

Davidson on Irrationality and Division

Miguel Amen¹,

Abstract: Donald Davidson in ‘Paradoxes of Irrationality’ (1982) claims that to understand irrationality one has to postulate a divided mind, absent which one could not make sense of the phenomenon. Here I want to defend his position against some objections advanced by John Heil in ‘Divided Minds’. Heil has two complaints against Davidson’s theory; first he seems to believe that when we cash out the metaphor of a divided mind to give an account of irrationality, the result is an implausible picture; secondly that there are other models to explain irrationality that do not rely on Division. The idea is that even if the concept of a divided mind could be sufficient to explain irrationality it is not necessary one, and in view of the cumbersome nature of Davidson’s explanation it would seem altogether superfluous. This is a serious attack, but to my mind entirely misguided. In this Paper I will show why. First by showing that Heil seems to develop an erroneous account of division and the function of partitioning. Secondly by showing that Heil’s model is not consistent with important doctrines that he seems to accept and are central to Davidson. In the end Heil’s counterexample is a failure at explanation.

Resumo: Em ‘Paradoxes of Irrationality’ Donald Davidson defende que para compreendermos a irracionalidade temos que postular uma mente dividida, sem o que não conseguiremos dar sentido ao fenómeno. Neste artigo quero defender a posição de Davidson contra algumas objecções avançadas por John Heil em ‘Divided Minds’. Heil tem duas grandes objecções à teoria davidsoniana: 1) ele parece acreditar que quando procuramos ‘converter’ a metáfora da mente dividida para termos uma concepção da irracionalidade, o resultado é um quadro implausível, 2) existem outros modelos que explicam a irracionalidade e que não se apoiam na divisão da mente. A ideia é que mesmo que o conceito de mente dividida seja suficiente para explicar a irracionalidade, ele não é necessário, e levando em consideração a natureza complicada da explicação de Davidson ele parece pura e simplesmente supérfluo. Este é um ataque sério, mas na minha opinião totalmente mal dirigido. Neste artigo procurarei mostrar o porquê. Primeiro mostrarei que Heil parece desenvolver uma descrição errônea da divisão e da função da partição. Em segundo lugar mostrarei que o modelo de Heil não é consistente com doutrinas importantes que ele parece aceitar e que são centrais para Davidson. Em última análise, o contraexemplo de Heil falha na tentativa de explicação.

¹ Membro e investigador do *Mind Language and Action Group* – MLAG – do Instituto de Filosofia da Universidade do Porto. Bolseiro de Doutoramento da FCT (Bolsa SFRH/BD/24582/2005).

Donald Davidson in 'Paradoxes of Irrationality' (1982) claims that to understand irrationality one has to postulate a divided mind, absent which one could not make sense of the phenomenon. Here I want to defend his position against some objections advanced by John Heil in 'Divided Minds'. Heil has two complaints against Davidson's theory; first that the conception of a divided mind "for all its apparent straightforwardness...when pushed, seems gratuitously complex" (p.580). He seems to believe that when we cash out the metaphor of a divided mind to give an account of irrationality, the result is an implausible picture; secondly that "there are already available to us simpler, far less cumbersome accounts of the phenomena, accounts that leave room for irrational thoughts and deeds within a framework broadly constrained by charity" (p.580). The idea is that even if the concept of a divided mind could be sufficient to explain irrationality it is not necessary one, and in view of the cumbersome nature of Davidson's explanation it would seem altogether superfluous. This is a serious attack, but to my mind entirely misguided. First Heil seems to develop the wrong model for our understanding of division, and the cumbersome nature already mentioned is thus a result of his conception and interpretation of Davidson, and not, in my view, a proper characterization of Davidson's work on irrationality and division. As we will see Davidson's account does not have the characteristic that makes Heil's reading of him implausible. In fact I think that Heil's account is not only a failure to provide the best model, but is inconsistent with important doctrines that he seems to accept and are central to Davidson. Not only do I think that Davidson's account provides a clear model to explain irrationality, it seems to me to provide a necessary one. I will not defend this latter claim in a completely satisfactory way, because I will settle with the weaker claim that Heil's counterexample is a failure at explanation.

Lets begin with the second objection, where it is claimed that "there are already available to us simpler, far less cumbersome accounts of irrationality". What is being asserted is that Davidson's claim that division is a necessary ingredient in certain accounts of irrationality is wrong, since there are available to us other ways of explaining it. In the following we will analyse a different model that Heil takes to be not only sufficient, but also, as we will see later on, a better model to explain irrationality. It will, I hope, be clear why I disagree.

