

# The role of dispositions in explanations<sup>1</sup>

Agustín VICENTE

BIBLID [0495-4548 (2004) 19: 51; pp. 301-310]

ABSTRACT: According to a model defended by some authors, dispositional predicates, or concepts, can be legitimately used in causal explanations, but such a use is not necessary. For every explanation couched in dispositional terms, there is always a better, and complete, explanation that makes use of a different vocabulary, that of categorial bases. In what follows, I will develop this view, and then argue that there is a kind of use of dispositions in explanations that does not fall within this model. That is, I will argue that we would miss some explanations if we were to forsake dispositional concepts and dispositional explanations.

Keywords: dispositions, explanation, teleology, exclusion problems.

## 1. *Dispositions as inert functional properties*

The view I am about to criticize has it that dispositions are functional properties. According to this view, C.B. Martin's and related examples have shown that dispositions cannot be analysed in terms of subjunctive conditionals. The most famous of these examples is that of the electro-fink which, when a live wire is touched, makes the wire go dead. This kind of counterexamples to the conditional analysis suggest a realist reading of dispositions, which, according to the present view, means a functionalist approach, where dispositions are functional properties whose occupants, or categorial bases, are said to "realise" them.

This general schema, where there is a distinction between two kinds of properties (dispositional and categorial) however, makes room for a (dangerous) problem of "dispositional causation" analogous to the exclusion problem for mental causation. In the present case, the causal exclusion problem would consist in the claims that

- (a) functional properties are causally relevant for some events;
- (b) categorial properties are well-established causes for those same events;
- (c) there is no functional-categorial overdetermination.<sup>2</sup>

Exclusion problems are governed by what Kim has called the 'principle of causal exclusion', which offers possible solutions apart from outright elimination, namely that one of the causes completes the other, or that the two causes are not independ-

---

<sup>1</sup> I have to thank those who attended the workshop *Dispositions and propensities in science* for their various and interesting comments. Many thanks also to Mauricio Suárez for inviting me to the workshop, for his encouragement and his (coming) suggestions, and to Nicholas Maxwell for his lively discussion and friendship. Research for this paper was funded by the BFF2002-03842 project of the Spanish Ministerio de Ciencia y Tecnología.

<sup>2</sup> The usual reason to deny that there is this kind of causal overdetermination is that its eventual defence would go against the assumption of the simplicity of nature. For the rest of the paper, I will assume that this is a good reason.



ent. Dispositions do not seem to complete the work of categorial properties; nor do categorial properties complete the work of dispositions. Neither do dispositions seem to be causes of categorial properties, or to be effects of categorial properties, a state of affairs that would provide a valid solution to the exclusion problem. Rather, dispositions and categorial bases are linked by what we can call a vertical dependency relation, the realisation relation. Hence, the issue of whether dispositions are causally relevant depends ultimately (*prima facie*, at least) on whether the realisation relation provides a solution to the exclusion problem that is respectful of facts (a-c). As Kim's discussion of the exclusion problem has made explicit it is not enough to say that two kinds of properties are linked by a vertical relation for them both to be causally relevant. One has to examine the very nature of the relation, so to speak, in order to decide whether it ensures causal relevance to both *relata*.

In this regard, realisation would be excellent news if there were a type-type identity relation lurking behind it. Identity, of course, is one of the relations that can make two apparently different properties causally relevant for the same effects. But, rather disappointingly, functional properties happen to be multiply realisable and, thus, not identical to their bases.

This fact of multiple realisability has led many to claim that dispositions are second-order functional properties. *Contra* D. Lewis or D. Armstrong, who claimed that functional properties are properties of the sort *having such and such causal effects (and/or causes)*, these philosophers (see, for example, Prior, Pargetter and Jackson, 1982, and Prior, 1985) take it that a dispositional property is best defined as the property of having a property that has such and such causal effects and/or causes. Hence, the relation between dispositional properties and their bases is not of type-identity, but of *supervenience*: given appropriate contexts and laws of nature, two objects instantiating identical categorial bases will instantiate the same functional properties.

Nonetheless, supervenience is no guarantee of causal relevance. Rather, it expresses a pattern of property covariation compatible with emergentism, epiphenomenalism and other nonsolutions to the exclusion problem (see Kim, 1993: *Postscripts*, 1998).

