

Dispositions, Rules, and Finks

Toby Handfield and Alexander Bird

Accepted 10 July 2007

Abstract

This paper discusses the prospects of a dispositional solution to the Kripke–Wittgenstein rule-following puzzle. Recent attempts to employ dispositional approaches to this puzzle have appealed to the ideas of finks and antidotes – interfering dispositions and conditions – to explain why the rule-following disposition is not always manifested. We argue that this approach fails: agents cannot be supposed to have straightforward dispositions to follow a rule which are in some fashion masked by other, contrary dispositions of the agent, because in all cases, at least some of the interfering dispositions are both relatively permanent and intrinsic to the agent. The presence of these intrinsic and relatively permanent states renders the ascription of a rule-following disposition to the agent false.

1

C. B. Martin and John Heil (1998) claim that they can explain what it is to possess a rule, and be following a rule, in terms of *dispositional capacities* of agents. In particular, they believe their account addresses the concerns raised by Saul Kripke (1982) – himself expounding ideas of Wittgenstein – that lead to a sort of scepticism about rule-following.

Kripke's sceptical problem concerning rule-following may be understood as originating with the following question. Don is a subject who intends by the term 'plus' exactly the normal addition function, plus. Van is a subject who intends by the term 'plus' something rather different, a bent addition function, quus, which may be explained thus: $x \text{ quus } y = x$ plus y when x and y are both less than 57, and 5 otherwise. So long as neither Don nor Van have been given an addition problem with x or y greater than or equal to 57, then the behaviour of each will be the same.

This is a pre-print of a paper forthcoming in *Philosophical Studies*, available from www.springerlink.com.

So the question is, what facts about Don and Van make it the case that they mean something different by ‘plus’ (that Don means plus and Van means quus)?¹

A *prima facie* attractive answer to this question is that Don and Van differ in their dispositions. Although neither has been give the sum ‘68 plus 57’, they are disposed to give different answers to it. Don’s disposition is such that he would answer 125, Van’s disposition is such that he would answer 5.

Kripke himself rejects the dispositional answer, raising three problems:

- (i) The dispositional approach cannot account for the distinction between making a mistake and adopting a new rule. If a subject Don were to answer ‘5’ in response to a request to sum 68 and 57, then he is so disposed. But as a consequence we cannot distinguish between the case where Don is attempting to follow the normal plus rule, but *making a mistake*, and the case where Don is *correctly* following Kripke’s bent quus rule.
- (ii) The dispositional approach cannot account for normativity. Dispositions govern what a subject actually does. How do they account for what a subject *ought* to do?
- (iii) Our dispositional states are finitely bounded. But the domain of rules is infinite. The plus rule applies to numbers that no human could even entertain let alone add.

The term ‘normativity’ may be thought to raise questions of justification, but if we put these on one side – as Martin and Heil do – we see that (i) and (ii) are two sides of the same coin.² Rules tell us what the *correct* answer is, what answer we *ought* to give; but dispositions explain what we actually do. So dispositions and rules cannot be the same. Conversely,

¹The discussion of these issues is sometime framed in terms of following a rule and sometimes in terms of meaning. For current purposes we take the relationship to be as follows: if S means or understands plus by ‘plus’ then when S attempts to answer correctly a question expressed in the following terms ‘what is x plus y ?’ S intends to follow the plus-rule.

²It should be noted that for Kripke, the issues of justification are in fact a crucial aspect of the sceptical problem he is attempting to raise: “Ultimately, almost all objections to the dispositional accounts boil down to this one” (1982, 24). In this paper, however, we are only focusing on the problem of providing an account of what fact about an agent makes it the case that they are following one rule rather than another. We have nothing to say about what *justifies the belief* that an agent is following one rule rather than another.

if we were following a rule which corresponds precisely to the disposition which governs our actual behaviour, then we could never be accused of making a mistake. For the rule corresponding to the disposition of one who apparently errs, would be a bent-looking rule, one that deviates from a straight rule on precisely the occasion of the seeming error.

The infinitude of rules is a different problem. Dispositional accounts might be thought to be well equipped to explain this feature of rules, because dispositions are necessarily “connected” to possible manifestations. The possible manifestations of a given disposition extend, as Martin and Heil put it, “indefinitely” (297). Thus there is some prospect of capturing the infinite applicability of rules in the indefinitely many possible manifestations of a disposition.