Heil discusses a case of Akrasia, and says that we do not need the model of a partitioned mind to account for irrationality. To take his

example: Wayne acts akratically when he insults a student, because he goes against his better judgment that tells him not to act that way. It is worth quoting Heil in full at this point, he say that

In such cases it may simply be that a certain desire, here the desire to insult Wayne, enjoys motivational clout disproportionate to my assessment of it. In judging what I have most reason to do, then, I assign the desire a relatively low ranking. The desire in question turns out, however, to possess strength disproportionate to its standing and, as a result, I acquire an intention to insult Wayne on its basis and subsequently act on that intention. My action is irrational, not because I fail to act on my strongest desire, but because I act against my considered better judgement, a better judgement that assigns a diminished ranking to that desire. (Heil 89, p.581)

Here we are given a model that, according to Heil, makes sense of Wayne's behaviour. What is not clear is why does he think that he has given an account that makes *sense*² of irrationality. Remember however, that what is required in this moment of the dialectics is a different model that is sufficient to *explain* irrationality. Be that as it may, I think that in a way he does not even address the problem. The way that he seems to think that irrationality enters the picture clearly makes the problem invisible. Heil says that an action is irrational because it goes against his best judgement. However this is a description of akrasia and not a description of the source of irrationality. The irrationality enters the picture only because akratic acts lead the agent to hold on to contradictory judgements. However an action, or an intention to perform an action that goes against one's best judgement does not lead one into contradiction. The contradiction enters the picture because one acts against the principle of continence. Maybe I am going a bit too fast here, so it is convenient now to go into a brief exposition of Davidson's views on the source of the problem of irrationality and his conception of practical reason that is adjacent to it. First to see what needs to be explained and why, because Heil seems to miss it; secondly to see why it is the principle of continence and not acting against one's best judgement that leads to a contradiction.

Why does irrationality pose a problem? Because it leads to paradox. This is explained by Davidson this way

² This is not so much a question of explaining irrationality, but as we will see in a moment, of giving an account of its paradoxical nature.

“The underlying paradox of irrationality, from which no theory can entirely escape, is this: if we explain it too well, we turn it into a concealed form of rationality; while if we assign incoherence too glibly, we merely compromise our ability to diagnose irrationality by withdrawing the background of rationality needed to justify any diagnosis at all.” (Paradoxes of Irrationality, p.184)

Now, what needs to be clear is why by assuming others to be incoherence or given to too many inconsistencies we lose the need to explain irrationality.

Succinctly the problem arises because of the nature of propositional attitudes. The fact that “there is a rational element at its core.” This comes out clearly in the fact that mental states and action fit into a pattern of logical relations in a way that, for example, an intention can be explained, i.e. rationalized, by referring to a desire and a belief; a desire by referring to the content of other beliefs and values. These logical relations provide explanations, in the form of reasons that not only tell us why someone would do an action by showing us that it makes sense in the light of the contents of his reasons, but that those reasons cause those actions, desires, beliefs, intentions etc. According to Davidson when we interpret someone we try making sense of him by attributing to him mental states in a way that they fit into a rational pattern, endorsing the principle of charity. The problem of inner inconsistency is that it goes against this pattern. To attribute to someone, tout court, that he believes ‘p and not p’ is to reach a breakdown in the process of making sense of someone in the light of reason. Now, this should not be seen as meaning that upon deviance from a principle of rationality, a rational interpretation is immediately lost. But such stark deviance has holding to ‘p and not p’ is for Davidson a clear crossing of the shadow line into the domain of the incomprehensible³. And once there we lose track of the mental, and once we lose the sense of the mental, questions of rationality have no jurisdiction, and so it makes no sense to pose the irrationality question. In a way for Davidson, to explain irrationality is to explain how come someone that *prima facie* seems to hold to ‘p and not p’ in fact holds p and holds not p, but doesn’t put the two together. Davidson is explicit –

³ Compare this with the case of someone that does not make an obvious inference. Here we are much more prepared to accept his failure.

to interpret someone as holding to 'p and not p' is to make a mistake of interpretation.

Now getting back to akrasia and the source of irrationality. If Heil thinks that acting or forming an intention against one best judgment is the source of inconsistency then he would think that somehow, best judgment and action or better, best judgment and intention, present a contradiction. Now, since this, as we will see in a moment, is not Davidson's view, he would have to have a different analysis on matters of practical reasoning. However he does not defend one and seems to accept most of Davidson's views. Heil's explanation might explain, by giving a mechanism, how come one goes against one best judgment, but since that is not the required explanandum it doesn't advance his criticism. To see this we now turn to Davidson's views on practical reasoning, only a brief look, but it will be enough to make matters more clear.