Confronted to this situation, some authors have defended a token-identity approach to the exclusion problem: it is a token of a dispositional property that is identical to a token of a categorial property, even if dispositions-types are not identical to bases-types. This answer, however, has to face the *in virtue* problem, namely, the question whether tokens of dispositional and categorial bases bring about their effects in virtue of being dispositional or in virtue of being categorial. Some answers have been attempted that confer causal relevance to the dispositional type (i.e. tokens inherit their causal relevance only, or also, from the dispositional property: see D. Ehring,

2003, for the latest development of this view),<sup>3</sup> but it is usually difficult to be satisfied with the particular responses produced under this general approach.<sup>4</sup>

Other authors, however, (Prior *et al.*, 1982, Kim, 1998, Heil, 1999) have simply conceded that functional properties, and dispositions with them, are causally irrelevant, that is, that all causal powers reside solely on categorial bases. This is the approach that I find more coherent, given the distinction between functional properties and categorial bases and the existence of an exclusion problem. In any case, it is the approach that I will be dealing with for the rest of the paper.

## 2. *An explanatory exclusion problem*

One ontological conclusion that is usually extracted from this view is that functional properties do not really exist, given their causal inertness.<sup>5</sup> Rather, what there is is just categorial bases that can be picked up either directly or via a functional concept. Functional concepts, on the other hand, although non referential, can and, to a certain extent, should, be retained, because they are useful to us.

Let me explain this last point. Just as there is a causal exclusion problem there is, at the explanatory level, an explanatory exclusion problem induced by the following propositions:

- (a') an explanation couched in functional/dispositional terms is a good causal explanation for *explanandum e*;
- (b') an explanation couched in categorial terms is a good causal explanation for *explanandum e*;
- (c') there is no functional-categorial overdetermination.

Explanatory exclusion problems are governed by a principle of explanatory exclusion, analogous to the principle of causal exclusion, which states that we do not accept two independent and complete causal explanations for one *explanandum*. Now, the difference when we move from the ontological to the explanatory level is that functional and categorial explanations do not really exclude each other. As was first noted by Block (1990) and later developed by Jackson and Pettit (1990), a functional explana-

---

<sup>3</sup> Ehring takes properties to be classes of tropes, and the relation between categorial and functional properties that of parts to wholes: classes of realizing categorial tropes are parts (subclasses) of the class of functional tropes. His response to the *in virtue* question is that “(a) if causal homogeneity [of the subclasses] is high, the class as a whole has the causal power at issue, but (b) if low, the class as a whole lacks the causal power at issue” (p. 372). That is, only if realizers are sufficiently homogeneous in their causal powers, should the common causal powers be attributed to the functional property. This seems to imply that one-way functional properties, like *being a sleeping pill*—whose realizers share just the power to make people sleep—are inert, while multiple-way functional properties are causally efficacious. This approach, then, draws a (sociological and unfair) class distinction within what seems to be a homogeneous group, that of functional properties.

<sup>4</sup> See also Mumford (1998) for a development of a token-identity approach to dispositional causation.

<sup>5</sup> This is not a conclusion that I endorse. As I am about to suggest, it may be that (at least some) functional properties are causally efficacious.

tion can be seen as a partially informative explanation that is completed by the categorial explanation. The use of a functional predicate brings with it the information (a) that one of its categorial bases brought about the effects to be explained, excluding thus the intervention of any other property, and (b) that the causally efficacious property followed the “causal path” marked by the dispositional concept (there are cases where one categorial basis realises several different dispositions). This ingenious model of functional explanation, labelled ‘program explanation model’, thus, can reconcile, on the one hand, the putative fact of the causal inertness of dispositions with their illuminative use in explanations, and on the other, the existence of two different causal explanations for one *explanandum*. To repeat: the functional explanation can be considered as a pointer to the categorial explanation, i.e. the one that cites directly the cause of the effect to be explained.

It must be mentioned, following this thread, that this model of functional explanations has an advantage over the realist reading proposed by, among others, Jerry Fodor.<sup>6</sup> According to Fodor, the best explanation (if not the only) that can be produced for the success of functional generalizations is that functional generalizations track causal relations between functional properties. However, this makes him unable to give an adequate response both to the causal and to the explanatory exclusion problems, for his model implies massive overdetermination. The program explanation model, on the other hand, eludes the problems of exclusion, and at the same time it responds to the Fodorean challenge by producing an alternative explanation to the success of functional generalizations, and so is able to block the inference to the best explanation argument for functional realism.

