Essays on Davidson: Actions and Events* (1985), p. 203.

an all-things-considered judgment entails an all-out judgment consistent with it. The analogy between the formation of this practical argument and the formation of a deductive argument consisting of the three propositions ‘If p , then q ’, ‘ p ’, and ‘ q ’ is obvious. Both arguments are construed as arguments with the structure of modus ponens, and the conclusions must be detached when the premises are given.

2.2 The principle of continence

In a reply to Grice and Baker, Davidson mentions that there is an analogy between the principle of continence and the principle of total evidence:

But the principle of continence is not a principle of logic. As Hempel says of the ‘requirement of total evidence’ on which the principle of continence was modelled: ‘The requirement of total evidence is not a postulate nor a theorem of inductive logic; it is not concerned with the formal validity of inductive arguments. Rather, as Carnap has stressed, it is a maxim for the *application* of inductive logic; we might say it states a necessary condition of rationality of any such application in a given “knowledge situation”.’⁸

Davidson later makes the following claim in his reply to Ariela Lazar:

... as she nicely puts it, ‘An agent who shows no concern for acting in accordance with her better judgment is not an agent.’ To think of it this way is, I agree, to hold the principle to be like a rule of inference, ...⁹

And in his reply to Jon Elster, Davidson says that

⁸ Davidson, ‘Replies to Essays I-IX’, p. 206.

⁹ Davidson, ‘Reply to Ariela Lazar’ (1999), *The Philosophy of Donald Davidson*, p. 404.

When I called the principle ‘second-order’ what I had in mind was its logical, not its psychological function. It serves logically to validate the inference from all-things-considered judgments to the unconditioned ‘This is what it is best to do’ which accompanies action or intention. It is the counterpart of modus ponens in my version of the practical syllogism (which is technically no longer a syllogism), and the counterpart of Hempel and Carnap’s rule of detachment of probabilistic reasoning.¹⁰

These passages are kind of puzzling when they are considered together. The earlier claim seems to suggest that the principle of continence, analogous to the principle of total evidence, is not concerned with the formal validity of practical arguments. As the principle of total evidence is a maxim which counsels an agent to accept the probabilistic belief supported by his total evidence, so the principle of continence is a higher-order principle which exhorts an agent to reason in a way that is in accordance with his best judgment. The latter claim, however, makes it clear that there is an analogy between modus ponens and the principle of continence. Like a rule of inference which validates a deductive argument, the principle of continence validates the inference from an all-things-considered judgment to an unconditional judgment in accordance with it. Validity is used in both claims that seem inconsistent.

Then, how should we understand these passages? What are the relations between the principle of continence, the principle of total evidence, and modus ponens? In the following sections, I suggest a way to read these passages involving the concept of validity which makes them consistent. I attempt to clarify the relations between these principles or rules in terms of understanding the concept of validity in different ways and distinguishing arguments of different kinds to which these principles or rules apply.

¹⁰ Davidson, ‘Reply to Jon Elster’ (1999), *The Philosophy of Donald Davidson*, p. 445.

2.2.1 The disanalogy between modus ponens and the principle of continence

It seems to me that these claims are not inconsistent when the concept of validity in both claims is understood in different ways. The ‘validity’ in the former claim, the reply to Grice and Baker, could refer to deductive validity. Arguments construed according to rules of inference in propositional logic and predicate logic are deductively valid. The validity of these arguments is a function of their logical forms. The structure of a deductive argument guarantees that, given all true premises, it is not possible for the conclusion of the argument to be false. The truth of its premises entails the truth of its conclusion.

When ‘validity’ in the former claim is so understood, the principle of continence that requires an agent to act on his best judgment is irrelevant to it. Modus ponens is the form of a valid argument. An argument with the structure of modus ponens is a deductively valid argument. However, the principle of continence has nothing to do with the form of a practical argument. An agent forms an argument with a correct conclusion when he follows the principle of continence to reason. The argument is right not because it has any specific form but because it properly responds to the agent’s reasons. Furthermore, an argument with the structure of modus ponens is necessarily truth-preserving. But practical arguments are non-deductive arguments because, some believe, the conclusion of a practical argument contains evaluative words such as ‘ought’ or ‘should’, which makes it possible that the conclusion of the argument is false when all of its premises are true. Although a practical argument which is formed in accordance with the principle of continence is an argument with true premises and conclusion,¹¹ as a practical argument, it is a non-deductive argument. The inference does not have the feature of formal

¹¹ This point will be explained later in the section 2.2.2.

validity. The relationship between the true premises and the true conclusion is not an entailment. Hence, an argument with the structure of modus ponens and a practical argument construed in light of the principle of continence are different kinds of reasoning.

Although the truth of the premises does not guarantee the truth of the conclusion in a non-deductive argument, a non-deductive argument can be a strong argument when, given all true premises, its conclusion is likely to be true. Practical arguments are non-deductive arguments, but a practical argument can be a good argument when, say, a practical argument is an argument whose conclusion causes an action that is expected to bring about maximum utility in the agent's calculation, a rational action.¹²

This is not to say that when an agent reasons about what to do, he needs not to follow rules of inference. On the contrary, when an agent reasons, theoretically or practically, he reasons among propositions, and for his inferences to make sense, he must frequently follow rules of inference. An agent who always violates rules of inference to reason seldom forms arguments that are intelligible. Deontic logic is designed for the application of normative concepts and normative reasoning, including concepts such as 'permissible', 'obligatory', 'better than', 'good', and so on. The system of deontic logic is obtained by adding to propositional logic several axioms, theorems, and rules of inference that regulate the logical relations among

¹² The criteria to determine what count as good practical arguments could be various. Decision theory is a theory about how decisions should be made in order to be rational, and it focuses on the evaluation-choice routine (See Sven Ove Hansson's 'Decision Theory: A Brief Introduction', 6, 12.) An agent's desires can be ranked according to how much he values the possible consequence brought about by taking an action of a kind, and his beliefs can be different in how likely he believes that the consequence can be realized through performing a certain action. The agent assigns subjective utilities and subjective probabilities to his desires and beliefs when he measures the relative weights among them. The expected utility is a product of the subjective utility and subjective probability of a pair of desire and belief. A rational agent forms an argument whose consequence causes an action that is expected to bring about the maximum utility. The argument is a good practical argument in the sense of maximizing expected utility.

sentences using normative concepts.¹³ Thus, rules of inference are applicable to practical reasoning as well.

Accordingly, modus ponens is applicable to practical reasoning, but the problem is that a conclusion about what to do cannot be reached by applying modus ponens in an argument. If a prima facie judgment is seen as ‘If r , then a is better than b ’, then when r is also present, it follows that a is better than b . But treating a prima facie judgment in this way is to overlook its ‘prima facie’ characteristic - r as a prima facie reason supports the judgment that an action of a kind is desirable. Thus, if we want to stress that prima facie judgments are conditionals, it would be better to express a prima facie judgment as ‘If r is the case, then r prima facie makes it true that a is better than b ’. Given the reason r , it follows that r prima facie makes it true that a is better than b . In other words, by applying modus ponens, what is deduced from the premises is also a prima facie judgment, not an intention to do something.

An inference from premises $pf(a$ is better than $b, r)$ and r to the conclusion a is better than b cannot be characterized as a deduction. Justin Gosling says that ‘“Prima facie” puts up the warning signs that there is more work to be done before the conclusion can be drawn.’¹⁴ When an agent holds a prima facie judgment, instead of directly inferring an unconditional judgment from it, he needs to consider whether or not it is the only relevant reason or is sufficient to detach an unconditional judgment. As Davidson clarifies, ‘a ‘ pf ’ judgment expresses a view of the relation between two propositions, and makes no commitment to the truth of either.’¹⁵ Although r , if true, would be a reason that supports the agent to judge that a is better than b , the agent may

¹³ For the axioms, theorems, and rules of inference, see Risto Hilpinen’s ‘Deontic Logic’ (2001), pp. 160-162.

¹⁴ Justin Gosling, ‘What is the Problem?’ (1990), p. 105.

¹⁵ Davidson, ‘Replies to Essays I-IX’, pp. 202-3.

not think either r or the judgment that a is better than b to be true when they are not related by the prima facie operator. Thus, directly drawing a conclusion about what to do from a prima facie judgment without deliberation is illegitimate.

Because practical arguments are not deductive arguments, the premises of a practical argument do not entail its conclusion. An agent who holds the principle of continence knows that he should reason and act in accordance with his best judgment, but since his practical reasoning is not a deductive reasoning, the agent is not forced to infer an unconditional judgment from his all-things-considered judgment. In other words, an agent could hold an all-things-considered judgment without holding a consistent unconditional judgment. The existence of an all-things-considered judgment does not imply the existence of a consistent all-out, unconditional judgment. In this sense, the principle of continence is not analogous to modus ponens in that the latter applies to arguments whose premises and conclusions have an entailment relation, but the former does not.

Applying the principle of continence in a practical reasoning is similar to applying the principle of total evidence in an inductive reasoning. In a probabilistic reasoning, an agent considers competing probabilistic judgments before forming a conclusion about what is almost certainly to occur. When the principle of total evidence is effective, the agent detaches a conclusion from the probabilistic judgment supported by the strongest evidence. The argument is a strong argument, and the agent who reasons consistently is a rational agent. Yet, whether or not the principle of total evidence is effective in a probabilistic reasoning, the argument is not deductive in form. As a non-deductive argument, the premises and the conclusion of the argument are not in a relation of entailment. An agent could hold a probabilistic judgment

supported by the strongest evidence without detaching a conclusion about what is almost certainly to occur from it. Similarly, an agent who holds an all-things-considered judgment does not necessarily detach a conclusion about what to do from it.

2.2.2 The analogy between modus ponens and the principle of continence

The ‘validity’ in the later claims, the replies to Lazar and Elster, on the other hand, seems to be a property of practical arguments where the normative force of the principle of continence is exhibited, i.e., a property of arguments which are inferences from all-things-considered judgments to unconditional judgments. The ‘validity’ here is not a matter of argument form but is understood in terms of truth preservation.

In the essay ‘The Objectivity of Values’, Davidson talks about the truth-evaluability of evaluative judgments in practical reasoning. He says,

When we reason about what to do or try to do, or about the value or morality of the actions of others, we must combine factual judgments with our values. ... Such reasoning, if properly laid out, is surely valid. But validity is defined as a truth preserving mode of reasoning. If practical reasoning can be shown to be (in some cases) valid, the premises and conclusion must have truth values. Our understanding of the logical constants brings out the same point. We say, *if* it is desirable that my jetlag be cured *and* taking Melatonin will cure it, *then* it is desirable that I take Melatonin. But what do ‘and’ and ‘if ... then’ mean here? ‘And’ is usually defined as yielding a true sentence if and only if each conjunct is true, and ‘if ... then’ is similarly defined as a truth functional connective. If we abandon these definitions, what should we put in their place?¹⁶

¹⁶ Donald Davidson, ‘The Objectivity of Values’ (1995), reprinted in *Problems of Rationality* (2004), p. 55.

This passage is to support the view that evaluative judgments are truth-evaluative, but it also reveals the conception that some practical arguments can be construed as valid arguments, and they are valid in the sense that the conclusion is true whenever the premises are true.

We know that practical arguments are not deductive arguments. Practical arguments are not deductive in form, and so they are not necessarily truth-preserving. On the contrary, the conclusion of a deductively valid argument is a logical consequence of its premise. Its logical form makes the conclusion necessarily follow from the premises. The truth of the premises guarantees the truth of the conclusion.

However, deductively valid arguments are not the only kind of argument that is truth preserving and valid. A non-deductive argument can be valid if it is metaphysically, conceptually, or analytically impossible for the premises to be true and the conclusion to be untrue. For example, arguments like ‘ x is water, therefore x is H_2O ’ and ‘Peter is Greg’s mother’s brother’s son, therefore Peter is Greg’s cousin’ are necessarily truth preserving, although they are not deductively valid.¹⁷

Similarly, a practical argument could be valid. The validity of a practical argument does not come from metaphysical, conceptual, or analytical necessity but from the rational force of the principle of continence. When the principle of continence is effective, an agent forms an argument in which he detaches an unconditional judgment from his all-things-considered judgment. For an argument of this kind, it is impossible that its true premises yield a false conclusion. And so it is a valid argument.

¹⁷ See Jc Beall and Greg Restall’s ‘Logical Consequence’, The Stanford Encyclopedia of Philosophy, Winter 2013 Edition.

The ideas of subjective utility and subjective probability in expected utility theory help to explain the truth of an all-things-considered judgment. In a decision making procedure, an agent measures the relative force of his desires and beliefs, and assigns subjective utilities and subjective probabilities to his desires and beliefs according to how much he values a consequence that would be brought about by taking an action of a kind and how much he believes that the consequence can be obtained by taking a certain action. Apparently, a desire that is expressed by an all-things-considered judgment is assigned the highest value by an agent. The all-things-considered judgment is a judgment true to the agent.¹⁸

The truth of an all-things-considered judgment guarantees the truth of the unconditional judgment detached from it in an argument. When an agent deliberates about what to do, he considers various competing prima facie judgments before making a decision. Only when he makes up his mind, can he take one of the competing prima facie judgments to form an argument whose conclusion is an unconditional judgment that causes him to act. How strongly the premises of this practical argument support its conclusion depends on how good the prima facie judgment that the agent takes to be the premise of his argument is. An all-things-considered judgment is supposed to provide an agent with the strongest reason to act. When it becomes a premise of a practical argument, the premises of the argument provide its conclusion with the

¹⁸ An agent assigns the highest value to his all-things-considered judgment, his best judgment, and the judgment is true to the agent. But the truth value of an evaluative judgment is not subjectively determined. Evaluative judgments can be objectively true or false. Davidson claims that objectivity 'depends on there being a systematic relationship between the attitude-causing properties of things and events, and the attitudes they cause'. (See Davidson's 'The Objectivity of Values', p. 47.) Accordingly, evaluative judgments are objectively true if the desire attitudes they represent are caused by objects that are attractive to thinking creatures. Since we share with other agents the desires to seek warmth, love, security, success and to avoid pain and distress, evaluative judgments caused by these objects are correct by interpersonal standards - we commonly regard these objectives as worth pursuing.

strongest justification or reason support, and the argument is a good argument. The agent takes the premises to be true, and the conclusion that is logically related to the premises is also true.

For arguments which are inferences from all-things-considered judgments to consistent unconditional judgments, they are validated by the principle of continence. Although practical arguments as non-deductive arguments are not necessarily truth-preserving, practical arguments could be valid in the sense that for an argument which has an all-things-considered judgment as one of its premises, it is not possible that its conclusion is false while its premises are true. The formation of an argument of this kind is a result of following the principle of continence. In this sense, the principle of continence is a rule of inference that is applicable to practical reasoning and validates a practical argument that is formed according to it.¹⁹ The principle of continence justifies an agent's inference from his all-things-considered judgment to a consistent unconditional judgment. It demonstrates that an agent who obeys it has sufficient reason to hold the unconditional judgment and to act on it. If the principle of continence is understood as a rule of inference that validates inferences from all-things-considered judgments to unconditional judgments, then there is an analogy between modus ponens and the principle of continence in that they are second-order principles that guide rational agents to reason and act. The principle of continence is analogous to modus ponens with respect to the fact that an agent who follows the principle of continence or modus ponens to reason does not get a false conclusion when he holds all true premises, i.e., an unconditional judgment is true whenever it is detached from his all-things-considered judgment.

¹⁹ Ian Rumfitt believes that "There is no suggestion that the only rules that are logically necessarily truth-preserving are the formal rules presented in logic book. For example, we are surely prepared to apply the rule of inference "From 'x is red,' infer 'x is colored'" in reasoning from any set of suppositions whatsoever." (See Ian Rumfitt, 'What is Logic?', pp. 155-156.) Analogously, the principle of continence is a rule of inference the application of which forces an agent to infer an unconditional judgment from his all-things-considered judgment. We frequently use such a principle in practical reasoning and make most of our actions intelligible to others.

By the same token, the principle of total evidence, when applied, can be seen as a rule of inference that validates inferences from inductive probabilistic judgments supported by the strongest evidence to conclusions about what is almost certainly to occur. The validity here is also defined as truth preservation rather than as deductive validity. As evaluative judgments which are truth-evaluable, inductive probabilistic judgments can also be true or false. Unlike a statistical probability which is understood in terms of the concept of frequency and is specified by a numerical value, an inductive probability judgment is a hypothesis established in relation to a body of evidence and is explicated as being higher or lower than another inductive probability judgment. The hypothesis is a prediction of a future event. The evidence is a set of known or assumed facts observed by an agent, the totality of relevant information available to an agent.²⁰ The stronger the evidence is, the more likely that the hypothesis is true. Hence, a future event indicated by a hypothesis supported by the strongest evidence is regarded by an agent as more probable than other future events to occur. When an agent infers from an inductive probabilistic judgment supported by the strongest evidence to a conclusion about what is almost certainly to occur, the principle of total evidence is effective, and the inference is valid in the sense that it is not possible for the conclusion to be false when the premises are true. The principle of total evidence functions as a rule of inference which validates arguments construed according to it.

Although the principle of continence, the principle of total evidence, and modus ponens are respectively applicable to non-deductive and deductive arguments, it is only when the force of rationality exhibits that these principles or rule will work. In practical reasoning or inductive reasoning, an agent does not necessarily detach a conclusion from his best or strongest judgment for there is no entailment between the two. An agent reasons consistently when he is rational. In

²⁰ Rudolf Carnap, 'Statistical and Inductive Probability' (1955), pp. 2-3.

theoretical reasoning, although there is an entailment between premises and the conclusion of a deductive argument, an agent's forming an argument with the structure of modus ponens, whether or not he knows such a rule of inference, relies on the performance of rationality. The agent must exercise his rational ability to discern what follows from the premises of the argument. When an agent's rationality shuts down, he can draw no inferences with correct conclusions. Consequently, the principle of continence, the principle of total evidence, and modus ponens are similar in the sense that when any of them is applied in reasoning, an agent is forced to reason logically and consistently. The operation of rationality makes it the case that the agent cannot avoid detaching a conclusion in accordance with premises endorsed by his rationality.

2.3 The distinction between two kinds of evaluative judgment

Grice and Baker, Peacocke, and Charlton take Davidson's regarding incontinence as a failure in reasoning to be problematic. They believe that when an agent holds an all-things-considered judgment that $pf(a \text{ is better than } b, r)$ and the reason r , he must detach an unconditional judgment that a is better than b . An all-things-considered judgment $pf(a \text{ is better than } b, r)$ implies an unconditional judgment that a is better than b . Grice and Baker explain this as if there is an analogy between modus ponens in theoretical reasoning and the principle of continence in practical reasoning. Through the analogy, they show that incontinence as a failure in reasoning is impossible. Just as an agent who holds that if p , then q and that p is forced to conclude that q , an agent who holds that considering all relevant reasons, a is better than b , cannot avoid judging unconditionally that a is better than b . It is not possible for an agent to hold an all-things-considered judgment that $pf(a \text{ is better than } b, r)$ and the reason r , but detaches an

unconditional judgment *b* is better than *a* from the other competing prima facie judgment that *b* is better than *a*.