As I said the contradiction consist in one going against the principle of continence, viz, that one should act according to one's best judgement. Davidson claims that irrationality does not enter when we hold reasons for and against an action. The point is that those reasons, or the conclusion of the practical syllogism that support conflicting actions or the formation of conflicting intentions, are not of the form of universal statements, say "all lies are wrong", since every action we perform has positive and negative things to be said for it. The conclusion of practical syllogisms present us with a conditional statement, and so one can have two conclusions, one of which supports a given action and the other that is against it, without leading or having the agent in a contradiction, because being *conditional* judgements, they support an action insofar as certain reasons are given. They are like the conclusions of a piece of statistical inductive reasoning, they are true in so far as the premises are held and cannot be separated from those premises. Basically, conditional judgements do not clash logically speaking. In the same way the akratic is not inconsistent because he is contradicting his best judgement. The point is that there isn't any contradiction there. An action or an intention to act being an unconditional judgement does not clash logically with a conditional judgement. Once again the conditional judgement, and in this particular case we are talking about our best judgement, cannot be separated from the reasons for it, and the unconditional judgement just say that something is good or desirable, period.

Since Davidson seems to think that an intention is an

unconditional judgement, it does not clash with one best judgement.

In the case of Akrasia *the inconsistency enters* when one considers one of the principles of rationality, the principle of continence. Since the principle is not conditional it *does* enter in contradiction with the akratic act!

So deeming some act akratic is to present one with a problem of interpretation. Someone, following Heil, would have to provide an explanation of how can the same mind that stands for such contradictory judgement or beliefs be explained without being said to hold an open contradiction like ‘p and not p’. Such propositional attitudes in a single, undivided mind, that are “present at once and in some sense in operation” would seem to require the attribution that ‘p and not p’, since not only would he be aware of the attitudes but they would stand in the same web of beliefs. But as it was said before, Davidson thinks that this would be a mistake of interpretation.

Now we reach a point in this work where some assessment of what we have been doing is useful. We have been analysing an example from John Heil’s? paper ‘Minds Divided’ that supposedly gives an explanation of an irrational action that does not presupposes the model of a divided mind. We have seen that this explanation is defective because it misunderstands the nature of practical reasoning and thus the source of irrationality in an akratic act. At same time we have seen that Davidson’s conception of practical reasoning, as exposed here, can accommodate, logically speaking, akratic acts and other irrational items. That is, his theory of practical reasoning is compatible with the existence of weakness of will and other irrational items. But this possibility, which Davidson takes very seriously⁴, brings a problem of interpretation, with which we will now deal, by presenting Davidson view on division, as giving a factual account of how irrationality is possible and how this account can save our interpretative efforts of the akratic agent.

It seems to me that any account of irrationality that presupposes a picture of practical reasoning with the preceding features needs to postulate a divided mind if is to make sense of inner inconsistency. Davidson claims that the following three theses, that are taken by him to be important conceptual elements of the Freudian thought on the mind, are necessary to provide an explanation of irrationality.

⁴ After all Davidson altered considerably his theory of practical reason over the years mostly in order to be able to account for the existence of akratic acts, which he took from granted.

Three Freudian Theses. ('Paradoxes of Irrationality', p.170, 171)

1. **Partitioning.** The mind contains a number of semi-independent structures, these structures being characterized by mental attributes like thoughts, desires, and memories.
2. **Structure.** Parts of the mind are in important respects like people, not only in having (or consisting of) beliefs, wants, and other psychological traits, but in that these factors can combine, as in intentional action, to cause further events in the mind or outside it.
3. **Causal relation.** Some of the dispositions, attitudes, and events that characterize the various substructures in the mind must be viewed on the model of physical dispositions and forces when they affect, or are affected by, other substructures in the mind.

By postulating a partitioning of this kind Davidson is basically extending the process of interpretation and charity, by attribution to the agent sub-divisions that are more rational than the agent as a whole. This process of division should be seen as a metaphor and not as a step into depth psychology – its justification is a result of the logic of the charity of interpretation. Its function is to provide a conceptual division *needed* to explain irrationality.

So let's see how it can handle the case of the aforementioned akratic agent. The akratic person has a reason R, constituted in part by a strong desire to insult Wayne. However his best judgement, having into account all of his reasons⁵ tells him not to act on that desire. Still he acts on R; R causes him to insult Wayne. Davidson explains this by saying that his desire to insult Wayne has a double role; first it is a reason to perform the insult, which however, as a reason, it is defeated by the sum of his other reasons not to perform the act. So here enters R's second role, the role that is responsible for the irrationality proper; by going against his best judgement, he is in contradiction with a principle that he holds, the principle of continence. Moreover his reason R is not a reason against the principle, it does not logically override the

⁵ Or having into account a substantial part of his reasons. Davidson sometimes speaks of the better judgement.

principle. In fact since the principle of continence is a principle of rationality, there are no reasons against it⁶. So in its second role R overrules the principle of continence, but R cannot do it by rational means. Davidson solution then is that R causally overrides the principle by being causally related with action or the intention to act without being a reason for the act or the intention. It causes them on the model of pure physical causation. Davidson puts in different non-overlapping parts of the mind those mental states that cannot enter in rational causation. Each partitioning then can be full rational. Akratic actions and irrationality in general arise from mental causes that are not rational causes, and so do not make reasonable the attitudes they cause.