The problem is not, however, one simply of an infinite *variety* of possible stimuli and manifestations, for a fragile vase may break in infinitely many slightly different ways in response to infinitely many slightly different strikings. In this case it is the *unbounded* nature of the possible stimuli and required manifestations that is problematic. If the disposition to behave in conformity to the “plus-rule” is understood in terms of a subjunctive conditional of the form: “Were the subject to be asked to sum any two numbers, she would give the correct answer”, then it is manifestly false that anyone possesses this disposition.³ No one possesses this disposition – or this “subjunctive property” – because of various possibilities of error and also because of various limitations of human agents. Someone can have grasped the rule for addition, but remain susceptible to making errors in particular cases. Moreover, even the most reliable human will not be able to add some numbers correctly, for the numbers might be so large that it takes more than a human lifetime to entertain them, let alone perform a calculation with them.

2

As the last paragraph suggests, Kripke’s criticisms of the dispositional account depend on something known as the simple conditional analysis of dispositions:

- (CA) S is disposed to yield manifestation *m* in response to stimulus *s* iff
were S to receive stimulus *s* it would yield manifestation *m*.

Let S be an individual of good but not abnormal intellectual skills. For enormously large *x* and *y* it is almost always true that in response to the

³As Kripke stresses: 1982: 26–7.

stimulus that is the request ‘what is x plus y ?’ S would fail to yield the manifestation of answering with ‘ z ’, where $z = x$ plus y . The contrapositive of the left-to-right implication of (CA) licenses the inference that S is not disposed to yield manifestation ‘ z ’ in response to stimulus ‘ x plus y ’. But that disposition (for all x , y , and z , where $z = x$ plus y) is what, on the dispositional account of meaning, meaning plus by ‘plus’ is supposed to be. This alleged refutation of the dispositional account of meaning depends crucially on (CA).

Similarly, Kripke’s criticisms (i) and (ii) have force only on the assumption that if a subject does exhibit m under stimulus conditions s , that is because the subject was *disposed* to m in response to s ; that is, Kripke’s criticisms assume the right-to-left implication of (CA). In response to the request ‘what is 189 plus 277?’ S answers ‘456’ (whereas 189 plus 277 equals 466). The right-to-left implication of (CA) licenses the inference that S was disposed to answer 456’ in response to ‘189 plus 277’ and so the dispositional view of rules implies that by ‘plus’ S meant some bent function for which 456 is the correct answer. Meanwhile the left-to-right implication of (CA) again shows that S was not disposed to give the answer 466, and so did not mean plus by ‘plus’ – and so S did not *err* in giving the answer ‘456’. (CA) is thus central to Kripke’s claim that the dispositionalist account of rules makes error impossible.

Martin and Heil are, however, able to reject Kripke’s criticisms of the dispositional account of rule following precisely because they are able to reject (CA), for good reasons that have been widely discussed. A disposition can be present in suitable stimulus conditions, but fail to manifest in virtue of interfering factors, known as “finks” (Martin 1994) and “antidotes” (Bird 1998).

A fink is a feature of a situation which, upon the disposition being exposed to its stimulus, will act so as to remove the disposition’s causal basis before it can manifest. For instance, in an example from David Lewis (1997), a sorcerer protects a fragile glass by waiting to see if it is struck. If ever a striking occurs, the sorcerer will very quickly cast a spell which makes the glass non-fragile. The spell works more quickly than the process by which the striking leads to breaking, thus the glass, if it were struck, would not break.

In a more prosaic example, a household fire, happily roaring in the fireplace, is disposed to burn the entire building down should it get too hot. However, the owners are well aware of this danger, so they have set up a safety mechanism in the fireplace. If the fire does get that hot, a heat

detector sets off a sprinkler system which douses the fire. The stimulus (the fire's being too hot) to the disposition also causes, by an indirect but faster route, the fire to lose the very disposition in question (the disposition to destroy the building).

Antidotes – also known as *masks* (Johnston 1992) – work differently. Rather than remove the disposition they interfere with its normal process leading to manifestation. The disposition remains in place but the conditions for its manifestation are removed. Continuing with the fire example, a building could be so protected that the heat detector causes fire-proof steel plates to fall into place around the fireplace that prevent the fire from spreading. The fire itself is not lessened, so its disposition remains, but it cannot have its normal effect.⁴

The possibility of finks and antidotes show that the left to right implications of (CA) are false. Each has its converse that shows the right to left implication to be false. A finkish circumstance may be set up so that a disposition is *absent*. But on the occurrence of its appropriate stimulus a finkish mechanism brings the disposition into existence quickly enough for it to combine with the stimulus to bring about the manifestation. Thus at the time t when the stimulus occurred, the disposition was absent, yet the manifestation subsequently occurred, and so it was true at t that were the stimulus to occur the manifestation would occur. A mimic (Johnston 1992) to a disposition will not bring the disposition into existence but nonetheless brings about its characteristic manifestation in response to its stimulus by other means. For example, an iron cooking pot is not fragile, but does break when struck, since attached to the pot is a grenade with a sensitive detonator. The counterfactual on the right of (CA) is true, but not the dispositional claim on its left.