Summing all this up, the program explanation model of functional explanations explains how it is possible to retain functional explanations, and to consider them genuine causal explanations, given that they inform about the cause of the *explanandum*. Moreover, it is also explained that in those (fairly common) cases where we do not know about the categorial base, or we are not interested in going very specific, we can be content with the functional explanation, but this can always be completed by the categorial explanation to which it implicitly refers.

It is easy to see that this model of functional explanations introduces an asymmetry between functional and categorial explanations, which results in a very different treatment of them. According to the program explanation model, functional explanations may be preferable in circumstances, but they are never complete. Categorial explanations, on the other hand, are complete: they do not require the information provided by the functional explanation, for they contain that information and more. This means that if we were to have a science made of categorial explanations, functional explanations would be redundant and dispensable. We could still be interested in using them, as we can be interested in using folk-psychological explanations even when having a perfectly scientific neuroscience, but the point is that no information about the

---

<sup>6</sup> See specially, Fodor (1990, 1997).

world would be missing if we did not use them. So, leaving at one side the pragmatic aspect and import of explanations, functional explanations, according to the program explanation model, happen to be entirely eliminable.

### 3. *Some missing explanations*

What I want to defend is that this model and its implications cannot account for a family of cases where dispositions are mentioned in causal explanations. It seems that in selective, or recruitment, processes, where functions are converted into teleofunctions, it is necessary to have an explanation pitched at the functional level, and that without such an explanation we would miss a crucial piece of information. Let me explain.

To take a typical example, a heart has the property of (being capable of) pumping blood as its teleological function; that is, it has been selected or designed to pump blood. This, according to Larry Wright's classical analysis, is to be analysed as (i) the heart pumps blood and (ii) it is there because it pumps blood. This second clause can well be understood as meaning that there is a causal story for the heart's being there (i.e. existing) in which the dispositional property of (being capable of) pumping blood appears as a causal factor. Were it not for the fact that it, or its ancestors, had that disposition, the state of affairs of the heart being there would not obtain.

This idea can be developed as follows. Suppose that we posed a causal exclusion problem between the blood pumping disposition and its categorial basis like this:

- (a) the pumping blood disposition was a causal factor for the heart's being there;
- (b) an intrinsic property of the heart was responsible for the heart's being there;
- (c) there was no overdetermination for the heart's being there.

According to what I propose, this problem is dissimilar to the usual exclusion problems, for it seems that it is the functional property that is in command this time, so to speak. The event to be explained is the presence of a heart. And the idea is that it "is there" because there has been a process of selection that has —so to say— respected it. Why? Because it did (or could do) some things that were an advantage for the creature possessing it, that is, because of what its properties, what they (their instantiation) caused. The process of selection did not take its intrinsic properties into account, only their effects. The heart is there because it belongs to the functional group of blood pumps; it is there *in virtue* of being a blood pump. The causal relevance of the intrinsic properties of the heart, then, depends on their being members of that functional group, of instantiating that functional property. Thus, (a) is true this time, and (b) is true only secondarily.

It is not difficult to grasp the difference between these cases and the usual cases of causal competition between dispositions and their bases (in which the former are defeated). A heart pumps blood not because it belongs to the group of blood pumps, not in virtue of being a blood pump; it pumps blood because it instantiates some (structural) properties. Because of this, it belongs to the functional group of blood

pumps, but this fact is, at this stage, causally inert. Now, my claim is that when selection processes take place, that is, when we are concerned with the “being there” of the heart, this functional fact becomes causally relevant.<sup>7</sup>

#### 4. *Functional properties as structuring causes*

An idea that comes close to what I am defending, and can give some flesh to it, is what Fred Dretske has to say about mental causation. Dretske (1988) has developed an ingenious account of mental causation that tries to make it compatible with physical causation by distinguishing between structuring and triggering causes. Take first a structure C with two kinds of properties, neurological on the one hand and backward-dispositional (carrying the information that F) on the other. Now, suppose that the objects instantiating F are a potential danger to the creature possessing C. Then, the neurological properties of C can be made to regularly cause or be causally connected to M, a fleeing movement. When this happens, says Dretske, the structure C has acquired the property of representing F, that is, it has acquired the function of indicating/carrying the information of F. Now, suppose the fleeing movement takes place: what is its cause? Surely the neurological properties of C were instantiated and triggered M, given that they were causally connected to it. But also, M happened because C had those indicative properties, which are responsible for the causal connection C-to-M. This second causal process is what Dretske calls ‘structuring causation’. Leaving some complexities aside, we can say that the idea is that some triggering causal processes depend for their instantiation on their structuring cause: If the property that is the structuring cause had not been instantiated, the property that triggers the effect would not have been connected to this effect, and hence would not have caused it.