In the above interpretation of the passages involving the concept of validity that I suggested, there seem to be both an analogy and a disanalogy between modus ponens in theoretical reasoning and the principle of continence in practical reasoning. As an argument construed in light of modus ponens cannot have a false conclusion whenever its premises are true, a practical argument which is a consequence of applying the principle of continence always has a true conclusion detached from an all-things-considered judgment that is true. Such a practical argument is valid in the sense that it is necessarily truth-preserving. A practical argument, however, is a non-deductive argument. Even if it is construed according to the principle of continence and is valid, the premises and the conclusion of the argument are not in a relation of entailment. On the contrary, an argument construed according to modus ponens is deductively valid. The premises of the argument entail the conclusion of the argument.

Suppose the analogy between modus ponens in theoretical reasoning and the principle of continence in practical reasoning is acceptable, the possibility of the failure to reason in accordance with an all-things-considered judgment is not then ruled out. Davidson claims that ‘Simple failures in deductive reasoning are fairly common, even among those who know some logic.’²¹ People who apply logical rules to reason could make mistakes, especially when reasoning processes are lengthy and complex. Sometimes people make mistakes in reasoning because they fail to grasp the correct logical forms of propositions which they entertain to form

²¹ Davidson, ‘Replies to Essays I-IX’, *Essays on Davidson: Actions and Events* (1985), p. 206.

arguments. Similarly, we should allow that an agent who holds the principle of continence sometimes makes a mistake, violates rules, and draws false conclusions in practical reasoning.

What is more important is the disanalogy between modus ponens in theoretical reasoning and the principle of continence in practical reasoning, which explains why an agent can hold an all-things-considered judgment without holding an unconditional judgment consistent with it.

Arguments formed in the light of modus ponens are deductively valid arguments. For arguments of this kind, the conclusion necessarily follows from the premises. Arguments with lengthy and complex steps apart, arguments with the structure of modus ponens can easily be formed, and so an agent who holds ‘if p , then q ’ and ‘ p ’ can hardly avoid forming a conclusion ‘ q ’.

If practical arguments are deductive arguments, an agent who holds premises that are essential to constitute a practical argument will usually be forced to form a corresponding conclusion since practical arguments of this kind are with brief steps and simple structure. Aristotle, for example, believes that the form of a practical argument is a syllogism, and practical syllogisms are similar to theoretical syllogisms in forms. Both kinds of syllogism can be constructed as formally valid deductive arguments. As Anscombe points out, Aristotle ‘is marked by an anxiety to make practical reasoning out to be as like as possible to speculative reasoning. “They work just the same’, ... seems to be referring to a necessitation of the conclusion.’²² This is because Aristotle hopes that the automatic-machine aspect of practical reasoning helps to explain how the syllogism sets the human animal in motion. The form of a practical argument promises a necessity in the connections among its premises and conclusion,

²² G. E. M. Anscombe, ‘Thought and Action in Aristotle’ (1965), p. 153.

and the major premise of the argument as a universally quantified conditional has a compulsiveness that an agent who accepts it takes it as a guide to act.²³ An agent must deduce a conclusion and take a corresponding action when he holds a universally quantified conditional, a general statement, and a relevant belief, a specific statement. The conclusion is an entailment of the premises.

If practical arguments are understood as deductions in which an all-out judgment necessarily follows from an all-things-considered judgment and the content of a relevant belief, then Grice and Baker are right to argue that an all-things-considered judgment is not separable from an all-out, unconditional judgment consistent with it because the former implies the latter.

However, practical arguments are not deductions. The lack of the entailment or automatic transition between prima facie, conditional judgments and the detached all-out, unconditional judgments explains why an agent can hold an all-things-considered judgment without holding an all-out, unconditional judgment consistent with it. An all-things-considered judgment is distinct from an all-out, unconditional judgment consistent with it in the sense that the existence of an all-things-considered judgment does not necessarily involve the existence of a consistent unconditional judgment. Since there is no entailment between the two kinds of evaluative judgment, an agent could fail to infer an all-out, unconditional judgment from his all-things-considered judgment. Incontinence as a failure in reasoning is possible.

3. THE MOTIVATIONAL FORCE OF A PRIMA FACIE JUDGMENT

²³ Ibid., pp. 152-154.

In a case of incontinent action, an agent fails to reason in accord with his all-things-considered judgment that $pf(a$ is better than $b, e)$ and the reason e , but detaches an unconditional judgment that b is better than a from another competing prima facie judgment supported by a relatively weaker reason, $pf(b$ is better than $a, r)$, and the reason r . In section 2, I have defended the possibility of incontinence as a failure in reasoning, and in this section I investigate the cause of incontinence: what causes an agent to fail to reason in accord with his all-things-considered judgment? The desire causing an incontinent agent to act is supported by a relatively weaker reason, but its motivational strength is stronger than that of the all-things-considered judgment. Then, this question can be put in another way: why does a prima facie judgment supported by a relatively weaker reason get a superior motivational force in an incontinent agent's reasoning?

Davidson identifies the desire causing an incontinent agent to act as the cause of his incontinence, and proposes the theory of partitioned mind to explain how such a mental cause exists in the mind. But many think that Davidson's identifying the effective desire as the cause of incontinence and proposing the theory of partitioned mind does not explain why the effective desire is motivational stronger, why the all-things-considered judgment is ineffective. The lack of self-control and emotions are suggested as the cause of an incontinent agent's failure in reasoning by some philosophers. However, I argue that no specific kind of mental states would be eligible to be the cause of incontinence which explains the stronger motivational force that the effective desire has in an incontinent action. Instead, any mental event or state that fits an incontinent agent's personal character can intervene in his reasoning process, cause him to be interested in the desire supported by a weaker reason, and thereby increase the motivational strength of the desire supported by a relatively weaker reason in a way that does not respond to all the reasons the agent has.

3.1 A challenge to the effective desire as the cause of incontinence

Under the assumption that propositional attitudes are causally and logically connected with each other, Davidson proposes the theory of partitioned mind to explain how an agent can simultaneously harbor inconsistent propositional attitudes in his mind. According to the theory of partitioned mind, an agent's mind consists of various parts. Each part is constituted by propositional attitudes which are causally and logically connected with one another. To the extent that the principle of logic fails to apply, there are only causal relations between propositional attitudes. Whenever a breakdown of the reason relation between two propositional attitudes occurs, they are divided into different parts of the mind. The two propositional attitudes are respectively supported by a set of interlocking propositional attitudes, but between them there is a non-logical causal relation. Because of the separation, an agent can hold inconsistent propositional attitudes in his mind.

Incontinent action is a case in which an agent holds inconsistent judgments. In an incontinent action, the incontinent agent's desire to perform the present action causes him to ignore the principle of continence and fail to reason in accordance with his all-things-considered judgment. Although the desire is with a relatively weaker reason, it causes an intention to perform the present action, which is inconsistent with the agent's all-things-considered judgment and the principle of continence. Hence, the incontinent agent is in a state of inner inconsistency. According to the theory of partitioned mind, the inconsistent judgments that coexist are separated into different parts of the mind. The principle of continence, the all-things-considered judgment, and other interlocking propositional attitudes constitute a part, while the effective desire, the intention to perform the present action, and other supporting propositional attitudes form another

part. Because of the separation, the principle of continence is temporarily ineffective to the part which gives a reason and causes the incontinent agent to act, and so is the agent's all-things-considered judgment.

However, many believe that in this account Davidson does not really explain why the desire to take the action that the incontinent agent performs is motivationally stronger than the desire with the best reason. David Pears claims that even though the theory of partitioned mind is able to cover all cases of akrasia, such as incontinent action or self-deception, it has no explanatory force because its success is achieved by definition. The theory is successful because a subdivision's boundary is defined as a line across which some element in a person's psyche fails to exercise its normal rational force to the elements belonging to different subdivisions. 'That definition guarantees a perfect fit between the functional theory of systems and the phenomenon of irrationality that the subject is competent to avoid, ...'²⁴ Irving Thalberg claims that the theory of partitioned mind cannot make sense of the idea that considerations which favor my doing *x* may seem more cogent to me but have less influence upon me than those favoring *y* because the theory does not improve our understanding of the motivational force of the prima facie judgment causing the unconditional judgment to do *y*. The lack of a proper interpretation of the motivated strength of the prima facie judgment causing an unconditional judgment makes the theory lose its explanatory force.²⁵ Ariela Lazar and Annette Baier also think that the theory fails to explain why an all-things-considered judgment fails to cause a corresponding unconditional

²⁴ David Pears, 'Paradoxes and Systems' (1984), p. 84.

²⁵ Irving Thalberg, 'Questions about Motivated Strength' (1985), pp. 101-103.

judgment and action, and they ask what decides when one subdivision rather than the other controls a behavior.²⁶

3.2 Is the lack of the virtue of self-control the cause of incontinence?

Apparently many philosophers are not satisfied with Davidson's explanation of the cause of incontinence. They feel that proposing the theory of partitioned mind and giving the agent's stronger desire to do what he has done do not really explain why the desire with a relatively weaker reason is motivationally stronger, why the desire supported by his best reason fails to cause a corresponding intention and action. In order to explain why an agent fails to reason and act in accord with his all-things-considered judgment, some philosophers try to identify a mental state that is responsible for the failure. Both Gary Watson and Alfred Mele suggest that an incontinent agent fails to reason in accord with his all-things-considered judgment because he lacks self-control. Then, is the incontinent agent's lacking self-control the cause of his incontinence?

Unlike Davidson who claims that there is an internal and necessary relation between unconditional judgments and actions,²⁷ Gary Watson denies such a connection. He even refuses to take desires and intentions as evaluative judgments. Watson adopts the view that evaluative judgments are rankings or preferences of the expected values of taking actions. The rational part

²⁶ Annette Baier, 'Rhyme and Reason: Reflections on Davidson's Version of Having Reasons' (1985), p. 121. Ariela Lazar (1999), 'Akrasia and the Principle of Continence or What the Tortoise Would Say to Achilles', 393-394.

²⁷ Davidson holds P1 'If an agent wants to do *x* more than he wants to do *y* and he believes himself free to do either *x* or *y*, then he will intentionally do *x* if he does either *x* or *y* intentionally' and P2 'If an agent judges that it would be better to do *x* than to do *y*, then he wants to do *x* more than he wants to do *y*' to be true, and in his analysis P1 and P2 are applicable to all intentional actions which are caused by conditional judgments, the outcomes of deliberation. This claim is a result of P1, P2, and such an analysis.

of the soul is the source of these evaluations. An agent could desire to act in accord with his evaluation, the good. Desires arising from appetites and emotions, however, come from the irrational part of the soul. They are blind in the sense that they are fundamentally independent of the agent's view of the good. Because evaluations are distinguished from desires, it is possible that 'one's evaluations and desires diverge in certain cases in such a way that one is led to do things which one does not think worth doing, or as much worth doing as some available option.'²⁸ When a desire caused by an appetite or emotion is strong enough, it moves an agent to act contrary to his evaluation.

Davidson believes that in a piece of practical reasoning, an agent deliberates about what to do, and the result of his reasoning is an unconditional, evaluative judgment that motivates him to act. However, Alfred Mele follows Watson to claim that when an agent has conflicting wants in his practical reasoning, he is attempting to rank his wants on some scale of values, and this is not to determine which of his competing wants has the strongest motivational force. On the contrary, evaluation and motivation can sometimes be out of line with one another.²⁹ What an agent wants more in the evaluational sense may be incompatible with what he wants more in the motivational sense. In this sense, it is possible that the motivational strength of the desire caused by evaluation is inferior to that of the desire caused by appetite or emotion. Davidson is wrong in that he fails to distinguish the two senses of 'wants more', evaluational and motivational, and so it is difficult for him to explain why an incontinent agent's desire assigned with the highest value is not motivationally strongest.

²⁸ Gary Watson, 'Skepticism about Weakness of Will' (1977), p. 320.

²⁹ Alfred Mele, 'Akratic Action: Causes, Reasons, Intentions' (1987), p. 37.

Watson identifies the lack of the virtue of self-control as the cause of incontinence.

Watson claims that an incontinent agent fails to act in accord with his best judgment because he lacks the virtue of self-control which is a capacity to counteract recalcitrant motivation contrary to one's better judgment. The agent fails to resist his desire contrary to his evaluation, and we blame him because the desire is resistible in the sense that if the agent had developed the virtue of self-control, he could have resisted the recalcitrant desire.³⁰

Mele basically accepts Davidson's causal theory of action, and believes that an incontinent action is caused by a pair of belief and desire. However, although the belief and desire explain why the incontinent agent takes the action that he performs, they do not explain why he does not take the alternative action judged to be better. An adequate explanation of an agent's action, Mele suggests, must go beyond his reason for acting. Prima facie, evaluative judgments have a motivational element, but whether an all-things-considered judgment results in action 'depends not only on the agent's being able to act accordingly and the absence of external prevention, but also on the extent to which he is self-controlled'.³¹ When a worse judgment gets stronger motivational force than a better judgment, an agent who fails to be perfectly self-controlled will not be able to elicit a reversal of the balance of his motivations.

In my view, Watson's and Mele's taking the lack of the virtue of self-control as the cause of incontinence encounters several problems. Watson thinks that a desire caused by an appetite or emotion is blind and fundamentally independent of an agent's evaluation. This seems to mean that a desire of this kind could be a force alien to an agent when it occurs in an agent's mind. Then, when the desire is motivationally strongest, the agent who acts on it seems to act

³⁰ Gary Watson, 'Skepticism about Weakness of Will', p. 323, pp. 333-4.

³¹ Alfred Mele, 'Akratic Action: Causes, Reasons, Intentions', p. 45.

involuntarily. He is moved by a desire he does not endorse. Incontinent actions as involuntary actions are not how we understand incontinent actions.

Furthermore, the explanatory force of Watson's explanation seems limited. Given that the desire causing an incontinent agent to act is understood to be blind and irresistible, the lack of the virtue of self-control may be able to explain cases in which an agent's rationality is dominated by his overwhelming desire or passion.³² But it cannot explain cases in which an agent calmly accepts his worse value judgment and acts on it. In cold cases of *akrasia*, an agent does not lose control of himself, but knowingly and intentionally acts against his best judgment. The desire causing him to act is not irresistible. If he wanted, he could have done otherwise. It is the cold cases of *akrasia* that puzzle us here. Lacking self-control seems not helpful to our understanding cold cases of *akrasia*.

However, even the suggestion that the lack of self-control explanation improves our understanding of hot cases of *akrasia* is doubtful. As Gosling clarifies, the Greek word "akrasia" means 'lack of control' and suggests the Socratic sort of examples in which an agent is overcome by passion to act on his false judgment.³³ Indicating the lack of self-control as the cause of incontinence seems to say no more than how the concept of incontinence is defined.

³² Watson distinguishes weakness of will from compulsion. In each of the two cases an agent is caused to act by an irresistible desire, but Watson indicates that they are distinguishable in terms of the virtue of self-control and the capacity of resistance acquired in the normal course of socialization and practice. Watson says, 'Weak agents fall short of standards of "reasonable and normal" self-control (for which we hold them responsible), whereas compulsive agents are motivated by desires which they could not resist even if they met those standards' (See Watson, 'Skepticism about Weakness of Will', 332). Although the desire that caused an incontinent agent to act could have been resisted, it is in fact irresistible to the agent when he acts. Thus, given that the desire causing an incontinent agent to act is blind and irresistible, it should be reasonable to treat the desire as an overwhelming desire and the case in which an agent is caused to act by the desire a hot case of *akrasia*.

³³ Justin Gosling, 'What is the Problem?', 97.

Although Mele takes the lack of self-control as one of the causes of incontinence, this less strong claim cannot avoid the challenge to the explanatory force of his view. Mele hopes to solve the problem why the desire causing an incontinent agent to act is motivationally stronger than his desire to act on his better judgment, but the answer that an agent is incontinent because he loses self-control does not tell us why but simply assumes the desire causing the agent to act to be motivationally strongest among all desires in the beginning. The incontinent agent is wrong because he fails to defeat the strongest desire contrary to his evaluation. The problem of weakness of will that Davidson proposes is puzzling because the desire expressed by an all-things-considered judgment is supposed to be causally and motivationally strongest, but in a case of incontinent action it turns out that another desire based on a relatively weaker reason causes the agent to act. The incontinent agent thinks that he should act on his best judgment but fails to do so. It seems that the problem of weakness of will is not genuine for Watson, Mele, and those who think that evaluation is distinguished from motivation because an incontinent agent who is caused to act by a strong desire that he does not endorse simply does not think that he wants to take the action that he performs.

3.3 Is emotion the cause of incontinence?

Although Watson and Mele are dissatisfied with Davidson's analysis, their answer to the problem about the cause of incontinence seems not better. Let us consider another suggestion: emotion. Christine Tappolet claims that emotions, at least in cases in which the desire causing an incontinent agent to act is caused by an emotion, have a causal role and make incontinent actions of this kind intelligible. Is emotion the cause of incontinence? Can emotion explain why the motivational force of a desire based on a relatively weaker reason is superior to that of the desire

supported by all relevant reasons, and so explain why an incontinent agent fails to detach an unconditional judgment from his all-things-considered judgment?

Emotions, according to Tappolet, are perceptions of values. An object that attracts someone's attention 'falls under some evaluative concepts, such as the concepts of the dangerous, the disgusting, the shameful, the irritating, the attractive, the admirable, or the love-worthy.'³⁴ An emotion arises when an agent perceives an object with any of these values. Tappolet takes emotions to be analogous to perceptual experiences in that both kinds of perception can be assessed in light of whether the perceived object or value is correctly represented and that the content of both kinds of perception is non-conceptual. She remarks that 'emotions are to a large extent isolated from our higher-order cognitive processes and thus from our deliberative faculty.'³⁵

An emotion can direct an agent's attention to an aspect of his circumstances and cause the agent to take an action. Tappolet shares with Ronald de Sousa the view that emotions direct our attention in the sense that 'emotions limit the range of information that the organism will take into account, the inferences actually drawn from a potential infinity, and the set of live options among which it will choose'.³⁶ Accordingly, emotions have a practical function. The emotion involved in a practical reasoning drives the agent to notice a certain aspect of his situation and ignore others. Corresponding to his attention, the agent forms an inference about what to do.

³⁴ Ibid., p. 110.

³⁵ Christine Tappolet, 'Emotions and the Intelligibility of Akkratic Action' (2003), pp. 109-110.

³⁶ Ibid., p. 104.