In the case of dispositions which are the basis of rule-possession, Martin and Heil suggest that it might be the case that these dispositions are also vulnerable to finkish or masking interference. It is their hope, then, that an agent who is following the plus-rule does indeed possess a

⁴Some cases seem to blur the distinction between finks and antidotes. For instance, consider a building which is disposed to burn down if a fire is lit in the lobby. A fire is lit in the lobby, and some fire-fighters start to spray the building with water. The activity of the fire-fighters is arguably fink-like, because they make the building saturated with water, hence take away its disposition to burn down. Arguably, however, the spraying of the water is also an antidote, as it interferes with the process by which the building would normally burn down. We see no objection to concluding that, depending on the details, the activity of the fire-fighters could be both a fink and an antidote. A condition would be both a fink and an antidote to D if it both removes D before D can manifest and also interferes with those conditions that are required for the normal manifestation of D.

disposition D, the manifestation of which is the correct answer. While it is true that not every agent who possesses D would give the correct answer to every sum, this is *not* because they lack the disposition. Rather, it is because of the presence of interfering finkish or masking circumstances.

3

The fact that finks and antidotes refute (CA) certainly opens up the possibility of reviving a dispositional account of meaning that does not fall to Kripke's criticisms. But for this possibility to be confirmed as actual it must be shown that the problematic cases alluded to can indeed be accounted for as finkish or antidote cases. We question whether matters are as straightforward in this regard as Martin and Heil seem to suggest.

The first issue requiring clarification is whether it is finkishness or antidotes to which Martin and Heil should be appealing. If it is the former, then the disposition in question should be lost; if it is the latter the disposition remains but has its normal operation interfered with.⁵ The dispositions are ones which correspond to rules: meaning addition by “plus”, intending to follow the rule for addition, etc. When someone makes an error in a calculation is that typically because they lose such dispositions? Or are they retained but interfered with? Both seem possible. Perhaps a very large addition sum so taxes the subject's brain that he mentally collapses and ceases to grasp the rule for addition at all. Perhaps the subject gets so confused that he starts following some other rule. These would be cases of finkishness. But in many cases one may suppose that the subject's dispositions are left intact. If while employed on an addition sum a loud noise distracts me and I make a slip in carrying that I would not otherwise have done, then I was following the same rule all along but was caused to fail in my execution of it by interfering circumstances. Thus some of the cases that show the falsity of (CA) with respect to rule-dispositions are finkish cases but many, perhaps most, will also be antidote cases. The discussion that follows will concentrate initially on finks and then return to antidotes.

A typical fink is an *extrinsic* feature of a disposition-bearer's situation. It is thus possible – in typical cases – to draw a clear distinction between the intrinsic causal basis of a disposition and its extrinsic fink. An object

⁵Strictly speaking the disposition *normally* remains in antidote cases – but need not in cases which are also fink cases – c.f. footnote 4. In what follows it is not necessary to consider cases of combined finks and antidotes.

possesses a disposition in virtue of its causal basis, and while an extrinsic fink may render the associated conditional sentence false, it does nothing to change the intrinsic causal basis of the disposition in question.⁶ Likewise, antidotes are also typically extrinsic interfering conditions, not intrinsic ones.

Regarding the sorts of disposition which Martin and Heil believe are the basis of rules, the factors in some cases are extrinsic – the distracting loud noise is a clear example. However, it is less easy to be confident that all the interfering circumstances are *always* extrinsic to the rule-follower. People often make mistakes that cannot be attributed to external causes, but are the result of their own failings. A putatively rule-following agent, Saul, is given the task of adding 187 to 23, under propitious circumstances of perfect silence, plentiful but not excessive supplies of coffee, not to mention pencils, paper and so forth. But he just forgets to carry or he loses concentration and makes some other error. The causes of the mistake are attributable to something intrinsic to Saul. Or let Saul be given a mental arithmetic problem involving numbers too large for him to hold in his mind. The failing again is Saul's: if he had a better memory, a better "head for figures", he could have entertained and maybe solved the problem that has defeated him.

If Martin and Heil are correct to say that Saul is a rule-follower in virtue of a disposition, and correct to say that his error is to be explained in terms of a fink to the relevant disposition, it would appear that both the fink *and* the disposition are intrinsic to the agent at the same time.