While partitioning and structure make sense of the idea of different parts of the mind with full-blooded intentionality (the intention to insult is a full blooded propositional attitude and thus is part of a web of propositional attitudes) the third of the Freudian thesis says that non-overlapping parts of the mind interact on the model of non-rational causes.

The function of the division is to show us how the agent can fail to stand for his own principles. So we can explain how one person can fail to follow the principle of continence by putting the principle and the intention that goes against it in different parts of the mind, parts that do not have rational access to each other but still interact causally.

Heil's complaint is that this model is too cumbersome. For him this is important. Recall the argument: First he provides a different model and then he tries to show that his model is simpler than Davidson's. We have seen that his model is a failure, but still this latter imputation is still more than he can properly defend. Here is how he pictures it

The cause of the irrational item – here an intention – cannot be a reason for it. A desire to A remains a reason to form the intention to A even when one judges it best not A, hence the desire to A cannot be the irrational cause. What is irrational, however, is my forming the intention to A on the basis of a desire to A together with (i) my judging that, all things considered, I ought not A, and (ii) my acceptance of the

⁶ Davidson is a little ambiguous about it, as he seems to say in 'Paradoxes of Irrationality' that only a person that holds the principle could be said to be inconsistent upon doing akratic acts. However in 'Incoherence and Irrationality' he is clear about the matter. The principle of continence, being one of the principles of rationality is not something that a rational creature might decide not to hold. He claim that it is constitutive of being a rational animal to accept those principles. This seems to be right and more in step with his other writings.

principle of continence...The cause of my forming the intention to A, then, is my being in the complex state comprising all (or most) of what is included in partition A [at least the principle of continence, the best judgement, the desire to A] (p.79)

I too think that this model is too complicated and difficult. In particular it does not seem plausible that my best judgement, that goes against A, and the principle of continence are part of the cause of the intention to A. I think however that this is not Davidson's view. Heil goes wrong on two accounts. First, because he pictures the desire to A in the web of propositional attitudes that constitute the intersection of the partitions, it follows that there is an element of rational cause in the process, whereas the model intends to explain irrationality by position a model of non rational cause. There is a partial cause that is also a rational cause, namely the desire A. But then there is still an element of contradiction lurking, for he still has a rational element causing something for which it isn't a reason in light of his principle of continence. Moreover, the fact that Heil puts the principle of continence to be part of the cause seems to me somewhat perverse. I see no justification for it.

If the desire to A is to be part of the cause the desire has to be outside of the partitioning that includes the intention. Heil does not see that this is necessary because he thinks that

The relevant cause and effect pair cannot be my desire to insult Wayne and my forming an intention to carry out the insult. The desire is, if anything, a reason for the intention. (p.579)

But the point of partitioning is to find the conceptual boundaries in the mind so that one can make sense of the overruling of a principle of rationality, and the boundary is between the relevant cause and effect. And the relevant cause is what in normal cases of intentional action is *the* reason for the action or formation of the intention. There is nothing else than the desire to A (or the reason that includes the desire A): that is the cause. The point of partitioning is to show how an attitude can cause without being a rational cause. So while or in the process of explaining irrationality the desire to A causes non-rationally. Heil seem for a while to miss this function of partitioning. He seems to look at the logical relation between the desire to A and the intention to A and thus rules them out as being related in a way the one causes the other without being a reason for it. But by being in different partitioning that is what happens, that is the function of thesis 3. I think this view is much more natural. The overruling of the principle of continence is

simple a matter of a reason for A to be a non rational cause of A⁷. The overruling of the principle comes about by a reason that causes non-rationally. It's a matter of blind causality, which does not see reason nor inconsistencies lurking.

Though I am not absolutely sure that this is what Davidson has in mind I find it consistent with what he has to say, and a much better model than that of Heil's.

References

Davidson, D.

1969, 'How is weakness of Will Possible?' In Davidson, 2001a

1970, 'Mental Events' in Davidson, 2001a

1974, 'Psychology as Philosophy' in Davidson, 2001a

1982, 'Paradoxes of Irrationality' in Davidson 2001b

1985, 'Incoherence and Irrationality' in Davidson 2001b

1986, 'Deception and Division' in Davidson 2001b

2001a, *Essays on Actions and Events*, Oxford: Clarendon Press

2001b, *Problems of Rationality*, Oxford: Clarendon Press

Heil, J., 1989, 'Minds Divided' , *Mind*, Vol 98

⁷ The paradox is thus resolved because being in different partitions, as in being in different minds, a mental cause can cause without being a rational cause.