An emotion can be assessed in terms of how it fits evaluative facts, which determines whether an action caused by it is rational or not. If an agent's emotion is compatible with his rational judgment about what the case is, then the emotion is appropriate, and the action is rational. In cases in which incontinent actions are caused by emotions, there is a conflict between a value perception and an evaluative judgment. The value perception, the emotion, appears to be inappropriate because it does not fit the evaluative fact that the agent recognizes, and so the action caused by the emotion is an irrational action.

Even if the emotion involved in an incontinent action is inappropriate, it explains why the incontinent agent fails to reason in accord with his best judgment and makes the incontinent action intelligible. Tappolet gives an example to illustrate this view. 'Suppose I am about to cross a narrow rope bridge hanging high up on a deep shaft. Though I feel fear, I judge that all things considered I ought to cross the bridge; I judge it to be sufficiently safe and going back would make for a much longer hike. If I end up not crossing the bridge, it will not be difficult to make sense of my action: the perceived danger, be it real or not, readily explains why I didn't cross the bridge.'³⁷ In other words, given the agent's judgment that all things considered crossing the bridge is better than not crossing, the fear to cross the bridge is inappropriate. However, the fear causes the agent to focus on another judgment that it would be better not to cross the bridge and detach an unconditional judgment from it. The fear explains the incontinent agent's failure in reasoning, and makes the incontinent action understandable.

Taking emotion as the cause of incontinence is indeed a way to explain why an incontinent agent fails to reason in accord with his best judgment, but it seems to me that how

³⁷ Ibid., 111.

Tappolet interprets the concept of emotion is problematic. Tappolet's taking the content of emotions to be non-conceptual makes the concept of emotion incomprehensible. To say that the content of emotions is non-conceptual seems to mean that emotions are just feelings. If emotions are just feelings, what could their content be? Is the content a specific feeling that constitutes an emotion? But we find that there is no necessary connection between an emotion and a specific feeling. My face turns red and my heart rate increases when I am embarrassed. But I may be embarrassed without feeling any of these changes in my body. Sometimes I experience these changes but do not feel embarrassed- my face turns red and my heart rate increases when I am running. Thus, feelings seem not proper to be the content of emotions.

Moreover, Tappolet believes that emotions are perceptions of values. Applying evaluative concepts to objects that are perceived, however, must involve cognition. If to have an emotion is to perceive an object that is thought to be, say, dangerous, dreadful, annoying, lovely, exciting and so forth, then concept and judgment are essential to having an emotion. Tappolet's holding emotions to be perceptions of values and the content of emotions to be non-conceptual would make her position unintelligible.

There is something wrong with Tappolet's explanation, but given Davidson's analysis of the concept of emotion, Davidson would agree that emotions have a role in the causal history of some actions. To Davidson, emotions are not simply feelings to which no concepts are applicable but are propositional attitudes. In the essay 'Hume's Cognitive Theory of Pride', Davidson takes pride as an example to analyze the nature of emotions. Pride can be considered in two ways. As an affective state, the impression of pride is an agreeable passion. According to Hume, the beliefs that cause the impression of pride always cause the impression of pleasure at the same

time, and so the two impressions are somehow inseparable. Pride, as a cognitive state, is expressible in sentences such as 'I am proud that I own a beautiful house'. Davidson regards the impression of pleasure as the attitude of approval and identifies the state in which someone is proud that p with his having the attitude of approving of himself because of p .³⁸ This is because someone's being proud of p is caused by a judgment that a person's exemplifying the property p prima facie makes him praiseworthy and a belief that he himself exemplifies that property p . In other words, self-approval which consists of a judgment of approval and a belief whose content involves the reference to self constitutes pride. Pride and its cause, when represented by propositional expressions, exhibit a logical relation.³⁹ In a similar way, other emotions can also be characterized with respect to the affective aspect and the cognitive aspect. When an emotion is understood in terms of the cognitive aspect, the emotion makes sense to us because a reason explanation of it is offered to rationalize it. Thus, given Davidson's view on emotion, if the desire causing an action is caused by an emotion, the emotion is a part of the rational pattern in which the action occurs.

Some may disagree with Davidson on how emotions are caused, but I think that, whether or not Davidson's analysis is acceptable, it is reasonable to take emotions as cognitive states. This is because when an agent has an emotion, his perceiving or recognizing the emotion is a process of thinking and involves the use of concepts and propositions. When the emotion causes

³⁸ Donald Davidson, 'Hume's Cognitive Theory of Pride' (1976), reprinted in *Essays on Actions and Events* (2001), p. 281, p. 287.

³⁹ The logical relation between pride and its cause can be recognized when the expressions of pride and its cause are construed as an argument. Davidson clarifies that the argument cannot be a syllogism, such as 'All who own a beautiful house is praiseworthy, I own a beautiful house, therefore I am praiseworthy'. On the one hand, we do not esteem someone just because he owns a beautiful house. On the other hand, the middle term 'owns a beautiful house' drops out in the conclusion, but it is needed as the content of propositional pride. To avoid these troubles, the major premise should not be a universal statement. Davidson suggests, 'My approval or judgment is prima facie in character and is given only relative to its ground' (Davidson, 'Hume's Cognitive Theory of Pride', p. 285). When the major premise is expressed as a prima facie judgment, the conclusion detached from it is also a prima facie judgment: I am praiseworthy in so far as I own a beautiful house.

any thoughts, the agent is in a process of reasoning. Since emotions involve thinking and reasoning, they are cognitive states. In the story of Romeo and Juliet, Romeo's recognizing his love to Juliet entails that he holds a belief about his love to Juliet. The love causes him to drink poison when he believes Juliet to be dead. Love, an emotion that interacts with other propositional attitudes in his thinking and reasoning, is a cognitive state. Only when emotions are cognitive states, can emotions be reasons that cause intentional actions.

To say that emotions have propositional content and are effects of beliefs is not to equate emotions with cognitions but to emphasize the built-in rationality of emotions. Because an emotion has propositional content and is a thought of a certain kind, when it is involved in the causal history of an incontinent action, it can be a reason favoring the desire that causes the incontinent agent's intention to perform the present action. In this sense, the emotion has a causal and logical relation with the incontinent action. On the contrary, Tappolet's taking emotions as non-cognitive states has the upshot that when an emotion is involved in the causal history of an incontinent action, it has no role to play in the rational pattern in which the incontinent action, an intentional action, is placed to be understood. It breaks the logical connection between the action and its cause and makes the action unintentional.

Besides explaining the rational aspect of an incontinent action, emotion can also explain the irrational aspect of an incontinent action. An emotion is eligible to be the cause of incontinence if and only if it intensifies the incontinent agent's interest in satisfying the desire causing him to act and makes the motivational force of the desire increase in a way that is not in proportion to the reason supporting the desire. The emotion is a cause of, but not a reason for, the agent's being incontinent because its occurrence prompts the agent to see the insufficient reason

of a desire as sufficient but fails to rationalize such a change. Hence, it seems reasonable to take the emotion as a cause which gives a reason explanation for why the incontinent agent takes the action that he performs and also gives a non-logical causal explanation to his being incontinent.

But the problem is that the attempt to take emotions to be the cause that accounts for all cases of incontinent action in which emotions are involved cannot be successful. We find that sometimes a person is intensely angry and is then caused to have certain thoughts, but the emotion neither rules his mind nor causes an incontinent action. Sometimes an agent acts incontinently and with an emotion, but the incontinent action is not caused by the emotion. Emotion could be the cause of incontinence to a case of incontinent action in which an emotion is involved, but it seems not able to account for all incontinent actions in which emotions are involved. And, of course, it cannot account for cases of incontinent action in which emotions are not involved either.

3.4 The cause of incontinence

Perhaps we should abandon the idea that there is a unique kind of mental state that is the cause of incontinence covering all cases of incontinent action. In other words, there is no mental state of a specific kind that whenever it presents, it makes an agent fail to reason in accordance with his best judgment and act incontinently. This view on the cause of incontinence echoes Davidson's anomalism that there are no psychological or psychophysical laws which regulate the connection between thoughts, and the connection between thoughts and actions. Since thoughts are not necessarily connected in certain ways, no mental state of a specific kind necessarily causes an agent's failure in reasoning.

This is not to suggest that there is no mental state that could cause the shift in the expected motivational forces of desires. On the contrary, any conception that occurs to an agent in his reasoning process could be such a cause when it arouses the agent's interest in satisfying the desire with an insufficient reason and thereby increases the motivational strength of the desire. When an intervening conception causes an agent to be more interested in doing *a* than doing *b*, even though the agent judges that all things considered, *b* is better than *a*, his strong interest in doing *a* could blind his reason and makes his all-things-considered judgment ineffective. That is, any intervening conception could be the cause of incontinence when it causes the agent's stronger interest in performing the action supported by a relatively weaker reason and increases the motivational strength of the desire in a way that does not properly respond to all reasons the agent has.

Obviously, emotions are not the only kind of mental state which could disturb an agent's reasoning process. Personal characters, customs, or physical impulses could also be the cause of an agent's failure in reasoning. A bold person enjoys the thrill of risk. When there is a typhoon, even though he judges that, all things considered, it would be better to stay at home than to go surfing, he chooses to go surfing. His adventurous personality reinforces the motivational strength of the desire to go surfing, and the stronger desire in turn causes him to ignore the principle of continence. A person is told by his doctor that he should stop consuming food with caffeine as he suffers from heart disease. However, he is used to drinking coffee or tea in the morning, and so although he judges that, all things considered, it would be better not to drink coffee, he serves himself a cup of coffee. It is his custom that adds motivational strength to the desire to drink coffee. A person who is implementing a low-fat eating plan occasionally finds that there is ice cream in the refrigerator. Although he knows that he should not eat ice cream, his

physical impulse increases the motivational force of the desire to eat ice cream. He eats ice cream. In these cases, personal character, custom, and physical impulse play a causal role in the history of an incontinent action as they respectively arouse an agent's strong interest in performing the action he judges to be worse and tilt the motivational force toward the desire to take the action.

Personal character, custom, and physical impulse, like emotion, can be a reason that explains the rational aspect of the action that an incontinent agent performs on the one hand, and be the cause that explains the incontinent agent's failure in reasoning on the other hand. Personal characters, customs, or physical impulses are an agent's tendencies or inclinations to choose in certain ways. When they are perceived and conceived by an agent, they exist as cognitive states in the agent's mind and interact causally and logically with other propositional attitudes in his thinking and reasoning. It is not necessary that their existence causes an agent to fail to detach an unconditional judgment from his all-things-considered judgment, but they could easily influence or interfere with an agent's reasoning.⁴⁰ When they cause an agent to be more interested in an action he judges to be worse and make the desire to take the action superior in motivational strength, they are the causes of incontinence.

Recall that Pears, Thalberg, Lazar, and Baier⁴¹ complain that Davidson does not explain why an incontinent agent's desire with an insufficient reason is motivationally stronger than the

⁴⁰ For example, acting on an emotion could be rational. An agent is rational to act on an emotion if the emotion is considered as a reason for taking the action and is consistent with his all-things-considered judgment and all-out, unconditional judgment. In this case, the emotion is a relevant reason for or against taking an action of a type that the agent considers.

⁴¹ Pears, Thalberg, Lazar, and Baier criticize that Davidson's theory of partitioned mind does not successfully explain the stronger motivational force of the desire with an insufficient reason in an incontinent agent's reasoning. How a desire with an insufficient reason gets the superior motivational force in an incontinent agent's reasoning is indeed a problem, which I have discussed in the present section. However, this problem is not what Davidson

desire with the best reason, why an incontinent agent fails to detach an unconditional judgment from his all-things-considered judgment in his reasoning. If to give an explanation of the motivational strength of the desire with an insufficient reason or the failure in reasoning is to give a specific kind of mental state that accounts for the phenomenon, as what Watson, Mele, de Sousa, and Tappolet have done, then Davidson does not need to do so. This is because identifying a kind of mental state as the cause of incontinence is inconsistent with Davidson's anomalism. And we also find that more than one kind of mental state could serve as causes of incontinence that explain the superior motivational force of the desire causing an incontinent agent to act and explain an incontinent agent's failure in reasoning.

4. PROBABILISTIC AKRASIA

Previously in section 2, I have examined the suggested analogy between modus ponens in theoretical reasoning and the principle of continence in practical reasoning, and here I would like to discuss another analogy proposed by Davidson himself, an analogy between practical reasoning and inductive reasoning, to explain the phenomenon of weakness of will. The appropriateness of this analogy is questioned by Susan Hurley who believes that beliefs are governed by truth alone and so no one could detach a conclusion about what is almost certainly to occur from a probabilistic judgment supported by less evidence. There is no such thing as probabilistic akrasia. In this section, I argue that probabilistic akrasia is possible because conflicting probabilistic judgments involved in a case of probabilistic akrasia could both be true and held by an agent. In section 3 I have explained that there is not a specific kind of mental state

attempts to solve through proposing the theory of partitioned mind. The explanatory force of the theory of partitioned mind will be discussed in chapter IV, section 2.

which functions as the cause of incontinence covering all cases of incontinent action. Emotion, personal character, custom, physical impulse, or any mental state that interferes with an incontinent agent's reasoning could be eligible to explain why an incontinent agent fails to reason in accordance with his best judgment. To the cause of probabilistic akrasia, I hold a similar claim. There is no specific kind of mental state that can explain all cases in which an agent fails to detach a conclusion about what is almost certainly to occur from the probabilistic judgment supported by the strongest evidence. Strong intuition, prejudice, stubborn belief, or any other belief that intervenes in an irrational agent's reasoning could be the cause of his probabilistic akrasia.

4.1 An analogy between practical reasoning and inductive reasoning

Probabilistic akrasia is an example of failure in reasoning that Davidson gives in order to illustrate an incontinent agent's failure in reasoning. This conception is basically inspired by Rudolf Carnap and Carl Hempel. Inductive probability, according to Carnap, is ascribed to a hypothesis with respect to a body of evidence. The hypothesis is a statement concerning unknown facts, such as a prediction of a future event, and the evidence is a statement about the information available to an agent. The relation between the hypothesis and the evidence in a probability judgment is a relation of support or confirmation. The stronger the evidence is, the more likely that the hypothesis is true. This relation of support or confirmation is a matter of degree.⁴² Any probabilistic reasoning in everyday affairs and in science ought to be guided by the principle of learning from experience that 'other things being equal, a future event is to be regarded as the more probable, the greater the relative frequency of similar events observed so

⁴² Rudolf Carnap, 'Statistical and Inductive Probability' (1955), p. 2.

far under similar circumstances'.⁴³ Given this principle, a man's behavior is 'unreasonable if his expectation of a future event were the higher the less frequently he saw it happen in the past, and also if he formed his expectations for the future without any regard to what he had observed in the past'.⁴⁴

Davidson believes that there is an analogy between reasoning about what to do and reasoning about what is almost certainly to occur. On the one hand, the major premise of each kind of reasoning is a judgment in a relational form. For a practical argument, the major premise provided by a desire is a prima facie judgment. A prima facie judgment that $pf(a$ is better than $b, r)$ is an evaluative judgment formed in relation to the reason r . Because the propositional expression of a desire is a prima facie judgment, the conclusion deduced from it and the reason r is also a prima facie judgment, that is, $pf(a$ is better than $b, pf(a$ is better than $b, r)\&r$). An agent who at the same time holds a conflicting desire, say, $pf(b$ is better than $a, e)$ and the reason e , does not thereby hold a contradiction, for what is deduced is a prima facie judgment established in relation to a different set of reason, that is, $pf(b$ is better than $a, pf(b$ is better than $a, e)\&e$). For an inductive reasoning, similarly, its major premise, a probabilistic judgment, is established in relation to a reason. What is directly inferred from a probabilistic judgment $pr(Rx, Sx)$ and a reason Sa is a judgment that $pr(Ra, pr(Rx, Sx)\&Sa)$. When an agent also holds another conflicting probabilistic judgment that $pr(-Rx, Qx)$ and a reason Qa , and he infers that $pr(-Ra, pr(-Rx, Qx)\&Qa)$, he does not hold a contradiction. Accordingly, in a practical reasoning or inductive reasoning, an agent who considers conflicting judgments does not hold a contradiction because the judgments are in the relational form.

⁴³ Ibid., 13.

⁴⁴ Ibid.

On the other hand, an agent who fails to detach an unconditional judgment about what to do from a prima-facie evaluative judgment supported by the best reason is like an agent who fails to detach a conclusion about what is almost certainly to occur from a probabilistic judgment supported by the strongest reason. When an agent holds two prima-facie evaluative judgments that respectively favor and disfavor an action, the conclusions deduced from them are also prima-facie evaluative judgments. None of them can directly cause an action. To decide what to do, an agent needs to take a further step to detach an unconditional judgment from one of the competing conditional judgments. When an agent follows the principle of continence to reason and detach a conclusion from his all-things-considered judgment, a prima-facie evaluative judgment supported by the best reason, he is rational. Otherwise, he is irrational. Analogously, when an agent considers conflicting probabilistic judgments supported by different reasons, a conclusion about what is almost certainly to occur is not deduced from either judgment. It is in a further step that an agent detaches a conclusion about what is almost certainly to occur from one of the competing probabilistic judgments. In the additional step, a rational agent follows the principle of total evidence to detach a conclusion about what is more probable from a probabilistic judgment supported by the strongest evidence, while an irrational agent fails to do so.⁴⁵

4.2 The possibility of probabilistic akrasia

Susan Hurley, however, believes that the suggested parallel between the two cases does not obtain because there is no such thing as probabilistic akrasia. An agent can hold conflicting evaluative judgments about what to do, but he cannot hold conflicting probabilistic beliefs

⁴⁵ Davidson, 'How is Weakness of the Will Possible?', pp. 37-8.

because 'in the case of what should be believed, truth alone governs and it can't be divided against itself or harbor conflicts'.⁴⁶ Once an agent knows what is more probable given all the evidence he has, he can no longer hold the conflicting probability judgment supported by less evidence. Hurley offers an explanation:

Consider someone makes two relational probability judgments, rPq and $eP-q$, and goes on to judge that in relation to all the relevant evidence there is, namely r and e , $-q$ is probable. He couldn't, despite having made the more inclusive relational judgment conjoined with the judgment of no better evidence, believe that q is probable because of the evidence provided by r ; r 's evidential force has been subsumed without remainder by the more inclusive evidence $r\&e$. Less inclusive probabilistic evidence has no constitutive reason-giving force that could hold out in the face of recognition that it's subsumed by the best probabilistic evidence, which favors the opposite conclusion.⁴⁷

Davidson believes that an agent can hold two conflicting probability judgments when they are supported by different bodies of evidence. But Hurley's argument shows that it is not possible for an agent to harbor two conflicting probability judgments because any weak probabilistic evidence is silent when the strongest probability evidence speaks. Hurley approves Pears' assertion that 'If the operation of the perversion of reason is detected and its result is recognized as irrational, it will be likely to collapse completely, ... a perversion of reason should cease the moment that a person becomes conscious of it'.⁴⁸

Following Davidson's principle of charity, Hurley wants us not to attribute probabilistic akrasia to an agent for we can always find a better explanation which makes his beliefs consistent. For example, an agent probably fails to recognize or remember that $-q$ is probable in

⁴⁶ Susan Hurley, 'Conflict, "Akrasia" and Cognitivism' (1985), p. 32.