One ought to be suspicious of this proposal: we suggest that there is good reason to think it is not possible for an intrinsic disposition to obtain in the presence of an intrinsic fink. This is plausible if one thinks about classic examples of finkish scenarios, and modifies them such that the fink is an intrinsic feature of the disposition-bearer. Suppose Lewis's sorcerer, protecting the fragile glass, decided that, to be on the safe side, he would enchant the glass in such a way that it has an intrinsic property which would cause it, when struck, to lose the molecular structure

⁶It has recently been argued by Jennifer McKittrick that (1) dispositions need not have any causal basis (2003b) and that (2) dispositions can be extrinsic (in particular, they may have extrinsic causal bases) (2003a). Martin and Heil reject the idea of a dispositional property having a *distinct* causal basis, but also reject the idea that a disposition could be "bare", lacking any causal basis. Thus it is easiest to treat their view as one whereby the dispositions they are interested in are intrinsic properties which are *identical* to their causal basis. While this view may encounter difficulties with respect to some dispositions, for current purposes Martin and Heil need only defend it with respect to the particular dispositions of human agents which are the basis of rule-possession.

which, under normal circumstances, is the causal basis of fragility. In other words, he gives the glass an additional intrinsic property, which is fink-like with respect to the original fragility of the glass. The glass, if struck, would not break, because of the enchantment. It seems plausible to say that such a glass is simply not fragile. It has lost the disposition of fragility because it has lost its intrinsic causal basis for fragility. This loss of the causal basis has occurred not by *removing* the molecular structure, but by adding another intrinsic property – the enchantment – which we do not typically associate with the molecular structure of glass.

It is important in such cases to be clear about the precise entity to which the disposition is being attributed. A uranium pile above critical mass is disposed to meltdown.⁷ But external to the pile are boron moderating rods connected to a fail-safe mechanism that in response to high levels of radiation drop the rods into the pile thereby stopping the meltdown. Consider the uranium alone. It retains its disposition for meltdown all along (the intrinsic basis of highly fissile U-235 atoms remains). Consider the uranium *plus* boron rods in the raised position. This has the disposition to meltdown, as would occur if this arrangement does not change. However the fail-safe mechanism changes the intrinsic relationship between rods and uranium, lowering the rods into the pile, so that now the composite entity has no meltdown disposition. Hence, in the presence of the fail-safe mechanism, the combination of uranium and moderating rods has its disposition to meltdown finkishly. Now consider the combination of uranium, boron moderating rods, *and* fail-safe mechanism. That combination has no disposition to meltdown (assuming that the fail-safe mechanism works as it should). The lesson is that an object x may have a disposition finkishly in virtue of some extrinsic mechanism y . But if y is part of some combined object x' , and so the fink is intrinsic to x' , then x' lacks the disposition altogether.⁸

The forgoing considerations concern intrinsic finks. Do they extend

⁷This example is borrowed from Bird 1998.

⁸See Choi 2005: 499–500, whose “nomic duplicate” heuristic for determining whether or not an object has a disposition coincides with this thought. Choi’s suggestion is that, if we are unclear whether an object has a disposition, perhaps due to strange, possibly finkish, circumstances, we ought to ask ourselves: is there a possible intrinsic duplicate of this object, subject to the same laws of nature (a “nomic duplicate”) which clearly does possess the disposition? Applying this test to an object like the enchanted glass will never yield a case where it is clear that the object has the disposition, because all possible intrinsic duplicates of the glass will also possess the enchantment.

See also Handfield Forthcoming where Choi’s point is developed so as to apply to all dispositions, both intrinsic and extrinsic.

to intrinsic antidotes/masks? It seems that they do. Consider Joon and Jim. Joon is lactose intolerant, which is to say that normally the ingestion of lactose (contained in milk) causes discomfort, since Joon does not produce the enzyme lactase which metabolizes lactose. In contrast Jim's body does produce lactase. Occasionally Joon drinks milky beverages, but in order to avoid the discomfort this would normally cause, he mixes lactase into his drink so that while the milk is in his gut the lactose it contains is broken down by the lactase he consumed with it, just as it would be in Jim's gut, except that in the latter case the lactase is produced naturally. Even when drinking milk mixed with lactase Joon still has the lactose intolerant disposition, only the lactase in the drink acts as an antidote. But why not say then that Jim is *also* lactose intolerant and that his (naturally-produced) lactase also acts as an antidote? The reason why not is that we regard Jim's lactase production as a process that is intrinsic to Jim whereas Joon's ingestion of lactase mixed into his drink is extrinsic to Joon. Jim's case shows that his possession of (what would be) an intrinsic antidote to a disposition D (lactose intolerance) means that he does not have D at all.