⁴⁷ *Ibid.*, 31.

⁴⁸ Susan Hurley, 'Conflict, "Akrasia" and Cognitivism', pp. 31-2. David Pears, 'Paradoxes of Systems' (1984), pp. 73-4.

relation to all the evidence. When he sees the mistake, he will either give up the judgment rPq or revise the judgment $(r\&e)P-q$.⁴⁹ When an agent seems to be in a case of probabilistic akrasia, his accepting the probability judgment with less evidence anyway is in fact a case of self-deception. The agent believes the probability judgment with less evidence to be true because he desires it to be true rather than because he accepts the evidence supporting the probability judgment.

In the rest of this section, I explicate the reason for the possibility of an agent's holding conflicting probabilistic judgments. I also offer examples to show that there are cases of probabilistic akrasia which should not be treated as cases of omission or self-deception. Finally, I conclude that, like an agent who fails to reason in accord with his all-things-considered judgment in a practical reasoning because of the disturbance of an emotion, personal character, custom, or physical impulse, an agent could fail to reason in accordance with the probabilistic judgment supported by the strongest evidence in an inductive reasoning because of the interference of a strong intuition, prejudice, or stubborn belief.

Hurley is right to claim that beliefs are governed by truth since to hold a belief is to take the content of the belief to be true, but this is not a reason against the possibility of an agent's holding conflicting probabilistic judgments. Hurley acknowledges the possibility of weakness of will. She believes that an incontinent agent holds conflicting evaluative judgments when he deliberates about what to do. But evaluative judgments, in fact, are also governed by truth. Since Hurley allows an agent to hold conflicting evaluative judgments, she should allow an agent to hold conflicting probabilistic judgments. Furthermore, an agent cannot hold a belief that he takes to be false, and so he cannot hold a contradiction that is obviously false. But two conflicting

⁴⁹ Susan Hurley, 'Conflict, "Akrasia" and Cognitivism', p. 33, p. 38.

evaluative judgments or probabilistic judgments can both be true to an agent, and so they can be held by an agent at the same time.

That an agent holds conflicting evaluative judgments or conflicting probabilistic judgments in a reasoning process is compatible with Hurley's claim that beliefs are governed by truth. When desires are expressed as evaluative judgments, they are truth-evaluable, and so they are governed by truth. Two conflicting evaluative judgments could both be true to an agent when the reasons supporting each of them are different. For example, 'It is desirable to go to the concern tonight because it is free' and 'It is not desirable to go to the concern tonight because no rock band will perform tonight' are conflicting evaluative judgments supported by different reasons. Both could be held true by an agent at the same time. Likewise, two probabilistic judgments that conflict with each other could both be true and held by an agent at the same time. Probabilistic judgments are relational judgments. They are formed in terms of certain pieces of evidence. Since the relation of support or confirmation between the judgment and the evidence in a probabilistic judgment is a matter of degree, each of two conflicting probabilistic judgments could be true to an agent to a degree. An agent can hold two conflicting probabilistic judgments as long as he holds both to be true. An agent, for instance, can hold both that 'Considering that no test was given last week, it is probable that there is a pop quiz tomorrow' and that 'Considering that teaching progress has fallen behind schedule, it is probable that there is not a pop quiz tomorrow' in his reasoning. The two probabilistic judgments are conflicting, but they are true to the agent, and so can be held by the agent at the same time.

It seems that there are some cases in which an agent accepts a probabilistic judgment because of, as Hurley suggests, his omission or his desire for the truth of the probabilistic

judgment, but there are also cases in which an agent accepts a probability judgment with less evidence simply because he accepts the evidence. Given the evidence we have nowadays, scientific or empirical, God probably does not exist. But people who share their Christian testimony show their faith in God. Merely because of the personal experience with God, such as hearing the voice from God, feeling that God is present, even just once, a person could totally change his religious beliefs. Many of these people could be atheists before they accept that God exists. And so they cannot be classified as cases of self-deception in which a person believes in God because he desires that God exists. They believe in God even if the evidence for the belief, the personal experience, is weaker than the evidence supporting an opposite judgment. Or consider the famous O. J. Simpson murder case. The evidence gathered by forensic science indicates that Simpson who is accused of murder is innocent, but many are not persuaded by this stronger evidence. They believe that Simpson is guilty not because they desire him to be guilty or forget the strong evidence for the opposite probabilistic judgment but simply because they tend to accept the evidence collected to prove that Simpson is the murderer.

The conception of the impossibility of probability akrasia seems to suggest that, omission and self-deception apart, an agent always accepts the probabilistic judgment with the strongest evidence in an inductive reasoning. If this is true, then when the same evidence is offered to different agents to judge what is more probable, they must detach a judgment from the probability judgment with the strongest evidence. They will necessarily form the same judgment. This, however, cannot be true. When agents have all the same evidence for a case that is considered, it is possible for them to accept different probability judgments and detach different conclusions. For example, the same evidence is presented in court for the jury to judge whether the defendant is probably innocent, and it may turn out that the defendant is found guilty by

some jurors and not by others. Since people could form different judgments upon the same evidence, it follows that it is not necessary for a person to reason and judge in a certain way. An agent could accept a hypothesis that is not supported by all evidence he has in an inductive reasoning.

It seems that intuition, prejudice, or stubborn belief occurring to an agent in his inductive reasoning could unreasonably enhance a hypothesis that is not supported by all evidence the agent has. When an agent considers what is probable, any of these mental states could interfere into his reasoning process, cause him to ignore the principle of total evidence and draw a false conclusion. An intuition, prejudice, or stubborn belief could be formed on the basis of an unconvincing reason or an unreasonable personal preference for someone or something. When this is the case, an agent who is influenced by his intuition, prejudice, or stubborn belief to draw a false conclusion in his reasoning can not give a rational explanation for such a failure. For example, a roulette player knows that the odds against winning for each number shown on the wheel is 37 to 1, but because he won \$700 yesterday by betting \$20 on number 26, he feels that number 26 will still be a lucky number today. The intuition makes him falsely judge that number 26 has the greater chance of winning, and he repeatedly places bets on that number. The intuition causes the player to draw a false inference but cannot make his false inference reasonable. The multiplicity of mental states that are causes of probabilistic akrasia can be explained by Davidson's anomalism that there are no psychophysical or psychological laws regulating the connection between two types of mental state.

5. CONCLUSION

In this chapter, I explored several problems raised because incontinence is understood as a failure in reasoning. These problems include problems of the possibility for an agent to fail to detach an all-out, unconditional judgment from his all-things-considered judgment, the explanation of the superior motivational force of the desire with an insufficient reason, and the parallel between incontinent action and probabilistic akrasia. In section 2, I defended the possibility of an incontinent agent's failure in reasoning through clarifying the role the principle of continence plays in practical reasoning. That the principle of continence applies to practical arguments which are non-deductive in form explains why an agent who holds the principle of continence in a practical reasoning is not forced to detach an all-out, unconditional judgment from his all-things-considered judgment. In section 3, I argue that there is no mental state of a specific kind the occurrence of which necessarily causes a shift in the expected motivational forces of desires. However, an emotion, personal character, custom, or physical impulse that intervenes in an agent's reasoning process could cause such a shift. In section 4, I discussed the parallel between incontinent action and probabilistic akrasia. I believe that just as an agent could hold conflicting evaluative judgments that are true to him in a practical reasoning, an agent could hold conflicting probabilistic judgments that are true to him in an inductive reasoning. Furthermore, just as an agent could fail to detach a conclusion about what to do from his all-things-considered judgment because of the influence of an emotion, personal character, custom, or physical impulse, an agent could fail to detach a conclusion about what is almost certainly to occur from the probabilistic judgment supported by the strongest evidence because of the influence of an intuition, prejudice, or stubborn belief.

What are crucial to the solutions of these problems are the fact that practical arguments and inductive arguments are non-deductive arguments and Davidson's anomalism thesis that

there are no psychophysical or psychological laws regulating the associations between different types of thoughts and actions. Because practical arguments and inductive arguments are non-deductive arguments, there is not an entailment between the premises and the conclusion of a practical or an inductive argument. The inference from a prima-facie evaluative judgment to a conclusion about what to do, or the inference from a probabilistic judgment to a conclusion about what is almost certainly to occur, is not an automatic transition. An agent does not necessarily detach a conclusion from his all-things-considered judgment or from the probabilistic judgment supported by the strongest evidence. Since there are no psychophysical or psychological laws which regulate how a thought is connected with another thought or action, any mental state interfering with an agent's reasoning could cause him to fail to reason in accordance with his all-things-considered judgment or with the probabilistic judgment supported by the strongest evidence.

IV. IRRATIONALITY AS AN INNER INCONSISTENCY

Donald Davidson proposes the theory of partitioned mind to explain how cases of irrationality are possible given the assumption that no single propositional attitude can exist without causally and logically connecting with other propositional attitudes. For each case of irrationality, there is a mental cause that is not a reason for the mental state it causes. It brings about an inner inconsistency within an agent, which makes the agent irrational. Hence, to Davidson, the problem of irrationality is a problem of inner inconsistency. The theory of partitioned mind explains the possibility of the existence of a mental cause that is not a reason for its effect and the inner inconsistency caused by it.

The theory of partitioned mind faces various criticisms. In this chapter, I discuss several lines of critical argument, aiming thereby to explicate Davidson's explaining irrationality as an inner inconsistency. First, some commentators question the explanatory force of the theory of partitioned mind. They think that drawing a boundary between two inconsistent propositional attitudes or dividing them into different parts of the mind does not tell us how an agent reasons irrationally. Second, some philosophers believe that an agent's holding two inconsistent propositional attitudes is intelligible to us when a coherent causal nexus in which the two inconsistent propositional attitudes occur is offered, and so there is no need to propose the theory of partitioned mind to explain the existence of a mental cause that is not a reason for the mental state it causes and the phenomenon of inner inconsistency. Third, many tend to deny the interdependent relations between reasons, consistency, inconsistency, rationality, and irrationality, and so challenge Davidson's taking inner consistency as the standard of rationality and inner inconsistency as the standard of irrationality.

I defend Davidson's theory of partitioned mind in this chapter. Section 1 begins with a brief review of Davidson's theory of partitioned mind. Section 2 investigates the explanatory force of the theory of partitioned mind. I argue that the theory of partitioned mind successfully explains the possibility of inner inconsistency because it rules out the situation in which inner inconsistency cannot occur, that is, the situation in which two inconsistent propositional attitudes are put together. Section 3 explains the necessity of the theory of partitioned mind. I argue that the theory of partitioned mind is necessary because it is only if two inconsistent propositional attitudes are located in different parts of the mind that no explicit or implicit contradiction is yielded by the conjunction of the inconsistent propositional attitudes, and so an agent can hold both beliefs at the same time. Section 4 explores the interdependent relations between reasons, consistency, inconsistency, rationality, and irrationality. I argue that it is appropriate to take inner consistency as a standard of rationality because a reasoning process consists of propositions with truth values to which consistency is a property that is attributable. Section 5 is a conclusion of this chapter.

1. A REVIEW OF DAVIDSON'S THEORY OF PARTITIONED MIND

Davidson's analysis of weakness of the will consists of two parts, namely, a description of the phenomenon of weakness of the will and an explanation of the cause of the phenomenon. Davidson describes the phenomenon of weakness of the will in a way that is compatible with the assumption that an agent always acts on what he judges to be better. And he identifies a mental state as the cause of the phenomenon. The mental state is a mental cause which bears no rational relation with what it causes. Davidson tries to explain how a mental cause that is not a reason for the mental state it causes exists in the mind under the assumption that propositional attitudes are

causally and logically connected with one another.

Davidson believes that an agent who acts intentionally acts on what he judges to be better. According to his causal theory of action, an intentional action is caused by a pair of belief and desire that rationalizes the action. The pair is not only a reason for but also a cause of the action. When we try to understand why an agent does something, identifying a pair of belief and desire that causes him to act will explain why he takes the action. Corresponding to an explanation of this kind, a practical argument can be constructed to show how an agent infers from a desire for an end and a belief that actions of a certain kind would help to reach the end to the conclusion that taking an action of that kind is desirable. This argument would show how he reasons from an end to a means. In a practical argument, the major premise provided by a desire is expressed as a prima facie, conditional judgment that $pf(a \text{ is better than } b, r)$, the minor premise provided by a paired belief is the content of the belief r , and the conclusion inferred from the two premises is an intention, an all-out, unconditional judgment that a is better than b . When an agent acts on the conclusion of his practical reasoning, he acts on what he judges to be better. An intentional action is a result of a practical reasoning, and so it is a case in which an agent acts on what he judges to be better.

Given the principle that an agent always acts on what he judges to be better, cases in which an agent acts against his best judgment are possible because the two evaluative judgments involved are of different kinds and the two kinds of evaluative judgments are not in a relation of entailment.¹ Before an agent acts, he weighs reasons for and against taking some sort of action. What he considers are competing prima facie, evaluative judgments conditioned by certain

¹ For a more detailed account of the possibility of weakness of the will, see Chapter III, section 2.

reasons. When the agent decides what to do, he detaches an all-out, unconditional judgment from one of the competing prima facie, conditional judgments. The competing prima facie judgments have the form '*pf(a is better than b, r)*', and the conclusion coupled with an action is an unconditional judgment with the form '*a is better than b*' or 'It would be better to do *a* than to do *b*'. Because the two kinds of evaluative judgments are not in a relation of entailment and the inference from a prima facie, conditional judgment to an all-out, unconditional judgment is a non-deductive argument, an agent who holds an all-things-considered judgment that *pf(a is better than b, e)* does not necessarily infer an unconditional judgment from it. He could fail to reason consistently and detach an unconditional judgment that *b* is better than *a* from a competing prima facie judgment that *pf(b is better than a, r)*. When the agent acts on the unconditional judgment inconsistent with his all-things-considered judgment, although he acts against his best judgment, he acts on his value judgment.

Davidson takes the cause of weakness of will to be a mental cause that is not a reason for the mental state that it causes, and he gives an account which explains how the existence of such a mental cause is compatible with the assumption that propositional attitudes that an agent has are causally and logically connected with each other. Moreover, Davidson believes that the mental cause brings about an inconsistency between an incontinent agent's propositional attitudes and makes him irrational, and so the account also provides an explanation of how an agent harbors inconsistent propositional attitudes in his mind. What follows is a brief outline of the account.

Davidson identifies the desire causing an incontinent agent to act as the cause of his being incontinent. The desire, on the one hand, exhibits the rational aspect of the action. It serves

as a cause of and a reason for the action that the incontinent agent performs. Because the desire rationalizes the action, there is a reason relation between the desire and the action. On the other hand, the desire makes the action an irrational action. It causes the agent to override the principle of continence and act against his all-things-considered judgment, but does not rationalize what it causes. Since the desire is a mental cause of, but not a reason for, the agent's acting against his best judgment, there is a breakdown of the reason relation between the agent's propositional attitudes.

Although the incontinent agent does not hold contradictory judgments in his reasoning, he holds inconsistent judgments when his strong desire to perform the present action causes him to ignore the principle of continence. The principle of continence, a principle of rationality, is intrinsic to an agent who can think and reason. It requires an agent to perform the action judged best on the basis of all available relevant reasons. The incontinent agent who holds the principle knows that he should act on his best judgment. He has an all-things-considered judgment about what is best for him to do, but he detaches an unconditional judgment from another competing, conflicting *prima facie* judgment. In other words, the incontinent agent holds a trio of judgments which are inconsistent. That is, (1) the principle of continence, (2) an all-things-considered judgment that, say, *pf(a is better than b, e)*, and (3) an unconditional judgment that *b* is better than *a*. The unconditional judgment is not contradictory to the all-things-considered judgment, but it is literally inconsistent with the all-things-considered judgment enjoined by the principle of continence.² Davidson takes inner consistency as a fundamental norm of rationality, and inner inconsistency a manifestation of irrationality. In this sense, irrationality lies in the inconsistency of the unconditional judgment with other attitudes and principles. The incontinent agent who

² Donald Davidson, 'Incoherence and Irrationality' (1985), reprinted in *Problems of Rationality* (2004), p. 194.

holds a set of inconsistent judgments is an irrational agent.

Davidson proposes the theory of partitioned mind to explain the compatibility between the assumption that no propositional attitude can exist without causally and logically connecting to other propositional attitudes, the existence of a mental cause that is not a reason for the mental state it causes, and the state of inner inconsistency. According to Davidson, the mind is divided into parts. The interaction between two propositional attitudes determines whether or not they belong to the same subdivision of the mind. The interactions among propositional attitudes in a subdivision are consequences of following the principle of logic. In a subdivision, a propositional attitude that causes another must also be a reason for holding it. The two propositional attitudes that are in a causal relation are in a reason relation. On the contrary, when the causal relation between two propositional attitudes remains while the logical relation between them is gone, the two propositional attitudes are apart and respectively belong to different subdivisions. Davidson remarks that ‘The breakdown of reason relations defines the boundary of a subdivision... The parts are defined in terms of function; ultimately, in terms of the concepts of reason and of cause.’³ Accordingly, a boundary is drawn between inconsistent propositional attitudes held by an irrational agent for they are in a non-logical causal relation.

The boundary drawn between inconsistent propositional attitudes limits the normative force of a propositional attitude to the subdivision where it is located and makes cases of irrationality possible. An agent does not believe the conjunction of two inconsistent beliefs that cannot be true at the same time. But the boundary separates the two inconsistent propositional attitudes and allows an agent to hold them simultaneously without putting them together.

³ Donald Davidson, ‘Paradoxes of irrationality’ (1982), reprinted in *Problems of Rationality* (2004), p. 185.

Because of the separation, the normative force of a principle of rationality and judgments compatible with it is restricted within a subdivision to which they belong. The normative force is not effective against an inconsistent judgment that is separated into a different subdivision. The agent is then able to hold the inconsistent judgments. As Davidson puts it, ‘The irrationality of the resulting state consists in the fact that it contains inconsistent beliefs; the irrational step is therefore the step that makes this possible, the drawing of the boundary that keeps the inconsistent beliefs apart.’⁴ In this sense, the boundary is what makes all cases of irrationality possible.