We can conclude therefore that as regards both fink and antidote cases if S contains an intrinsic fink or antidote to some disposition D, then S does not possess D. If that conclusion is correct, then Martin and Heil's defence of the dispositionalist account of rule-following cannot work in many of the relevant cases, those where the failure of the rule-follower to act in accord with the rule is a consequence of factors *intrinsic* to the rule-follower.

4

The conclusion we reached in the preceding section may be contested. According to Michael Fara⁹, at least for some objects and for some dispositions that we ascribe to those objects, there are some cases where it is plausible that an object can instantiate both a disposition and an intrinsic mask to that disposition. Fara uses examples such as the disposition to get a stomach-ache from eating highly acid foods like lemons. Suppose I have that disposition. However, if I consume some milk before having a lemon, my disposition will be – for a short time – masked by the milk: I won't get a stomach-ache at all. Plausibly, the milk acts as an *intrinsic* mask to my disposition to get a stomach-ache. It changes my intrinsic properties such

⁹"Masked Abilities and Compatibilism", unpublished ms. Draft version of 9 February, 2007.

that the process from ingesting lemon juice to having a stomach-ache is interfered with.

Immediately after consuming the milk, but before I consume any lemons, am I disposed to get a stomach-ache from eating lemons? Yes, claims Fara. On the other hand, he concedes, if the change in my intrinsic properties is relatively *permanent* – achieved for instance by an operation on my digestive tract to improve my ability to digest acidic foods – then I will indeed lose the disposition to get a stomach-ache.

We are not sure that our intuitions track Fara's in every case. Must we concede that the ingested milk is intrinsic to me? Something's being located internally to x does not make it intrinsic to x – it is not an intrinsic feature of my office that it has me in it. Likewise an indigestible lead pellet that enters my stomach along with food need not have a duplicate in the stomach of every one of my intrinsic duplicates. Insofar as the milk has not been digested but is simply acting as a barrier lining my stomach, perhaps it too is not yet intrinsic to me. On the other hand, one might feel that to the extent that the digested milk has become intrinsic it does, albeit temporarily, remove the disposition to get a stomach ache.

Let us suppose, nonetheless, that Fara is correct. In that case, our earlier claim simply needs only a small refinement. As he has conceded, an object can possess disposition D and an intrinsic fink or antidote to D together, but only if the fink or antidote is relatively temporary. And so if an object which putatively instantiates a disposition also instantiates a relatively permanent, intrinsic fink, or a relatively permanent, intrinsic mask, then it simply does not possess the disposition at all.

Fara's point, if accepted, will evidently assist Martin and Heil in explaining some cases of failure to follow a rule. But his concession means that their explanation fails for other cases.

Here is a case to which their explanation will still apply. Suppose that Saul fails to successfully compute a sum because he is temporarily under the influence of alcohol. Without the alcohol in his system, let us simply grant Martin and Heil that he does have the disposition to follow the rule for addition. But while intoxicated, he appears to have an intrinsic fink or mask for this disposition. According to our earlier discussion in Section 3 we would therefore conclude that he has simply lost the disposition. If Fara is correct, however, provided that the intoxication is a relatively temporary property of Saul, it would still be correct to say that he has the disposition to follow the rule, even while he is so intoxicated that he cannot manifest this disposition.

But not all cases are like this. Some of the intrinsic factors that interfere with our ability to do sums correctly are both intrinsic and relatively permanent. Most obviously, my inability to hold twenty digit numbers in my head or my inability to write down numerals with 10^{10} digits are intrinsic masks to the plus-disposition that are relatively permanent. An imperfect power of concentration may also be permanent. So our complaint against Martin and Heil goes through: if the finkish or antidote-like features of an agent which prevent the agent from correctly following a rule are both intrinsic and permanent, then it is not correct to say that the agent nonetheless has a disposition to follow the rule.¹⁰

5

But let us try to allow Martin and Heil that an intrinsic disposition can be co-instantiated with an intrinsic fink for that disposition. Consider the disposition of a human body to wither and decay within the next 200 years. This disposition is, plausibly, intrinsic and is certainly highly enduring. One could perhaps inhibit the expression of this disposition by a cryogenic procedure, but that would not remove the body's disposition to decay; it would rather be an antidote to that disposition. Thus it is true that:

1. Saul is disposed to wither and decay within the next 200 years.

Now consider an instance of the plus rule, x plus y , which requires at least 200 years to apply, due to the very great magnitude of the numbers x and y involved. Martin and Heil wish to maintain that an agent such as Saul

¹⁰To be fair to Martin and Heil, they do appear to prefer to talk of *capacities* rather than dispositions. And while clearly the concepts of disposition and of capacity are similar, there might be crucial differences, at least in ordinary parlance. In particular, it seems that ascribing a capacity to someone is to make a weaker claim than to ascribe a disposition to them. I am certainly capable of smoking. I have that capacity (that power, that ability). But I am not disposed to do so. Conversely, however, it is hard to imagine someone having a disposition to do something for which they lack the capacity.

Even if capacities are weaker than dispositions, they nonetheless suffer also from finks and antidotes. If we were to rephrase the current discussion in terms of capacities, it could still proceed as currently. Let us consider the capacity to follow the plus-rule. The intrinsic and enduring features of humans, in virtue of which we die and decay (see Section 5), are surely incompatible with having the capacity to follow this rule for very large x and y . So even if Martin and Heil can point to intrinsic differences between Don and Van in virtue of which they have different capacities to add small numbers, it remains the case that the capacities of Don do not straightforwardly show that he is following plus, rather than some other rule, which will return deviant values for sums such as x plus y .

could possess the plus rule, in virtue of a disposition to give the correct answer to this sum.

2. Saul is disposed to give the correct answer to the question ‘what is x plus y ?’.

Giving the correct answer, however, is not compatible with withering and decaying. And giving the correct answer is not physically possible in less than two hundred years.

So Saul is – supposedly – both disposed to wither and decay within two hundred years, but also disposed to give the correct answer to the sum of x and y , even though to give that answer would require living for at least 200 more years.

Is it possible for an agent to be “oppositely disposed” in this fashion? To put it more bluntly, could an object be disposed both to ϕ and not to ϕ , under the *same* stimulus conditions?¹¹

Remarkably, David Lewis (1997, 157) entertained this possibility. He suggested that an object which was disposed to break when struck could also possess an intrinsic fink to its fragility. Such an object, he said, would be both disposed to break when struck, and disposed not to break when struck. While doubtful about the coherence of this response, we shall grant that it is not immediately objectionable to see what its implications are for Martin and Heil.

If Martin and Heil were to embrace this account of the compatibility of Saul’s intrinsic disposition to decay and his rule-disposition D, they would have to give some account of why possessing D suffices for possession of the rule, *despite* the presence of the contrary disposition to decay.

Given the foregoing, the simplest and most natural account of rule possession clearly will not work:

- (R) S possesses a given rule just in case S possesses a disposition to give responses in conformity with that rule.

For on this simple account, an agent like Saul possesses not only the plus-rule, but also various deviant rules. For Saul no doubt instantiates very many finks (and antidotes) for the disposition for plus. He has limited memory, and so will not be able to entertain large numbers or he will forget things while calculating. He makes mistakes due to his tendency for daydreaming. And so on. Each of these finks gives rise to a disposition

¹¹We will assume that the dispositions under consideration are deterministic. It is much less implausible that an object could have both of two *probabilistic* dispositions with incompatible manifestations. Thanks to Allen Hazen for this point.

to follow a rule contrary to the plus-rule. Suppose that Saul does possess the plus rule but in carrying out the addition 189 plus 277 makes a mistake and gives the answer ‘456’ on account of some intrinsic and relatively permanent fink, such as a poor memory, that means that he forgets to carry. Allowing for dispositions with incompatible manifestations means it can still be true that Saul has the disposition to answer with the correct answer ‘466’. Nonetheless it is also true that Saul has the disposition to answer ‘456’. Given (R), however, this means that Saul possesses – and, presumably, was following on this occasion – both the plus-rule *and* some deviant rule for which ‘456’ *is* the correct answer. So Saul appears to be conforming to multiple contrary rules at the same time. This is a pyrrhic victory for Martin and Heil.

What is required is some reason to suppose that one of Saul’s many properties is the basis for a *privileged* disposition which constitutes his possessing the plus-rule. And given that many properties are causally relevant to Saul’s behaviour, there needs to be some principled basis for distinguishing one disposition as deserving of this title. This we call the *privileging problem*, and we believe it is a pressing one for this approach.

6

One possible line of answer to the privileging problem is to suggest that Saul’s possession of the rule is marked out by some *component* of his nervous system having the relevant disposition, while all the finks and antidotes to this disposition are thought to be somehow separate from this component – thus generating the required separation between disposition and interfering factors. Indeed, Martin and Heil distinguish between an object’s overall “dispositional condition” and the “components” of said condition. The components are themselves “a sort of capacity”, involving “possession of a particular dispositional condition” (300–1). Their thought appears to be that while the agent’s dispositional condition might not directly include the disposition that is the basis of possessing a rule, the agent’s condition might have a suitable *component* which bears the right disposition.