2. THE EXPLANATORY FORCE OF THE THEORY OF PARTITIONED MIND

The theory of partitioned mind depicts a model in which an agent who holds principles of rationality could occasionally be irrational, which includes cases of incontinent action. Nevertheless, the theory of partitioned mind is frequently criticized with respect to its putative explanatory force. Commentators, such as David Pears, Irving Thalberg, Annette Baier, and Ariela Lazar, deny that the theory of partitioned mind successfully explains incontinent action or other cases of irrationality. They think that dividing inconsistent propositional attitudes into different parts of the mind does not tell us why a judgment supported by the strongest reason is overcome by a competing judgment with an insufficient reason in an incontinent action. In this section, I examine arguments against the explanatory force of the theory of partitioned mind. I will first clarify that the main purpose of the theory of partitioned mind is not to explain the superior motivational force of a desire with an insufficient reason in a case of incontinent action, but to explain the possibility of inner inconsistency. Then, I discuss the explanatory force of the

⁴ Donald Davidson, ‘Deception and Division’ (1986), reprinted in *Problems of Rationality* (2004), p. 211.

theory of partitioned mind. I will argue that the theory of partitioned mind does explain the possibility of inner inconsistency. Its explanatory force lies in the fact that it rules out the situation in which an agent cannot hold inconsistent propositional attitudes, that is, the situation in which two inconsistent propositional attitudes are put together and form an explicit or implicit contradiction.

David Pears believes that Davidson's theory of partitioned mind is capable of covering all cases of irrationality which are cases of inner inconsistency, but he argues that such an accomplishment is merely achieved by definition, which deprives the theory of its explanatory power. In a case of irrationality, Pears claims, an agent knows that the judgment he favors violates a rational constraint and endorsing it will make him irrational. The proper function of the cautionary belief is to intervene and stop irrationality. When it fails to do so, a boundary is drawn between inconsistent beliefs, which enables an agent to hold them simultaneously and be irrational. In this sense, the theory of partitioned mind covers all cases of irrationality for 'In such cases the irrationality is always the result of the cautionary belief's failure to intervene and stop it and, whenever a failure of this kind occurs, a line is drawn between main system and sub-system'.⁵ However, Pears argues that this achievement is obtained because a subdivision's boundary is defined as a line across which some element in a person's psyche fails to exercise its normal rational effect on the elements belonging to different subdivisions. 'That definition guarantees a perfect fit between the functional theory of systems and the phenomenon of irrationality that the subject is competent to avoid, ...'⁶ Since the achievement of the theory is obtained by definition, the theory has no explanatory power at all.

⁵ David Pears, 'Paradoxes and Systems' (1984), 84.

⁶ Ibid.

If the theory of partitioned mind is to have explanatory power, Pears argues that it must give a permissive cause of irrationality which is not tied to definition but is ‘an empirically discoverable cause and one whose presence can be established independently of its supposed effect, the irrationality.’⁷ Freud, for example, identifies the failure of consciousness as the permissive cause of irrationality. If an agent consciously believes that he is irrational to form a belief, this cautionary belief will prevent him from forming it. Otherwise, he will form the belief and be irrational. As a permissive cause of irrationality, the failure of consciousness is empirically discoverable and distinguished from its effect. Its explanatory force is not given by definition. Unfortunately, drawing the boundary between inconsistent beliefs as the permissive cause of irrationality has none of these features.

Furthermore, according to Pears, the theory of partitioned mind should explain why the cautionary belief fails to intervene and stop irrationality effectively, but it does not. The boundary drawn between inconsistent beliefs restricts the function of the cautionary belief within the subdivision to which it belongs so that the cautionary belief does not exhibit its rational force across the line. Pears writes that ‘we might hope for a psychological account of the junction between the two systems which would explain the blockage. But that would be an empty dream... The criterion used by the functional theory when it draws the line is simply that that is where the blockage occurs.’⁸ Irving Thalberg, Annette Baier, and Ariela Lazar share with Pears a similar view.⁹ They seem to attribute the function of Pears’ cautionary belief to an all-things-considered judgment, and think that the theory of partitioned mind does not tell us why the best

⁷ Ibid.

⁸ Ibid., 87.

⁹ See Irving Thalberg’s ‘Questions about Motivated Strength’ (1985), 101-103, Annette Baier’s ‘Rhyme and Reason: Reflections on Davidson’s Version of Having Reasons’ (1985), 121, and Ariela Lazar’s ‘Akrasia and the Principle of Continence or What the Tortoise Would Say to Achilles’ (1999), 393-394.

judgment enjoined by the principle of continence fails to make an agent stick to it. That is, the theory of partitioned mind does not show us why a desire with an insufficient reason gets superior motivational force in a case of incontinent action.

Previously, in Chapter III, Section 3, I have argued that there are no psychological states of a specific kind whose occurrence guarantees the occurrence of cases in which an agent fails to reason consistently about what to do. Although Davidson points out that it is the strong desire causing an incontinent agent to act that causes him to override the principle of continence and so be irrational, Davidson does not explain why the desire with an insufficient reason gets superior motivational strength in the reasoning process. Hence, to commentators who complain that the theory of partitioned mind does not explain why an incontinent agent's all-things-considered judgment is overridden by another competing *prima facie* judgment formed on the basis of a relatively weaker reason, i.e., why the judgment enjoined by the principle of continence fails to prevent an incontinent agent from being irrational, merely identifying the strong desire as the cause of incontinence is not satisfactory. They expect to find a mental state that can explain the inefficacy of the all-things-considered judgment. My response is that there is no unique kind of mental state that serves as the cause of incontinence, but any mental state the agent has could interfere with the agent's reasoning when he is aware of it. When the intervening mental state is recognized by the agent, it could increase the motivational strength of the desire with an insufficient reason in a way that is not proportional to the reasons he has. Emotion, personal character, custom, physical impulse, or other mental states could play such a role in a reasoning process and be the cause of incontinence.

However, explaining the superior motivational strength of the desire with an insufficient

reason in a case of incontinent action is not what Davidson intends to do when he comes up with the theory of partitioned mind. The theory of partitioned mind mainly aims to explain how a mental state can exist as a cause of, but not a reason for, another mental state - how an agent can harbor inconsistent beliefs simultaneously in his mind. Because irrationality is understood in terms of inner inconsistency, the theory of partitioned mind which explains cases of inner inconsistency explains cases of irrationality, including weakness of the will, self-deception, probabilistic akrasia, and any phenomenon that consists of a set of inconsistent propositional attitudes. According to the theory of partitioned mind, whenever a boundary is drawn between two inconsistent propositional attitudes, it is possible that the agent holds them simultaneously. Otherwise, it is not. Thus, drawing the boundary gives a condition in which cases of irrationality are possible.

Pears claims that for the theory of partitioned mind to have explanatory force, it must give a permissive cause that is empirically discoverable, distinguished from its effect, and not tied to definition, but drawing the boundary does not have these characteristics. The permissive cause of irrationality that Pears mentions seems to be an event or state whose occurrence allows irrationality to occur, and he seems to take drawing the boundary as the permissive cause given by the theory of partitioned mind. If this is the case, then Pears is right to think that drawing the boundary is not empirically discoverable because the boundary, to Davidson, is a concept of the theory. Davidson remarks that 'I postulate such a boundary somewhere between any (obviously) conflicting beliefs. Such boundaries are not discovered by introspection; they are conceptual aids to the coherent description of genuine irrationalities.'¹⁰ In other words, the boundary is merely an idea which helps to delineate the division of the mind. When an agent holds inconsistent beliefs,

¹⁰ Davidson, 'Deception and Division', p. 211.

the inconsistent beliefs are separated as if there is a boundary drawn between them. The boundary is neither a mental event nor an object, and so it is not a permissive cause of irrationality.

Although the theory of partitioned mind does not give a permissive cause of irrationality as Pears demands, the theory does not thereby lose its explanatory force. What the theory aims to explain is how an agent can harbor inconsistent beliefs in his mind, and it gives the answer that separating inconsistent beliefs into different parts of the mind enables an agent to hold them at the same time. This answer indeed solves the problem because it rules out the situation that prevents an agent from holding inconsistent beliefs, that is, the situation in which two inconsistent beliefs are put together and form an explicit or implicit contradiction. Beliefs that are inconsistent are necessarily not all true. That two beliefs are inconsistent entails that they cannot be true at the same time. When one is true, the other must be false. It could be the case that two inconsistent beliefs are contradictory with each other when, say, one is P and the other is its denial $\neg P$. The conjunction of P and $\neg P$ is an explicit contradiction ($P \& \neg P$). It could also be the case that two beliefs are inconsistent but not contradictory with each other when one is P and the other is a statement from which, together with other beliefs that an agent accepts, $\neg P$ can be inferred. The two beliefs constitute an implicit contradiction. An agent does not believe the conjunction of two inconsistent beliefs for the conjunction is a contradiction, explicit or implicit, and cannot be true. Therefore, putting inconsistent beliefs together is a situation in which an agent cannot hold them at the same time. And the theory excludes this situation.

Accordingly, cases of irrationality, such as self-deception or weakness of will, are possible because the involved inconsistent beliefs are apart from each other. A self-deceiver

holds opposite judgments. A belief that he does not want to believe to be true sustains another belief that he inclines to believe to be true.¹¹ For example, a man knows that he is being cheated by his wife. Because he does not want to accept the truth, he deceives himself into believing that he is not being cheated by his wife. The former causes the latter that contradicts it, and so they are inconsistent and contradictory with each other. The two beliefs together form an explicit contradiction which cannot be held by an agent. Only when they are separated, can an agent hold them at the same time. An incontinent agent does not hold contradictory beliefs because his conflicting all-things-considered judgment and unconditional judgment are with different forms and are not in an entailment relation. The incontinent agent holds an all-things-considered judgment that $pf(a \text{ is better than } b, e)$, a relevant reason e , the principle of continence, and other supporting beliefs, but he does not form a conclusion about what to do from this set of beliefs. If he continued to reason, he would infer that a is better than b . However, the incontinent agent detaches an unconditional judgment that b is better than a from another set of beliefs. In other words, his unconditional judgment is implicitly contradictory with the conclusion that could be inferred from the all-things-considered judgment and other beliefs that he accepted. The conjunction of all of the incontinent agent's beliefs is an implicit contradiction which is not acceptable to an agent. Yet, when the inconsistent beliefs are apart from each other, an agent can hold them at the same time. For other cases of irrationality, similarly, separating inconsistent beliefs prevents the formation of an explicit or implicit contradiction and makes it possible for an agent to hold inconsistent beliefs at the same time.

Two inconsistent beliefs do not form a contradiction when they are separated from each other, but why doesn't an agent's holding two beliefs entail his holding the conjunction of them?

¹¹ Davidson, 'Deception and Division', p. 199.

John Williams says that ‘To believe that p and to believe that q is not necessarily to believe that p and q. Were this *not* the case then all cases of holding beliefs inconsistent with each other would be cases of holding self-contradictory beliefs.’¹² He enumerates examples to show that inconsistency is not the same as contradiction. For instance, an agent may present an argument containing many premises, including p and including $\neg p$. He could believe p and believe $\neg p$ without believing $p \& \neg p$ if he has only considered the two contradictory beliefs separately and sincerely believes each of them. Moreover, an agent who holds a large number of beliefs may not be able to understand or consider their conjunction, and therefore without being able to believe it. In these cases, inconsistent sets of beliefs are not only possible but also hard to avoid especially when there are a large number of beliefs in the set. An agent is allowed to hold inconsistent beliefs without holding a contradiction when he does not or is not able to discover the fact that the conjunction of the inconsistent beliefs he accepts is a contradiction.

Williams seems to think that, due to omission or incompetence, an agent is allowed to hold inconsistent beliefs without holding a contradiction. Once the agent discovers that beliefs he holds are contradictory, he should give up one or the other. This is because for contradictory beliefs p and $\neg p$, any evidence for the truth of p will prove $\neg p$ to be false and vice versa. The agent is not justified in continuing to hold contradictory beliefs.¹³ However, in the various types of irrationality that Davidson deals with, irrational agents are assumed to be aware of all beliefs that are inconsistent. In an extreme case of irrationality, such as a case of self-deception, an agent remains aware of the fact that his evidence favors the belief that he does not want to accept. For it is this awareness that motivates the agent to hold an opposite belief. Thus, it is quite possible

¹² John Williams, ‘Inconsistency and Contradiction’ (1981), p. 600.

¹³ *Ibid.*, p. 601.

that the agent appreciates that beliefs which he holds are contradictory. If so, how can an agent consciously hold beliefs that are inconsistent or even contradictory with each other without holding their conjunction?

The legitimacy of the separation of inconsistent beliefs could come from the fact that each of the inconsistent beliefs is supported by a set of interlocking beliefs held true by an agent. In a case of irrationality, the main reasons that an agent has for each of the inconsistent beliefs are usually different but are held true by the agent. For example, an agent judges that having a piece of French bread with garlic spread would be desirable because it is delicious. However, thinking about the fact that garlic spread consists of hydrogenated oil which could increase the risk of heart disease and cause other health problems, the agent also judges that it would be better not to consume it. Suppose that the agent takes the latter to be his best judgment which is enjoined by the principle of continence he holds but finally judges unconditionally that it would be better to eat the bread. The agent then holds inconsistent beliefs. Because the agent has reasons for each of his inconsistent beliefs, he can still hold them even if he recognizes that the beliefs he holds are inconsistent. The situation in which the agent would get rid of one of the beliefs and makes his beliefs consistent is that he finds out that the reason supporting it is wrong. He no longer holds the belief because he has no reason to do so. Hence, as long as the agent has reasons for what he believes, he can hold the beliefs even if they are inconsistent.

In an extreme case of irrationality, such as self-deception, an agent holds p and $\neg p$. How can the agent hold obviously contradictory beliefs without holding the conjunction of them? Williams denies that an agent who discovers the beliefs he holds are contradictory is justified in continuing to hold them for any of the agent's evidence for one belief will be evidence against

the other and vice versa. This seems not right. For whether an agent believes p or $\neg p$ is not always determined by a single piece of evidence the acceptance of which represents that to the agent only p or only $\neg p$ is true. Sometimes an agent holds p and $\neg p$ because they are respectively supported by different sets of evidence. In a case of self-deception, the agent has stronger evidence for a judgment he knows to be true. However, in order to avoid the truth that hurts but accept an opposite judgment, the agent irrationally overemphasizes the weaker evidence for the opposite judgment that he favors. Because the evidence for two opposite judgments is different and mistakenly appears to have equal force to the agent, they are allowed to coexist. The only thing that is impossible is that the agent is caused to hold that $p \& \neg p$ by the same set of evidence.

Accordingly, Davidson's theory of partitioned mind aims to explain the possibility of an agent's holding inconsistent beliefs at the same time, and what does the trick is the separation of a belief from another that is inconsistent with it and the supporting relation between beliefs and different sets of reasons. Davidson understands irrationality as an inconsistency within a person. Because two inconsistent beliefs together form an explicit or implicit contradiction which is unacceptable to an agent, the theory of partitioned mind excludes the situation in which an agent is not allowed to hold inconsistent beliefs, that is, the situation in which they are put together. Only when inconsistent beliefs are separated from each other, can an agent hold them simultaneously. The separation is neither an arbitrary step nor a stipulation of the theory. Its legitimacy relies on the fact that inconsistent beliefs coexisting are respectively supported by different sets of reasons. No one can sincerely believe a contradiction for it is obviously false. But two inconsistent beliefs caused by different sets of reasons could both appear true to an agent. As an agent has reasons for each of his beliefs that are inconsistent, he is justified in holding them without holding their conjunction.

Here the compatibility between Davidson's rational assumptions for thinking creatures and cases of irrationality is evident. Davidson claims that no propositional attitude can be isolated from other propositional attitudes but exist independently in the mind. Instead, a propositional attitude is causally and logically connected with other propositional attitudes, and is identified by its relation with them. I have explicated that what allows an irrational agent to hold inconsistent beliefs without holding their conjunction is that the agent has different sets of reasons for each of the beliefs. That a belief is supported by an interlocking set of propositional attitudes partially makes irrationality possible. In this sense, the rational assumption that propositional attitudes are causally and logically connected with each other is not only compatible with but also necessary for cases of irrationality.

To sum up, Pears believes that Davidson's theory of partitioned mind does not have explanatory force because, in his view, the boundary has the function of restricting the rational force of a propositional attitude within a subdivision to which it belongs simply because it is so defined. Some other philosophers also complain that the theory of partitioned mind does not improve our knowledge about why judgments with sufficient reasons or principles of rationality are ineffective in cases of irrationality. This is an illusion. I have clarified that the aim of the theory of partitioned mind is to give a condition in which inconsistent beliefs can coexist, rather than that of identifying psychological states of a specific kind to explain how a desire with an insufficient reason gets superior motivational force to override the principle of rationality. Furthermore, I believe that there are no psychological states of a specific kind the occurrence of which guarantees the occurrence of incontinence. It is futile to look for such a psychological state.

It seems to me that it would be easier for us to see the explanatory force of the theory of partitioned mind if we focus on the separation of two inconsistent beliefs rather than on the imaginary boundary drawn between inconsistent beliefs. The boundary between inconsistent beliefs is a conceptual aid which helps to sketch a picture of divided mind in which irrationality is possible. The boundary, in fact, represents the separation of a belief from another that is inconsistent or even contradictory with it in an agent's mind. The theory of partitioned mind aims to explain the possibility of inner inconsistency, and it explains that an inner inconsistency occurs when beliefs that are inconsistent are separated into different parts of the mind. I have argued that such an account does have explanatory power because it rules out the situation in which inner inconsistency is impossible, that is, the situation in which two inconsistent beliefs are put together and form an explicit or implicit contradiction. I also explicated that reasons supporting each of the inconsistent beliefs legitimate an agent's holding them separately. As long as he has reasons to hold each of the beliefs that are inconsistent, he is not forced to give up one or the other. The situation is that he is then in a state of being inconsistent.

3. THE NECESSITY OF THE THEORY OF PARTITIONED MIND

To Davidson, cases of irrationality are possible when the inconsistent beliefs involved are separated into parts. For when inconsistent beliefs are put together, an explicit or implicit contradiction will be formed. No one can hold a contradiction. Only when inconsistent beliefs are apart from each other, can an agent hold them at the same time. Hence, dividing the mind into parts is necessary for an account of the possibility of irrationality.

The necessity of dividing mind into parts to explain irrationality is doubted by Alfred Mele. Mele agrees with Davidson that the existence of a mental cause that is not a reason for the mental state it causes is required for the explanation of irrationality, but he denies that a theory like the theory of partitioned mind is necessary for accommodating such a mental cause. Mele believes that in his account the mental cause and its effect which bear no rational relation to each other are elements in a coherent action-generating causal nexus, and so constructing the theory of partitioned mind to explain such a mental cause is superfluous.¹⁴

Although Mele provides us with an alternative way to think about the problem of irrationality, his analysis seems not to be more convincing than Davidson's theory of partitioned mind. In this section, I defend the necessity of the theory of partitioned mind. I will argue that what is essential to the phenomenon of irrationality is the irrational agent's being inconsistent and that the theory of partitioned mind is required to explain such an inconsistency. Moreover, it is the existence of inner inconsistency in a reasoning process that confuses us. With respect to this problem, Mele's argument does not eliminate our confusion but Davidson's theory of partitioned mind does.

Mele begins his argument with enumerating several cases in which the mental cause and its effect are in a non-logical causal relation but can be attributed to the same non-divided agent without being paradoxical. For example, Bart uses an acronym in preparing for the test. He first thinks the acronym, which causes him to think words corresponding to the letters consisting of the acronym. His thinking of the acronym is the cause of, but not a reason for, his thinking the words. In this case, although the mental cause is not a reason for its effect, the mental cause

¹⁴ Alfred Mele, 'Strict Akritic Action: A Paradox of Irrationality' (1987), p. 79.

operates in an intelligible way to generate the effect as an intended result. The case has coherence and intelligibility of its own. Mele argues that, if in a case of ‘reason-cause’ it is coherence that inclines us to place cause and effect in the same mental system, we should be equally so inclined in the case of Bart. Assigning the mental cause and its intended effect in the same non-divided agent is not at all paradoxical. Since a mental cause that is not a reason for its effect can be placed in the same non-divided agent without causing any paradox, dividing the mind into parts to explain the existence of such a mental cause is unnecessary.¹⁵

Similarly, in an incontinent action, Mele believes that the mental cause of irrationality and the agent’s acting irrationally are elements in a coherent action-generating causal nexus, and so separating them into different parts of the mind to explain the incontinent action is not necessary. Mele claims that in an incontinent action, the mental cause of irrationality, the stronger desire causing the agent to act, has an influence on the agent’s motivational condition. For instance, ‘a person’s focusing on and vividly representing the prospective pleasant results of an action, *A*, that is contrary to his better judgment may increase his motivation to do *A* even though these mental events are not *reasons* for him to be more motivated to do *A*. At the same time, such mental events may block various routes of resistance.’¹⁶ In this illustration, the incontinent agent is attracted by the imaginary pleasure brought about by doing *A*. Although the agent judges that *A* is worse than other available actions, the desire to do *A* and the pleasure associated with *A* causes him to do *A*. The mental cause is the desire to do *A*, and the effect is the agent’s irrationally performing *A*. The mental cause and its effect are in a coherent action-

¹⁵ Ibid., 77-78.

¹⁶ Ibid., 79.

generating causal nexus and their relation appears intelligible to us. Separating them into parts is unnecessary.

It seems to me that Mele's argument against the necessity of the theory of partitioned mind is problematic. Bart's case is simply irrelevant to the problem of irrationality. A case of irrationality basically contains competing and conflicting beliefs, but no competing or conflicting beliefs are involved in Bart's case. A case of irrationality is a process of reasoning that fails to apply a certain principle of rationality. For instance, an incontinent action is a failure of reasoning in which an agent fails to follow the principle of continence to detach a conclusion about what to do from his all-things-considered judgment. Probabilistic akrasia is a failure of reasoning in which an agent fails to follow the principle of total evidence to detach a conclusion about what is most likely to occur from a probabilistic judgment supported by the strongest evidence. But whether Bart's case has a reasoning process or not is unclear. Even if it does, it is not a reasoning violating any principle of rationality. The problem of irrationality is puzzling because an irrational agent intentionally reasons against principles of rationality that he accepts or holds. Apparently, Bart's case has nothing to do with irrationality and is not puzzling at all.

Mele seems to think that Davidson's theory of partitioned mind is designed to cover every mental cause that is not a reason for its effect, and so any case which shows that such a mental cause can be explained without dividing the mind into parts will prove that the theory of partitioned mind is not necessary. This is incorrect. For each case of irrationality, there is a mental cause that is not a reason for its effect which brings about an inner inconsistency within a person. The theory of partitioned mind attempts to explain cases of irrationality which are cases of inner inconsistency. However, the existence of a mental cause that is not a reason for its effect

is not sufficient to form a case of irrationality. An agent who arbitrarily associates two mental events in his mind is not an irrational agent for the association does not cause an inner inconsistency within the agent. Bart's case is an arbitrary association of this kind. Relevant beliefs that Bart holds are all true at the same time. Since there is no inner inconsistency in Bart's case, the theory of partitioned mind is not applicable to it. Thus, Bart's case or other similar cases are not examples against the necessity of the theory of partitioned mind.

In addition, Mele believes that in his description, without appealing to the theory of partitioned mind, the mental cause of irrationality and its effect are elements in a coherent action-generating causal nexus and appear intelligible to us. Unfortunately, the way Mele explains incontinent action says nothing more than what Davidson has told us, and leaves the problem of the possibility of irrationality unsolved. In Mele's description, imagining that pleasure will come with *A* or holding other thoughts compatible with the desire to perform *A* are in fact a set of interlocking propositional attitudes supporting the desire to perform *A*. Davidson makes it clear that the desire causing an incontinent agent to act plays two roles. On the one hand, it is a reason that causes the agent to take the action that he performs, and on the other hand, it causes the agent to ignore the principle of rationality but is not a reason for his doing so. Mele claims that vividly representing the prospective pleasant results of *A* may block various routes of resistance. Since the imaginary pleasure is a mental event supporting the desire to do *A*, when the desire to do *A* causes the agent to ignore the principle of continence, the psychological force preventing the agent from doing *A* is gone. However, the principle of continence, the agent's all-things-considered judgment favoring an action other than *A*, and other supporting reasons are still there. These beliefs and the agent's unconditional judgment to do *A* together cause an inner

inconsistency in the agent's mind. Mele's account of irrationality does not explain the possibility of such an inner inconsistency.

So far I have shown that neither Bart's case nor the explanation of incontinent action given by Mele can support the rejection of the necessity of the theory of partitioned mind. The main problem that Mele's argument presents is that Mele fails to grasp the nature of irrationality and so he cannot appreciate the necessity of the theory of partitioned mind. What is essential to a case of irrationality is the breakdown of the reason relation between beliefs, the state of inner inconsistency. In a case of irrationality, beliefs involved in an agent's reasoning are reasons the agent considers in order to form a conclusion. As reasons considered in a piece of reasoning, they are supposed to connect with each other in a logical way. The principle of continence and the all-things-considered judgment supported by the best reasons, for example, are assumed to have the strongest causal power. In an incontinent action, they temporarily lose their normative force. If they worked normally, they would cause a consistent unconditional judgment. But the expected function shuts down. The incontinent agent forms a conclusion about what to do against the principle of continence and the all-things-considered judgment, which breaks the proper logical relation between beliefs he holds and causes an inner inconsistency. Mele's examples do not capture this significant characteristic of irrationality. Bart's associating an acronym with words represented by letters consisting of the acronym does not break any logical relation and cause an inconsistency since no specific logical relation or any rational force is assumed to exist between the two mental events. Mele's account of incontinent action simply ignores the problem of inner inconsistency. Once the fact that inner inconsistency is essential to a case of irrationality is recognized, the necessity of the theory of partitioned mind will be seen.

The theory of partitioned mind is necessary in order to explain the inner inconsistency caused by the breakdown of the reason relation between beliefs in a reasoning process. But is it appropriate to take inner consistency as a criterion of rationality, and inner inconsistency a mark of irrationality? The following section is a consideration of this question.

4. CONSISTENCY, RATIONALITY, AND IRRATIONALITY

Davidson takes irrationality to be an inner inconsistency in an agent's propositional attitudes, and his theory of partitioned mind is proposed to explain the possibility of this sort of irrationality. In section 2, I clarified that the theory of partitioned mind explains the possibility of irrationality by separating inconsistent beliefs the conjunction of which entails an explicit or implicit contradiction. Only when inconsistent beliefs are apart from each other, can an agent hold them at the same time. In section 3, in order to defend the necessity of the theory of partitioned mind, I indicated that what is essential to a case of irrationality is the state of being inner inconsistent which accompanies the breakdown of the reason relation between two beliefs. Since a case of irrationality is a case of inner inconsistency, and an inner inconsistency within a person is possible when inconsistent beliefs are divided into different parts of the mind, the theory of partitioned mind is necessary to account for irrationality. However, if one refuses to accept that consistency is a fundamental norm of rationality and that inner inconsistency is a manifestation of irrationality, the theory of partitioned mind will be useless.

This section examines the appropriateness of Davidson's characterization of the concepts of rationality and irrationality. What follows begins with giving reasons to explain why Davidson takes inner consistency as a standard of rationality. I think that Davidson's reasons would be

relevant to the fact that practical reasonings or other kinds of non-deductive reasonings in question are constituted by propositional attitudes. Consistency is attributable to a set of propositions, and so the reasons I state will focus on the relations between rationality, consistency, and propositional attitudes. Then, I will discuss several lines of argument that could threaten the view that inner consistency is a standard of rationality. First, propositional attitudes as reasons are necessary for the exhibition of rationality or irrationality. But some argue that rationality is possible without reasons. If this is the case, then consistency will have nothing to do with rationality. Second, some agree that rationality requires reasons, but deny that the lack of consistency can deprive a person of his title of rational agent. In this sense, consistency is not a criterion of rationality. Third, the inner consistency that is associated with rationality by Davidson is an idea of synchronic consistency. In a piece of practical reasoning, it is attributable to an agent when his all-things-considered judgment, the principle of continence, and the unconditional judgment detached from a prima facie judgment are all active. Some philosophers doubt the rational authority of all-things-considered judgments and the normativity of the principle of continence. They tend to take diachronic consistency as the standard of rationality and the principle of continence as an evaluative tool. I shall provide reasons against these lines of arguments.

Davidson's reason for taking consistency between beliefs as a standard of being rational could be relevant to his regarding cases of irrationality, such as weakness of the will, probabilistic akrasia, and self-perception, as failures in reasoning which consists of propositional attitudes. As Sebastian Gardner remarks, '... these forms of ordinary irrationality, the forms of irrationality- those recognized in ordinary psychology- are *propositionally transparent*, by which it is meant that they are *constituted and defined by a particular structure of propositional*

attitudes'.¹⁷ Consistency is a property that can be attributed to a set of propositions. In a reasoning process, an agent entertains various propositions to reach a conclusion. Thus, it is natural to think about whether the agent reasons consistently or not. When the agent follows principles of rationality to reason, beliefs involved in his reasoning process are consistent. All the beliefs are taken to be true at the same time, and the agent commits himself to no explicit or implicit contradiction. Since consistency occurs when an agent follows principles of rationality to reason, consistency can be a standard of rationality.

What makes inner consistency a fundamental norm of rationality could also be explained in terms of the nature of deductive and non-deductive arguments. For an argument that is construed according to rules of deductive logic, the truth of its premises guarantees the truth of its conclusion. The argument is deductively valid and logically consistent. In contrast, an argument is invalid if it is possible for the premises to be true and the conclusion false. When an argument has true premises but a false conclusion, the argument is inconsistent since propositions that constitute the argument cannot be true at the same time. Although usually we do not talk about whether or not a non-deductive argument is valid when validity is considered as a function of arguments with deductive forms, we can talk about whether or not a non-deductive argument is good or strong. A non-deductive argument is good or strong when, given all true premises, its conclusion is likely to be true. Logical consistency can be attributed to a good or strong non-deductive argument whose premises and conclusion are true at the same time. I believe that it is natural to associate rationality with the construction of deductively valid arguments and good or strong non-deductive arguments. Practical reasoning and inductive

¹⁷ Sebastian Gardner, 'Ordinary Irrationality' (1993), p. 16.

reasoning are non-deductive reasoning. Hence, in non-deductive reasoning, he who reasons consistently is a rational agent. Otherwise, he is an irrational agent.

Accordingly, rationality and irrationality presuppose the existence of propositional attitudes. Because an agent exhibits rationality or irrationality when he reasons, and his reasoning consists of propositional attitudes, propositional attitudes are required for us to judge whether an agent who entertains them in a process of reasoning is rational or not. In a practical reasoning, these propositional attitudes constitute the agent's reasons for or against a course of action that he considers. Whether the agent is rational or not is determined by whether he reasons consistently, that is, by whether he properly responds to reasons in the reasoning procedure.

However, some argue that rationality does not require reasons, which challenges taking inner consistency as a standard of rationality. Judith Baker claims that many actions which are not understood as acting for reasons are rational, intelligible, and open to rational criticism. Baker denies the popular way which explains an action by reconstructing the agent's thoughts before taking the action as a reason-giving argument, but adopts the view that in most cases an agent's thoughts before acting are inclinations to act in certain ways which are not reasons. According to Baker, reasons are evaluations of ends or objects of actions, but inclinations are propensions. A propension is 'a more, or relatively, primitive capacity that has survived in humans, and that continues to guide action.'¹⁸ Propensions such as doing what one believes to be enjoyable, pleasant, relaxing, avoiding what is unpleasant or uncomfortable, acting out of curiosity, pursuing something intriguing, and going for what has some kind of appeal are

¹⁸ Judith Baker, 'Rationality without Reasons' (2008), p. 773.

dispositions that are good to have.¹⁹ Yet, Baker argues, ‘neither going for what is appealing nor being disposed to turn away from what one finds ‘scary’ is automatically to be reconstructed in terms of acting for what one finds valuable or acting for reasons. In some cases, however, the action may be rational, in that it is positively OK, but no more than that.’²⁰ For example, ‘you turn back from a mountain path that just feels scary. Even acting on a whim may be rational without there being reasons for doing so.’²¹

It seems that Christine Tappolet who claims that the content of an emotion is non-conceptual also disconnects reasons from rationality. Tappolet believes that an incontinent action based on emotion can actually be more rational than actions which conform to one’s better judgment. She claims that ‘... far from being only disruptive, emotions can also help us to behave more adequately than if we only trusted our deliberative faculty. My belief that there is no danger and that I thus ought to behave in a certain way might be simply wrong, so that it would be thanks to my fear that I am able to escape the danger that threatens me.’²²

If reasons are not necessary for rationality or irrationality, inner consistency will no longer be a standard of rationality. But I find Baker and Tappolet’s taking inclinations and emotions to be non-cognitive mental states problematic. Baker seems to hold inconsistent claims about inclinations. She regards reasons as evaluations, and denies that inclinations are reasons for they are not evaluations. However, she also claims that inclinations are propensions such as doing what one believes to be pleasant or relaxing, avoiding what is unpleasant or uncomfortable, and so on. It seems to me that what Baker enumerates as inclinations are evaluative judgments.

¹⁹ Ibid., 774.

²⁰ Ibid., 765.

²¹ Ibid., 764.

²² Christine Tappolet, ‘Emotions and Akratic Action’ (2003), p. 115.

They are positive or negative attitudes that we have toward objects, and can roughly be seen as judgments about what are good or bad. Hence, when we act on our inclinations, we act on what we judge to be good and avoid what we judge to be bad. In this sense, inclinations are evaluations which can be reasons for actions. Furthermore, even if Baker insists that inclinations are not evaluations, inclinations are cognitive states that can be reasons for actions. Believing what is pleasant, being curious about something, and other examples of inclination that Baker gives, when thought of by an agent, are expressed by propositions. As cognitive states, they can be reasons. If the inclination that motivates an agent is a belief, we know that there is a paired desire which together moves the agent to act. If an agent acts because he is curious about something, we know that the desire to satisfy his curiosity is his reason for acting.

Similarly, because emotions are not simply feelings but are propositional attitudes, when an emotion plays a role in the causal history of an action, it is a part of the rational pattern in which the action occurs and appears intelligible to us. Hence, acting from an emotion is acting for a reason. Suppose an agent accelerates his walking because of fear. His action makes sense to us when we know that he is scared because he judges that, say, walking alone in the dark is dangerous and he believes that he is in that situation. The desire to avoid danger and the belief that walking fast would avoid danger can explain why the agent walks faster. An agent does not have an emotion without having a reason for it. Only when emotions are cognitive states and logically connect with other mental states, can actions caused by emotions be intentional actions and be intelligible to us. Since acting for inclinations or emotions are acting for reasons, consistency is attributable to an agent who acts for an inclination or emotion.

Davidson claims that ‘All (objective) irrationality is a matter of inner inconsistency.’²³ Some agree with Davidson that an incontinent agent holds inconsistent practical judgments, but reject the claim that the connection between akrasia and inner inconsistency can deprive a person of his title of rational agent. Justin Gosling, for example, thinks that such a claim, that those whose mental states are inconsistent are irrational, is too strong. Instead, he takes rational agents to be those who are able to articulate reasons of their actions, and irrational agents to be those who have no idea about what they are up to.²⁴ Other philosophers, such as David Wiggins and Niko Kolodny,²⁵ also have different views about what rationality and irrationality are.

Of course there is more than one way to elaborate the concepts of rationality and irrationality, but I believe that Davidson’s taking consistency as a standard of rationality points out a necessary condition of an agent’s being rational. That is, for an agent who has propositional attitudes and is capable of thinking and reasoning, logical consistency is a basic requirement which merely asks the agent not to hold contradictory propositional attitudes. Hence, to those who hold different opinions about the nature of rationality, as long as they agree that rationality presupposes propositional attitudes, they should admit that inner consistency is associated with rationality. Inner consistency, in this sense, is a minimum requirement for the performance of rationality. Although Gosling seems to hold a different opinion about rationality, I doubt that he can sincerely reject the association between inner consistency and rationality. Gosling takes rational agents to be those who are able to articulate reasons of their actions. Since actions are based on reasons whose verbal expressions are propositions, rationality can be attributed to an

²³ Davidson, ‘Incoherence and Irrationality’, 194.

²⁴ Justin Gosling, ‘Akrasia and Irrationality’ (1990), pp. 122-123.

²⁵ David Wiggins, ‘Weakness of Will Commensurability, and the Objects of Deliberation and Desire’ (1979), p. 257. Niko Kolodny, ‘Why Be Rational?’ (2005), p. 554.

agent who reasons consistently. Moreover, it is difficult to take articulating reasons for actions as the standard of rationality. To articulate the reason of an action is to express the reason by an utterance which can be understood through interpretation. If an agent cannot articulate the reason of his action, his action is not intelligible to an interpreter and even to himself. Defining irrationality as incapability of giving the reason of an action renders irrational action a mysterious object beyond our ability to comprehend.

Consistency is a property of a set of propositions that are true at the same time, and the inner consistency that is associated with rationality is an idea of synchronic consistency. The concept of irrationality, in contrast, is an idea of synchronic inconsistency. In a process of practical reasoning, an agent's all-thing-considered judgment, the principle of continence, and his all-out, unconditional judgment are all active. We judge whether or not the agent is rational according to whether these beliefs are true at the same time, that is, the conjunction of these beliefs yields no contradiction. An all-things-considered judgment is a judgment formed on the basis of all relevant reasons that the agent has. The principle of continence demands that agent perform the action judged best on the basis of all available relevant reasons. Hence, when an agent detaches an unconditional judgment from his all-thing-considered judgment, his beliefs are consistent, and he is rational. An agent who fails to do so is irrational for he holds inconsistent beliefs.