Consider a pocket calculator with a ten-digit display, a corresponding ten-digit limit to both input and output, and an equally limited memory. One might think of the process of performing an addition on this calculator as a matter of deploying various distinct modules. The display records the input and shows the output; the memory stores the input and vari-

ous products of calculation; and a processor carries out a calculation on the numbers stored (encoded) in the memory. One might think that the third of these is where the plus-rule is really carried out. The display and memory are tangential. It is true that the calculator is limited in what it can output and therefore it is limited in which additions it can perform. On the other hand that limitation is not due to any fact about the processor, which is the module we regard as responsible for the plus-rule. That module can add any two numbers given to it. To extend the calculator's adding capacity the display and memory would need to be upgraded, but not the processor. Likewise we might hope to find some component of Saul's nervous system that embodies the plus-rule and plus-disposition that is distinct from those parts that act as finks and antidotes to the plus-disposition.

While we think that some development of this proposal may be illuminating, as it stands it is still inadequate.

One way of construing the proposal is to suppose that, by focusing on components, we are introducing a *physical* separation between the part of the nervous system that instantiates the disposition which grounds the rule and the parts of the nervous system which instantiate various defeating dispositions such as finks and antidotes. As a result, any finks or antidotes would be extrinsic to the component which instantiates the rule-grounding disposition, and the dispositional account would not be threatened.

This way of understanding the proposal, however, makes at best incomplete progress in answering the objection. It might help with antidotes or finks that arise from an inadequate memory for large input numbers. But it cannot help with the problem of decay that was raised in the preceding section. For whatever part of the brain is said to bear the rule-grounding disposition, that part of the brain will itself be a physical entity made of grey organic matter, subject to various processes of decay and decomposition. Therefore, as before, for any putative rule-grounding disposition in any physical part of the brain, there will be factors *intrinsic* to the relevant part of the brain which render it false that the disposition obtains.

A further problem concerns, again, the precise entity to which the disposition is being attributed. Hitherto we have been understanding the relevant dispositions as behavioural dispositions – dispositions manifested by answers to questions. That may be too restrictive a conception of the those dispositions, which can also be manifested in acts of silent

mental arithmetic. But, either way, the manifestations are still those that are attributable only to a whole person, not to sub-personal components. So there remains the problem of linking the intended whole-person disposition to a subpersonal disposition of a rather different kind. Note that the subpersonal disposition may not be a simple subpersonal analogue of the whole-person disposition. Indeed, in the absence of the memory components (which have been factored out since they have a masking effect), what results may not look much like a disposition to add multi-digit numbers. The overall algorithm will be recursive, the principal adding step will involve, on the n th iteration, adding two single-digit numbers (the n th digits in the two multi-digit numbers to be added – and adding 1 if required, when there is carrying from the $(n - 1)$ th iteration). To get from that process to what can be regarded as the adding of two multi-digit numbers is likely to require the employment of memory, the very component we were excluding from consideration.

The final issue is that the privileging problem has not been answered, but has simply been transferred to a different location. The requirement remains, that a principled reason should be given for circumscribing some particular part of the subject as that which fixes the dispositions to be attributed to the whole. Why is it the processor alone that fixes the disposition of the whole and not the part that is the processor plus memory, which would ascribe a different disposition to the whole? The privileging problem has been moved from being a problem of selecting the privileged disposition of the whole to the problem of selecting the privileged part of the whole.

We have looked at the response that seeks to attribute a disposition to the whole on the basis of the dispositions of a proper physical part of the whole. Another way of interpreting this line of reply is to suppose that the separation between components is a separation into distinct *properties* of the whole. But what kinds of properties? If single dispositional properties, then we are exactly at the point we reached at the end of the preceding section. Certainly, there may be many properties of the person that may be manifested in the behaviour of an agent. Abstracting the properties away from each other, each property can perhaps be associated with a different rule. But in virtue of what do Martin and Heil say that just *some* of these properties constitute *the* disposition to follow the rule in question? So it looks as if the properties in question must be property-complexes that are the causal bases of the dispositions we are looking for. For example, it may be that it is certain properties of the neurons plus their ar-

rangement that allows one to compute the plus-function. This complex of properties may be abstracted away from the property of the neurons that is their being made out of organic material disposed to decay. This complex of properties might, in principle, be a complex of properties also possessed by a non-biological entity. This complex will support the ascription of the plus-disposition. Still, the principal problem remains, that of finding a principled reason for privileging this complex of properties rather than any other. So for example, Saul also possesses the complex of properties that is the combination of this one plus the property of being made out of decay-prone organic material, which has an inbuilt fink or antidote to the plus-rule and so may be considered to be the basis for a deviant alternative (in answer to ‘what is x plus y ?’ answer with x plus y for small and moderately large x or y , otherwise get sick and die’).