Although Davidson takes an agent's acting in accordance with his all-things-considered judgment as a manifestation of rationality, some argue against the rational authority Davidson assigns to all-things-considered judgments. Opponents such as Nomy Arpaly even think that it is not always irrational to act against one's all-things-considered judgment. According to Arpaly,

setting the coherence of mental states as a core of rationality does not show that acting against one's all-things-considered judgment is irrational. Instead, it supports the idea that a person may be more rational when he fails to form an unconditional judgment consistent with his all-things-considered judgment. For the failure may render a larger number of his desires, the overall desire set, more coherent.²⁶ Arpaly rejects treating all-things-considered judgments as something special. An all-things-considered judgment is just another belief. When something conflicts with an all-things-considered judgment, the conflict is just like other ordinary inconsistency between other mental events or states. When discussing beliefs, all beliefs an agent actually has are equally important.²⁷ Annette Baier also asks whether an agent should trust his all-things-considered judgment as a guide to action when it is at odds with what he most wants to do. She says that, 'I do not find it obvious that our "best judgment" is our best guide to action'.²⁸

Arpaly seems to be more inclined to treat rationality as diachronic consistency, and I believe that it is important to maintain the diachronic consistency between an agent's propositional attitudes. An agent who is prudent pursues what he values the most in the long term. A decision he makes now would influence what is going to happen to him in the future. Hence, the consistency of his overall desire set is significant. But we can ask several questions about taking diachronic consistency as the standard of rationality. For an agent who does not know what his ultimate goal is or changes his mind frequently, it seems difficult for us to judge whether his overall desire set is consistent. If an agent's ultimate goal is not the basis according to which we evaluate the consistency of a set of desires, I wonder what would be more proper to be the standard. In addition, unlike synchronic consistency that is the consistency between

²⁶ Nomy Arpaly, 'On Acting Rationally against One's Best Judgment' (2000), p. 496.

²⁷ Ibid., p. 512.

²⁸ Annette Baier, 'Rhyme and Reason: Reflections on Davidson's Version of Having Reasons' (1985), p. 122.

propositional attitudes active at a moment, diachronic consistency is the consistency between propositional attitudes active in a period of time. Then, there must be a point at which we can go back to examine whether an agent's mental states are consistent and whether he acts rationally. But when is that point? It seems not unproblematic to take diachronic consistency as the standard of rationality.

Arpaly and Baier deny the rational authority of all-things-considered judgments, which should be astonishing to those who believe that there are logical relations in propositional attitudes. Propositional attitudes are components of thought. The logical relations, the reason relations between propositional attitudes, influence the way we think and reason. For instance, an agent who knows that more than 50% of voters in Scotland voted against Scottish independence in September 2014 should know that Scotland is not going to be an independent country according to this ballot. If the agent instead holds the belief that Scotland thereby becomes an independent country, we would doubt that he is aware of the mistake in his thought or he knows what voting is. An agent who sees something at a far distance but cannot discern its shape may temporarily take it to be a circle or a square. He does not take it to be both a circle and a square for this is not conceivable to him. If two propositional attitudes are randomly associated with each other, such as believing that it is going to rain and thinking that it is desirable to eat some broccoli tonight, an agent can hold them at the same time when he takes them to be true. The logical or reason relations between propositional attitudes have force guiding an agent's thinking and reasoning. It is rational for an agent to think and reason logically. Because an all-things-considered judgment is supported by an agent's best reason, it is supposed to have the strongest rational force. The rational authority of all-things-considered judgment is not arbitrarily assigned but comes from the nature of propositional attitudes.

Lastly, I would like to consider a problem about the normative force of the principle of continence, which Davidson requires being active in a process of practical reasoning. Principles of rationality are assumed as constitutive elements of thinking creatures by Davidson. The assumption of principles of rationality represents the rational aspect of thinking creatures. Those who hold principles of rationality but disobey them are irrational. For these claims to be true, principles of rationality must be normative. Principles of rationality are normative in the sense that they are principles 'saying something about what one ought, or has reason, to do or to refrain from doing in given circumstances.'²⁹ They guide agents to act or to think rationally, and they serve as standards to judge whether an agent is rational or not. If this assumption is shown to be false, following principles of rationality or not is irrelevant to rationality, an agent's disobeying principles of rationality does not lead to an inner inconsistency, and the partitioning mind theory becomes unnecessary.

Are principles of rationality normative? Justin Gosling does not think so. He takes following the principle of continence as something like following someone's announced ordering of preference since it is someone's announced ordering that determines that reason lies on that side or that that is the better judgment. The idea of the principle of continence then loses its charm. It is odd to declare that it is always more rational for someone to abide by the initially preferred course of action.³⁰

Niko Kolodny argues that principles of rationality are not normative but evaluative. There are no reasons to be rational. Kolodny does not deny that thinking creatures are subject to principles of rationality, but he thinks that principles of rationality only have seeming normative

²⁹ Jason Bridges, 'Rationality, Normativity, and Transparency' (2009), p. 353.

³⁰ Justin Gosling, 'What is the Problem?' (1990), p. 112.

force. The pressure that an agent feels when he obeys principles of rationality comes from the reasons he already has for a belief he may accept or for an action he may take. In other words, when we tell someone that it is rational for him to do x , we are not appealing to rational requirements to persuade him to do x . We are simply drawing his attention to the reasons he already has for x . Principles of rationality are standards of assessment or evaluation for an evaluator to judge, from a third-person point of view, whether an agent meets some standard.³¹ Davidson takes principles of rationality as constitutive elements for thinking creatures, but Kolodny claims that ‘it rarely threatens one’s survival as a believer or agent to violate a rational requirement in any particular case.’³²

I do not think that these objections are convincing enough to deny the normativity of principles of rationality. Even if following the principle of continence, as Gosling claims, is just like following someone’s announced ordering of preference, this does not make the principle of continence charmless. Instead, an agent’s following his preference to act is a manifestation of autonomy. The agent decides what is worth pursuing and pursues what he thinks is valuable. Given the self-goal choice assumption that ‘the rational person maximizes the satisfaction of her goals irrespective of others’ goals’³³, an agent is rational when he follows his preference ordering to act. Davidson does not specifically indicate what are included in the considerations on the basis of which an agent forms his all things-considered judgment. Since Davidson claims that an agent must assume that other thinking creatures are as rational as he is so that he can interpret and understand others’ speech acts, the agent must consider not only himself but also others’

³¹ Niko Kolodny, ‘Why Be Rational?’, pp. 542-560.

³² *Ibid.*, p. 544.

³³ Ann Cudd, ‘Commitment as Motivation: Amartya Sen’s Theory of Agency and the Explanation of Behavior’ (2014), p. 41.

thoughts when he reasons to act. Hence, the considerations involved in the all-things-considered judgment would include the agent's own goals and others' goals, and the different types of action in the preference ordering would include the agent's own goals and others' goal with which the agent shares. When an agent follows the principle of continence to reason, he detaches an unconditional conclusion from his all-things-considered judgment, and acts according to what he ranks first in his preference ordering. His action maximizes the satisfaction of his goals and expresses the autonomy-agency.³⁴ In this sense, the normativity of the principle of continence and someone's announced ordering of preference is grounded on the maximization framework. The rational force of the principle of continence can not be denied unless one denies the self-goal choice assumption.

As to Kolodny's argument, I agree with him that principles of rationality are evaluative in that they are taken as standards to evaluate whether an agent is rational or not, but I do not think that principles of rationality have no normative force. Kolodny is right to claim that an agent reaches a conclusion because of the reasons he has for the conclusion. Principles of rationality are not extra premises of arguments. The reasons an agent has in reasoning have the rational force to cause a corresponding conclusion. However, we know that competing reasons an agent has in his reasoning could cause different arguments and conclusions. If principles of rationality

³⁴ Ann Cudd (2014) believes that besides actions performed in light of one's own goals, there are actions motivated by goal-displacing commitments. Cudd proposes a broader theory of human agency which allows both types of action to be seen as expressions of agency, that is, the duality of autonomy-agency and identity-agency. The former covers cases in which an agent acts according to his preference ordering, and the later covers cases in which an agent acts not for his own reasons, such as internalized social norms or tacit commitments which constitute the agent's identities that are ascribed to him non-voluntarily and non-intentionally. The maximization framework and the intentional rational choice explanation accommodate actions that express autonomy-agency, but cannot accommodate actions which express identity-agency, that is, actions which are motivated by internalized social norms or tacit commitments that are never reflected or questioned and therefore do not represent actual, explicit, intentional commitments. On the contrary, the broader theory of human agency can accommodate both types of action. However, Cudd clarifies that once tacit commitments are recognized as guiding one's behaviors and embraced explicitly, they are no longer tacit, and behaviors explained by tacit commitments are now explained in the standard intentional rational choice manner.

have no normative force, and the formation of an argument simply depends on the rational force of reasons, doesn't it follow that an agent should always detach a conclusion from his best reason and always be rational? If principles of rationality have no normative force, and the rational force of reasons is not the only factor which determines whether or not an agent is going to endorse a reason, is it just lucky when an agent forms a conclusion in accordance with principles of rationality? It seems that the rational force of reasons alone is not enough to explain rationality and irrationality. On the contrary, when we admit the normative force of principles of rationality, principles of rationality as second-order principles that can be violated makes it possible for an agent to be irrational. We also know that when an agent reasons in accordance with his best reason, this is not an accident but a consequence of following principles of rationality and a performance of rationality.

In fact, principles of rationality are beliefs which have inherent rational force. The principle of continence is a belief that I should act on the judgment supported by the best reasons. The principle of total evidence is a belief that I should reason in accordance with the judgment supported by the strongest reasons. Other principles of rationality are also beliefs which have propositional content. As propositional attitudes, principles of rationality are logically connected to an agent's other propositional attitudes. Although principles of rationality are not reasons considered in reasoning about what to do or what is almost likely to occur, because they are beliefs held by an agent, they have rational force in his reasoning. They are effective when an agent is influenced to adopt his best reasons among competing reasons available to him. If the rational power fails to exhibit in a reasoning process, principles of rationality will cause an agent to notice that he is irrational to reason in this way.

To sum up, through the above discussion, I have clarified the following points. First, an agent always acts on certain reasons to which consistency as a standard of rationality applies. Inclinations, emotions, and any mental states could be expressed by propositions when they enter an agent's thought. Consistency is a property which is attributable to an agent's reasons for acting which could include inclinations, emotions, and any mental states that he is conscious of. Second, inner consistency which requires that propositional attitudes which an agent has are true at the same time is a minimum requirement for rationality. Being consistent is not enough to form a valid deductive argument or a good or strong non-deductive argument, but without consistency no argument is acceptable for an explicit or implicit contradiction will be yielded. Third, the inner consistency associated with rationality is synchronic consistency. In a practical reasoning, for instance, it requires that an agent's all-things-considered judgment, the principle of rationality, and the unconditional judgment to be present at the same time. Neither the rational authority of all-things-considered judgments nor the normative force of the principle of continence is arbitrarily assigned. As propositional attitudes, they have intrinsic rational force to cause corresponding propositional attitudes.

5. CONCLUSION

Davidson proposes the theory of partitioned mind to explain the cause of incontinence. With regard to the irrational aspect of an incontinent action, the strong desire causing an agent to act is the cause of the agent's being incontinent. It is a mental cause that is not a reason for the mental state it causes. When it causes an agent to ignore the principle of continence and his all-things-considered judgment, there is a breakdown of the reason relation between the agent's propositional attitudes, an inconsistency within the agent. Under the rational assumption that no

propositional attitude can exist independently from other propositional attitudes and have no logical relations with them, the theory of partitioned mind not only explains the possibility of the existence of a mental cause which is not a reason for its effect, but also explains irrational phenomena which contain inner inconsistencies in general.

In this chapter, I have explored several questions concerning Davidson's theory of partitioned mind, questions concerning the explanatory force of the theory, the necessity of the theory, and the relations between consistency, rationality, and irrationality. I have clarified that what the theory of partitioned mind aims to explain is not how a desire supported by a relatively weaker reason gets superior motivational force in an incontinent action but rather the existence of such a desire, a mental cause that is not a reason for its effect, and the inner inconsistency caused by it. The theory of partitioned mind gives an explanation which rules out the situation in which inner inconsistency is possible, that is, the situation in which inconsistent beliefs are put together and form a contradiction. When inconsistent beliefs are separated from each other into different parts of the mind, no explicit or implicit contradiction will be yielded from the conjunction of the inconsistent beliefs, and an agent can hold them at the same time. An agent is justified in holding inconsistent or even contradictory beliefs as long as he has reasons for each of the beliefs.

The theory of partitioned mind is necessary because only when inconsistent beliefs are separated into different parts of the mind, can an agent hold them at the same time. The necessity of the theory of partitioned mind is associated with the nature of irrationality. Although the phenomenon of irrationality is caused by a mental cause that is not a reason for its effect, the existence of such a mental cause is not enough to form a case of irrationality. For two mental

states that are arbitrarily connected in a causal relation could be true at the same time, break no reason relation, and cause no inconsistency. The destroyed logical relation between two propositional attitudes, the inner inconsistency, however, is sufficient for irrationality.³⁵ In a case of irrationality, an agent draws a false conclusion in his reasoning. Propositional attitudes involved in his reasoning are supposed to have certain logical relations and rational force. When a mental state causes another which cannot be rationalized by it, the given rational force of propositional attitudes is undermined. There is a breakdown of the reason relation between propositional attitudes, and the agent is in a state of inner inconsistency. The theory of partitioned mind is required to explain a case of irrationality which is a case of inner inconsistency.

The concepts of reason, consistency, inconsistency, rationality, and irrationality are interdependent for Davidson takes inner consistency to be a fundamental norm of rationality, and inner inconsistency a manifestation of irrationality. An agent's performing rationally or irrationally presumes the existence of reasons. Because all mental states that enter an agent's thought can be expressed by propositions, when they move an agent to act, they are the agent's reasons for acting. Consistency is a property that is attributable to a set of propositions. Reasons involved in a practical reasoning are propositional attitudes which can be either consistent or inconsistent. When they are all true, there is a logical consistency in the reasoning. If any propositional attitude fails to exhibit its rational force properly and causes a wrong propositional attitude, the reasoning lacks logical consistency. I believe that it is natural to accept that an agent

³⁵ Davidson would try to avoid giving a sufficient condition for the existence of a kind of mental state because doing so would violate his anomalism thesis that there are no psychophysical or psychological laws which connect a kind of mental state to another kind of mental or physical state. However, taking the breakdown of the reason relation between two propositional attitudes as a sufficient condition for irrationality is not against Davidson's anomalism thesis. This is because this claim does not give a necessary connection between two kinds of mental state. It just says that whenever there is a breakdown of the reason relation between two propositional attitudes, whatever they are, there is a case of irrationality.

who reasons consistently is rational, and that he who fails to do so is irrational.

V. CONCLUSION

Donald Davidson's analysis of the phenomenon of weakness of will is an investigation of the possibility of incontinent action and the cause of an agent's being incontinent. Davidson describes incontinence as a failure in reasoning. An incontinent agent holds competing *prima facie* judgments and fails to detach an all-out, unconditional judgment from his all-things-considered judgment. The incontinent agent considers (A) an all-things-considered judgment that *pf*(*a* is better than *b*, *e*) and a reason *e*, and (B) a judgment that *pf*(*b* is better than *a*, *r*) and a reason *r*. His reasoning stops at (A) on the one hand, and he detaches an all-out, unconditional judgment that *b* is better than *a* from (B) and does *b* on the other hand. In this analysis, incontinence is possible because the incontinent agent does not hold a contradiction. And this description is compatible with Davidson's causal theory of action and the assumption that an agent always does what he judges to be better because the all-out, unconditional judgment causing the incontinent agent to act is a value judgment.

Davidson gives the theory of partitioned mind to explain the cause of incontinence, a mental state which bears a non-logical causal relation with its effect, and other cases of irrationality which are cases of inner inconsistencies. Davidson identifies desire causing an incontinent agent to act as the cause of his being incontinent. The desire causes the agent to ignore his principle of continence but does not rationalize his doing so. It is a mental cause which is not a reason for its effect. When the agent ignores the principle of continence and forms an all-out, unconditional judgment not in accordance with his all-things-considered judgment, the agent holds an inconsistent set of propositional attitudes. He is in a state of inner inconsistency. Given the assumption that no propositional attitude can exist without causally and logically connecting

to other propositional attitudes, a mental state can cause another to which it bears no rational relation and the state of inner inconsistency is possible when the propositional attitudes that are inconsistent are divided into different parts of the mind. According to the theory of partitioned mind, the mind is to be regarded as having two or more semi-autonomous structures. A propositional attitude in a subdivision can be explained by other propositional attitudes in the same subdivision that are causally and logically connected with it. Two propositional attitudes that are inconsistent are apart from each other and respectively supported by a set of interlocking propositional attitudes. But between the parts they belong to there is only a non-logical causal relation. In a divided mind, a mental state is causally and logically connected with other mental states in the subdivision where it is located, and can cause another to which it bears no rational relation when the mental state and its effect are divided into different subdivisions. Irrationality as an inner inconsistency is possible when inconsistent propositional attitudes are separated from each other.

In this dissertation, I have defended Davidson's analysis of the phenomenon of weakness of will. I discussed several lines of argument with respect to Davidson's describing incontinence as a failure in reasoning and taking irrationality as an inner inconsistency. These lines of argument and my responses are summarized as follows.

(1) Some deny that Davidson's account of weakness of will is successful because they deny the non-implication relation between an all-things-considered judgment and a consistent all-out, unconditional judgment. They claim that the principle of continence is analogous to modus ponens and should rationally compel an agent to reason consistently. An agent who refuses to infer from an all-things-considered judgment that $pf(a \text{ is better than } b, e)$ to an all-out,

unconditional judgment that *a* is better than *b* is like an agent who refuses to infer from ‘If *p*, then *q*’ and ‘*p*’ to ‘*q*’. They argue that because an all-things-considered judgment that *pf*(*a* is better than *b*, *e*) entails a consistent all-out judgment that *a* is better than *b*, an incontinent agent who holds an all-things-considered judgment that *pf*(*a* is better than *b*, *e*) and an inconsistent all-out, unconditional judgment that *b* is better than *a* thereby holds contradictory judgments.

Incontinence as a failure in reasoning is impossible since there is an implication between the two kinds of evaluative judgment.

I defended the non-implication relation between two kinds of evaluative judgment by clarifying the role the principle of continence plays in practical reasoning. Davidson replies to those who think that there is an analogy between the principle of continence and modus ponens, and I attempted to give an interpretation which makes Davidson’s responses consistent. I suggested understanding ‘validity’ in different ways. When the validity of an argument is seen as a function of its logical form, the principle of continence is not analogous to modus ponens because formal validity is irrelevant to practical arguments which are non-deductive in form. However, validity is a property which can be attributed to arguments other than deductive arguments when ‘validity’ is understood as a property of arguments which are truth-preserving. When ‘validity’ is so understood, a practical argument which follows the principle of continence is established as valid in the sense that it is impossible for its conclusion to be false when all of its premises are true.

(2) Given Davidson’s causal theory of action, reasons are causes, and the strongest reason is supposed to have the strongest causal power. In an incontinent action, the agent’s all-things-considered judgment fails to cause an all-out, unconditional judgment in accordance with it.

Instead, another competing prima facie judgment with a relatively weaker reason causes a corresponding all-out, unconditional judgment which in turn causes the agent to act. Some criticize Davidson's theory of partitioned mind because it does not explain why a prima facie judgment with a relatively weaker reason gets superior motivational force in an incontinent agent. They think that merely separating inconsistent beliefs into different parts of the mind cannot explain the shift in the expected motivational forces of desires.

Although Davidson's theory of partitioned mind is not proposed to explain the superior motivational force of the prima facie judgment which causes an all-out, unconditional judgment in an incontinent action, why a prima facie judgment supported by a relatively weaker reason gets stronger motivational force in an incontinent action is indeed a problem. Some philosophers try to identify a specific mental state, such as the lack of self-control or an emotion, to be the mental cause which is responsible for an incontinent agent's reasoning error. However, I have argued that neither the lack of self-control nor the emotion involved in an incontinent action can be the cause of incontinence which covers all cases of incontinent action. In fact, no mental state of a specific kind can be the cause of incontinence which explains all cases of incontinent action. Davidson needs not to identify a mental state of a specific kind to explain the superior motivational force of the prima facie judgment, for doing so is inconsistent with his anomalism that there are no psychophysical or psychological laws governing the connection between different types of mental state. But this does not mean that there is no mental cause which can explain why an incontinent agent's all-things-considered judgment fails to cause an all-out, unconditional judgment in accordance with it. Instead, an emotion, personal character, custom, or physical impulse could interfere with his reasoning and cause him to draw a false conclusion and be incontinent.

(3) Davidson believes that there is a parallel between a case in which an agent fails to follow the principle of continence to detach a conclusion about what to do from his all-things-considered judgment and a case in which an agent fails to follow the principle of total evidence to detach a conclusion about what is almost certainly to occur from his probabilistic judgment supported by the strongest evidence. However, this parallel is rejected because probabilistic akrasia is thought to be impossible. Beliefs are governed by truth. When an agent has stronger evidence to believe p , he will not believe $\neg p$. No one can harbor conflicting probabilistic judgments. There is no such thing as probabilistic akrasia.

I agreed that beliefs are governed by truth, but this is not a reason against the possibility of probabilistic akrasia. As long as the probabilistic judgments that conflict with each other are supported by different evidence, they can be true to an agent to different degrees. Conflicting probabilistic judgments held true by an agent can exist at the same time. In a normal situation, a probabilistic judgment supported by the strongest evidence should be taken to be more likely to be true than other competing probabilistic judgments. But an irrational agent fails to think in this way. To the cause of probabilistic akrasia, I suggested an answer similar to that for the cause of incontinence. That is, no mental state of a specific kind can be the mental cause which explains all cases of probabilistic akrasia. However, an intuition, prejudice, or stubborn beliefs could intervene in his reasoning and cause him to draw a false conclusion and be irrational.

(4) Some argue that Davidson's theory of partitioned mind has no explanatory force because although it identifies the strong desire causing an incontinent agent to act as the cause of his being incontinent, it does not explain why the desire with an insufficient reason gets superior motivational force in the incontinent agent's reasoning. The theory of partitioned mind claims

that when an agent has inconsistent beliefs, they are divided into different parts of the mind, and each of them is supported by a set of interlocking propositional attitudes. Between them there is a non-logical causal relation. This seems able to cover all cases of irrationality, but such an achievement is obtained merely because the subdivision boundary is defined as a line which limits the normative force of a propositional attitude to the subdivision where it is located.

I have clarified that the theory of partitioned mind is not proposed to explain the superior motivational force of the desire causing an incontinent agent to act, but to explain the possibility of inner inconsistency. The explanatory force of the theory lies in the separation of inconsistent beliefs, rather than the subdivision boundary which is used as a conceptual aid. I argued that the theory of partitioned mind does give an account of the possibility of inner inconsistency because it rules out the situation in which inner inconsistency is possible, that is, the situation in which inconsistent beliefs are put together and form a contradiction. An agent could hold each of the inconsistent beliefs to be true when he considers them separately. What legitimates such a separation is the fact that the agent has different reasons for each of the inconsistent beliefs. Hence, as long as the agent has reasons for each of the beliefs that are inconsistent, he is not forced to give up one or the other.

(5) The theory of partitioned mind is regarded as unnecessary to explain the cause of irrationality, a mental state which bears no rational relation with another that it causes, because such a mental cause and its effect are intelligible to us when a coherent causal nexus in which they occur is offered. There is no need to divide them into different parts of the mind.

Reasons that an agent considers in his reasoning are supposed to link in certain ways according to the rational relations between them. In a case of incontinent action, the rational

relation assumed to exist between two propositional attitudes is destroyed. The agent who holds them at the same time is in a state of inner inconsistency. The theory of partitioned mind explains the compatibility between the state of inner inconsistency and the assumption that no propositional attitude can exist without causally and logically connecting to other propositional attitudes. This compatibility, in my view, cannot be explained by giving the causal nexus which shows how the two inconsistent propositional attitudes are generated. The causal history of two inconsistent propositional attitudes does not tell us how irrationality as a case of inner inconsistency is possible.

(6) Davidson takes inner consistency as a standard of rationality, and inner inconsistency as the manifestation of irrationality. In an intentional action, whether the agent is rational or not depends on whether his all-things-considered judgment, the principle of continence, and the all-out, unconditional judgment are consistent. An incontinent agent holds these propositional attitudes that can not be true at the same time when they are considered together. Davidson develops the theory of partitioned mind to explain the possibility of irrationality as a case of inner inconsistency. However, his position is threatened by the following claims. First, rationality is possible without reasons. Second, the lack of consistency cannot deprive a person of his title of rational agent. Third, all-things-considered judgment does not have the attributed rational authority, and the principle of continence has no normative force. If any of these claims is true, inner consistency is irrelevant to rationality, and the interdependent relations between reasons, consistency, inconsistency, rationality, and irrationality become illusionary.

My responses to these claims are as follows. First, any mental state that an agent is conscious of can be expressed by a proposition. When it causes an agent to take an intentional

action, it is a reason for the action. Hence, an agent always acts on certain reasons to which consistency as a standard of rationality applies. Second, consistency is not sufficient to form a valid or strong argument, but propositions consisting of a valid or strong argument must be consistent. If forming a valid or strong argument is a performance of rationality, consistency is a requirement for rationality. Third, the rational authority of an all-things-considered judgment and the rational force of the principle of continence come from the nature of propositional attitudes. As propositional attitudes, they have intrinsic rational force to cause corresponding propositional attitudes.

In this dissertation, my arguments mainly relied on these assumptions: (a) practical arguments are non-deductive arguments, (b) there are no laws governing the connections between different types of mental event, and (c) mental states or events an agent perceives are cognitive states which can be represented by propositional expressions. Assumptions (a) and (b) explain the non-implication relation between all-things-considered judgment and all-out, unconditional judgment and reject the necessity that an agent detaches an all-out, unconditional judgment from his all-things-considered judgment. This makes incontinence as a failure in reasoning possible. A mental state, such as an emotion, personal character, custom, or physical impulse could interfere with his reasoning and cause him to draw a false conclusion. Assumption (b) gives a reason for the multiplicity of the mental causes which could explain such a failure, and assumption (c) excludes the situation that the mental state is a non-cognitive state and the action caused by it thereby becomes a non-intentional action. Assumption (c) is also essential to interpreting irrationality as an inner inconsistency. When mental events or states that an agent perceives are cognitive states, mental events or states involved in a course of action can be

expressed as propositions and be consistent or inconsistent with other propositional attitudes involved in the agent's reasoning for the action.

For those who believe that the form of practical arguments is not non-deductive and believe that there are laws or rules governing the connections between different types of thoughts and actions, the explanation of the possibility and the cause of incontinence as a failure in reasoning suggested in this dissertation would not be acceptable. For example, Paul Churchland claims that action-explanations are of the deductive-nomological mold, and he gives conditions the satisfaction of which necessarily causes an agent to act.¹ But if the relation between a reason and the action it causes, as Churchland claims, is deductive and nomological, then no one would ever make a mistake in practical reasoning, and the problem whether an agent could fail to detach an all-out judgment from his all-things-considered judgment will not arise.

Claiming that any mental state that is perceived and recognized by an agent is a cognitive state and that any mental state intervening in an agent's reason could be the cause of his failure in reasoning could cause someone to mistakenly regard the mental state as a reason for his acting incontinently. The mental state, when it is considered by an agent and thereby represented to the agent as a thought, becomes a reason favoring the incontinent action. Because the mental state, as the cause of the agent's incontinence, also rationalizes his incontinent action, it could be taken as a reason for his acting incontinently. But it is not. The agent has no reason for being incontinent. The mental state probably makes the agent appear to be more rational, but as long as the agent still believes that there is a better option available to him when he acts, his action is an incontinent action.

¹ Paul Churchland, 'The Logical Character of Action-Explanations' (1970), pp. 221-222.

REFERENCES

- Anscombe, G. E. M. (1965) Thought and Action in Aristotle. *New Essays on Plato and Aristotle*, edited by Renford Bambrogh, 143-158. New York: The Humanity Press.
- Aristotle. *Nicomachean Ethics*, Book VII, Chapter 1-10, edited by Richard McKeon (2001). New York: The Modern Library.
- _____. *De Anima*, translated by Mark Shiffman (2011). Newburyport, MA: Focus Publishing.
- Arpaly, Nomy. (2000) On Acting Rationally against One's Best Judgment. *Ethics*, 110, 488-513.
- Baier, Annette. (1985) Rhyme and Reason: Reflections on Davidson's Version of Having Reasons. *Action and Events: Perspectives on the Philosophy of Donald Davidson*, edited by Ernest Lepore and Brian McLaughlin, 116-129. Oxford: Basil Blackwell.
- Baker, Judith. (2008) Rationality without Reasons. *Mind*, Vol. 117, Issue 468, 763-782.
- Beall, Jc & Restall, Greg. (2013) Logical Consequence. The Stanford Encyclopedia of Philosophy.
- Bricke, John. (2011) Hume and Davidson: Passion, Evaluation, and Truth. Paper presented at conference Hume and Modern Philosophy, University of Moscow, Russia.
- Bridge, Jason. (2009) Rationality, Normativity, and Transparency. *Mind*, Vol. 18, Issue 470, 353-367.
- Carnap, Rudolf. (1955) Statistical and Inductive Probability. 1-17.
- Charlton, William. (1988) Davidson and His Critics. *Weakness of Will*, 114-134. Oxford: Basil Blackwell.
- Churchland, Paul. (1970) The Logical Character of Action-Explanations. *The Philosophical Review*, Vol. 79, No. 2, 214-236.
- Cudd, Ann. (2014) Commitment as Motivation: Amartya Sen's Theory of Agency and the Explanation of Behavior. *Economics and Philosophy*, Vol. 30, 35-56.

- Davidson, Donald. (1963) *Actions, Reasons, and Causes*. Reprinted in *Essays on Actions and Events* (2001), 3-19. Oxford: Oxford University Press.
- _____. (1967) *Truth and Meaning*. Reprinted in *Inquiries into Truth and Interpretation* (2009), 17-42. Oxford: Oxford University Press.
- _____. (1969) *How is Weakness of the Will Possible?* Reprinted in *Essays on Actions and Events* (2001), 21-42. Oxford: Oxford University Press.
- _____. (1970) *Mental Events*. Reprinted in *Essays on Actions and Events* (2001), 207-225. Oxford: Oxford University Press.
- _____. (1973) *Radical Interpretation*. Reprinted in *Inquiries into Truth and Interpretation* (2009), 125-140. Oxford: Oxford University Press.
- _____. (1973) *The Material Mind*. Reprinted in *Essays on Actions and Events* (2001), 245-259. Oxford: Oxford University Press.
- _____. (1974) *Psychology as Philosophy*. Reprinted in *Essays on Actions and Events* (2001), 229-244. Oxford: Oxford University Press.
- _____. (1975) *Thought and Talk*. Reprinted in *Inquiries into Truth and Interpretation* (2009), 155-170. Oxford: Oxford University Press.
- _____. (1976) *Hume's Cognitive Theory of Pride*. Reprinted in *Essays on Actions and Events* (2001), 277-290. Oxford: Oxford University Press.
- _____. (1978) *Intending*. Reprinted in *Essays on Actions and Events* (2001), 83-102. Oxford: Oxford University Press.
- _____. (1980) *A Unified Theory of Thought, Meaning, and Action*. Reprinted in *Problems of Rationality* (2004), 151-166. Oxford: Oxford University Press.
- _____. (1982) *Paradoxes of Irrationality*. Reprinted in *Problems of Rationality* (2004), 169-187. Oxford: Oxford University Press.
- _____. (1984) *First Person Authority*. Reprinted in *Subjective, Intersubjective, Objective*, (2001), 3-14. Oxford: Oxford University Press.
- _____. (1985) *Incoherence and Irrationality*. Reprinted in *Problems of Rationality* (2004), 189-198. Oxford: Oxford University Press.
- _____. (1986) *Deception and Division*. Reprinted in *Problems of Rationality* (2004), 199-212. Oxford: Oxford University Press.
- _____. (1995) *The Problem of Objectivity*. Reprinted in *Problems of Rationality* (2004), 3-18. Oxford: Oxford University Press.
- _____. (1995) *The Objectivity of Values*. Reprinted in *Problems of Rationality* (2004), 39-52. Oxford: Oxford University Press.

- _____. (2005) The Content of the Concept of Truth. *Truth and Predication*, edited by Kevin Sharpe, 49-75. Harvard University Press.
- Elster, Jon. (1999) Davidson on Weakness of Will and Self-Deception. *The Philosophy of Donald Davidson*, edited by Lewis Edwin Hahn, 425-442. Peru: Open Court.
- Gardner, Sebastian. (1993) Ordinary Irrationality. *Irrationality and the Philosophy of Psychoanalysis*, 15-39. Cambridge University Press.
- Gosling, Justin. (1990) What is the Problem? *Weakness of the Will*, 97-118. London: Routledge.
- _____. (1990) Akrasia and Irrationality. *Weakness of the Will*, 119-131.
- Grice, Paul & Baker, Judith. (1987) Davidson on 'Weakness of Will'. *Essays on Davidson: Actions and Events*, edited by Bruce Vermazen and Merrill Hintikka, 27-50. Oxford University Press.
- Hare, R. M. (2001) Weakness of Will. *Encyclopedia of Ethics*, 2nd Edition, Vol. III, edited by L. Becker and C. Becker, 1789-1792. New York: Routledge.
- Hilpinen, Risto. (2001) Deontic Logic. *The Blackwell Guide to Philosophic Logic*, edited by Lou Goble, 159-182. Blackwell Publishers Ltd.
- Hurley, Susan. (1985) Conflict, "Akrasia", and Cognitivism. *Proceedings of the Aristotelian Society*, Vol. 86, 23-49.
- Irwin, Terence. (2008) Aristotle reads the *Protagoras*. *Weakness of Will from Plato to the Present*, edited by Tobias Hoffman, 22-41. Catholic University of America Press.
- Kim, Jaegwon. (2003) Philosophy of Mind and Psychology. *Donald Davidson*, edited by Kirk Ludwig, 113-136. Cambridge University Press.
- Kolondy, Niko. (2005) Why Be Rational? *Mind*, Vol. 114, Issue 455, 509-562.
- Lazar, Ariela. (1999) Akrasia and the Principle of Continence or What the Tortoise Would Say to Achilles. *The Philosophy of Donald Davidson*, edited by Lewis Edwin Hahn, 381-401. Peru: Open Court.

- Mele, Alfred. (1987) Akratic Action: Causes, Reasons, Intentions. In his *Irrationality*, 31-49. New York: Oxford University Press.
- _____. (1987) Strict Akratic Action: A Paradox of Irrationality. *Irrationality*, 75-95.
- _____. (2003) Philosophy of Action. *Donald Davidson*, edited by Kirk Ludwig, 64-84. Cambridge University Press.
- Peacocke, Christopher. (1987) Intention and Akrasia. *Essays on Davidson: Actions and Events*, edited by Bruce Vermazen and Merrill Hintikka, 51-74. Oxford University Press.
- Pears, David. (1984) Paradoxes and Systems. In his *Motivated Irrationality*, 67-106. Oxford: Oxford University Press.
- _____. (1985) Reply to Annette Baier: Rhyme and Reason. *Action and Events: Perspectives on the Philosophy of Donald Davidson*, edited by Ernest Lepore and Brian McLaughlin, 130-136. Oxford: Basil Blackwell.
- Plato. *Protagoras*, 348c- 362a. *Readings in Ancient Greek Philosophy: From Thales to Aristotle*, edited by S. Marc Cohen, Patricia Curd, and C. D. C. Reeve (2005).
- Price, A. W. (2006) Acrasia and Self-Control. Aristotle's *Nicomachean Ethics*, edited by Richard Kraut, 234-249. Blackwell Publishing Ltd.
- Rumfitt, Ian. (2011) What is Logic? *Truth, Reference and Realism*, edited by Zsolt Novak and Andras Simonyi, 125-175. Central European University Press.
- Russell, Bertrand. (1959) On Induction. *The Problems of Philosophy*. Oxford University Press.
- Santas, Gerasimos. (1966) Plato's Protagoras and Explanations of Weakness. *The Philosophical Review*, Vol. 75, No. 1, 3-33.
- Tappolet, Christine. (2003) Emotions and the Intelligibility of Akratic Action. *Weakness of Will and Practical Irrationality*, edited by Sarah Stroud and Christine Tappolet, 97-120. Oxford: Oxford University Press.
- Thalberg, Irving. (1985) Questions About Motivated Strength. *Action and Events: Perspectives on the Philosophy of Donald Davidson*, edited by Ernest Lepore and Brian McLaughlin, 89-103. Oxford: Basil Blackwell.

- Waismann, F. (1965) *The Causal Interpretation of Language. The Principles of Linguistic Philosophy*, edited by R. Harre. New York: St Martin's Press.
- Watson, Gary. (1977) Skepticism about Weakness of Will. *The Philosophical Review*, Vol. 86, No. 3, 316-339.
- Wiggins, David. (1979) Weakness of Will, Commensurability, and the Objects of Deliberation and Desire. *Proceedings of the Aristotelian Society*, Vol. 79, 251-277.
- Williams, John. (1981) Inconsistency and Contradiction. *Mind*, Vol. 90, No. 360, 600-602.
- Wittgenstein, Ludwig. (1967) *Zettel*, edited by G.E.M. Anscombe & G.H. von Wright. Oxford: Blackwell.
- _____. (2001) *Wittgenstein's Lectures: Cambridge, 1932-1935*, edited by Alice Ambrose. Prometheus Books.