Note that our objection is not that Martin and Heil *cannot* succeed in answering the privileging problem. It is indeed plausible that some future neuropsychologists will be in a position to answer questions like this. Indeed those who think that the mind is modular also need to come up with an answer to an analogue of the privileging problem – what is it that requires us to identify certain parts of the brain or certain dispositions or property-complexes as being modules, whereas other ways of carving up the brain or its properties do not correspond to modules. An evolutionary account might be one possible answer, which would link the current question to that of evolutionary accounts of biological function. Our aim here is certainly not to argue that a dispositional account of rule-following is impossible. But neither is it to suggest how an answer to the privileging problem might rehabilitate the dispositional account. Rather we wish to show that the philosophical task of distinguishing between the good and bad dispositional factors has not even been commenced by Martin and Heil, nor, to our knowledge, by any other defender of the dispositional account of rule-following.

7

We are sympathetic to Martin and Heil’s anti-scepticism with respect to rule-following. Certainly they are right to point out that Kripke’s criticisms of the dispositional account assumed (but not explicitly) the simple conditional analysis of dispositions, (CA). And they were right also to point out the falsity of (CA), a fact which clearly goes some way towards undermining Kripke’s sceptical claims. But only so far. That is because (CA) devi-

ates from the truth only in certain now well-understood kinds of case. It is still true, as Martin (1996, 178) himself tells us, that there is *some* connection between dispositions and conditionals. The kinds of case in which that connection is broken are finks and antidotes (masks). The problem for Martin and Heil is that the canonical cases of finks and antidotes are *extrinsic* to the object which possesses the disposition in question. But in many cases of a break-down in rule-following, the cause is *intrinsic* to the object. Consequently, it looks *prima facie* that such cases are not cases of disposition-plus-an-antidote (or disposition-plus-a-fink) but are cases of the absence of that disposition – or, even more pressingly, cases of the absence of that disposition *and* the presence of some other disposition with the same stimulus but incompatible manifestation. Which is, unfortunately for the dispositional account of rule-following, exactly what Kripke was claiming in the first place.

If, however, we do permit the existence of a disposition in an object that also has an intrinsic fink or antidote to that disposition, then we are forced to concede that there are many other dispositions that an object has, including dispositions that have manifestations inconsistent with the manifestation of the disposition we are interested in. In which case we now face the privileging problem: on what principle do we single out one of these as the disposition that corresponds to the rule that the subject is following? The privileging problem may have an answer. Finding that answer is a major challenge that defenders of the dispositional account of rule-following have yet to address.¹²

TOBY HANDFIELD

Monash University

ALEXANDER BIRD

University of Bristol

References

Bird, A. 1998. Dispositions and Antidotes. *The Philosophical Quarterly* 48: 227–34.

¹²Both authors gratefully acknowledge the support for this research received from the British Academy and the Australian Academy of the Humanities, in the form of a joint projects grant. Handfield was also supported by a British Academy Visiting Fellowship and an Australian Research Council APD Fellowship while undertaking this research. Thanks to Allen Hazen, John Heil, Bruce Langtry, and audiences at University of Melbourne for helpful discussion. Thanks also to the Institute for Advanced Studies at the University of Bristol.

- Choi, S. 2005. Do Categorical Ascriptions Entail Counterfactual Conditionals? *The Philosophical Quarterly* 55: 495–503.
- Handfield, T. Forthcoming. Unfinkable dispositions. *Synthese*.
- Johnston, M. 1992. How to Speak of the Colours. *Philosophical Studies* 68: 221–63.
- Kripke, S. A. 1982. *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
- Lewis, D. 1997. Finkish Dispositions. *The Philosophical Quarterly* 47: 143–58.
- Martin, C. B. 1994. Dispositions and Conditionals. *The Philosophical Quarterly* 44: 1–8.
- . 1996. Final replies to place and armstrong. In *Dispositions: A Debate*, edited by T. Crane, 163–92. London: Routledge.
- Martin, C. B. and J. Heil. 1998. Rules and Powers. *Philosophical Perspectives* 12: 283–312.
- McKittrick, J. 2003a. A Case for Extrinsic Dispositions. *Australasian Journal of Philosophy* 81: 155–74.
- . 2003b. The Bare Metaphysical Possibility of Bare Dispositions. *Philosophy and Phenomenological Research* 66: 349–69.