Note that Dretske’s account, though in the line of Wright’s analysis, takes a different twist. As said above, what Dretske proposes is that mental properties are properties of the sort *having the function of carrying the information that F*. This, Wright would say, amounts to (i) C carries the information that F and (ii) C is there because it carries the information that F. However, Dretske tells us, C’s carrying the information that F does more than that: it is also responsible for C’s causing M. As this is not part of any analysis of having the function of indicating F, one may wonder how this element came in for the ride. The answer is that, as a matter of fact, what Dretske calls the recruitment process is double-faced: On the one hand, C acquires the function of indicating F, but on the other, it also acquires the function of causing M. That is, C is recruited to do both things, indicating F and causing M. That Dretske focuses on the

---

<sup>7</sup> Think also of a sleeping pill’s existence (maybe its “being there” in a chemist’s). Now what normally is the case is that if we have (a) the sleeping pill did this or that because it instantiates such and such intrinsic structural properties, and (b) the sleeping pill did this or that because it instantiates the property of making people sleep, we have to put a stick in (a) and discard (b) —no matter what relationship we discover between functional properties and their bases. But in these cases, where what we want to know is why the sleeping pill “is there”, we have to say that it is there *in virtue* of its being able to make people sleep, it is there because it belongs to the functional group of sleeping pills, not because it belongs to the group delimited by its categorial properties.

backward looking disposition instead of on the forward looking disposition is to do with his being a theory of the causal efficacy of the representational content of mental properties. But there is no reason, I think, why he could not complement his theory by looking at the causal role that C's forward looking dispositions play in the story.

Robert van Gulick (1993) has developed another account of mental causation that takes the complementary approach. That is, it is the forward looking dispositions of neuronal structures that make them cause what they cause. In van Gulick's view, a brain, being a dynamic self-preserving and self-reproducing organ, recruits physical structures according to their causal capacities, and because they have such causal capacities, and then *selectively activates* them. Take an injured brain, says van Gulick: if the neurones that have been injured were displaying an important function for the organism, the brain will *recruit* some other neurones so that they take on, if possible, the causal role that was temporarily lost.

This account, then, takes it that mental properties are properties of the sort having the function of having such and such a causal role (i.e. something is a mental property iff it has been recruited for a certain causal role). According to Wright's analysis this means: (i) mental properties fulfil such and such a causal role; (ii) they are "there" (instantiated) because they can fulfil that role. That is, there has been a process of selection (here recruitment) in which the dispositional property acted as a causal factor. So van Gulick's approach perfectly concords with Wright's views. And again, this means that the exclusion problem between dispositions and their bases is solved by making the causal relevance of the latter dependent on the causal relevance of the former. That is, if the objects instantiating mental properties had not instantiated the relevant dispositions, their intrinsic properties would not have caused the effects they caused. They would not have been recruited and their causal powers would not have been selectively activated.

Functional properties, then, can be considered (a) causes, or causal factors, of the result of recruitment or selection processes (i.e. X's "being there", or some of its properties being instantiated), and (b) structuring causes of triggering causation processes (i.e. X's properties causing E).

##### 5. *The role of functional explanations*

So far so good. However, I have been offering an ontological reading of the issue that may be unmotivated. In effect, what I have been defending is not my original claim that some functional explanations are indispensable, but the stronger claim that some functional properties are causally relevant. Now, there is a problem with this stronger claim, which, in a nutshell, is the following: suppose you accept that dispositional properties pass the exclusion problem with/against their categorial bases in the way proposed. Is there not another exclusion problem which they cannot pass so easily, namely, the one posed by the existence of one local cause for every effect? That is, dispositions, once they have passed the hurdle of categorial bases, have to face another, consisting in:

- (a) disposition D was causally relevant for effects E by virtue of being causally relevant for the instantiation of intrinsic property F that caused E;
- (b) F causing E, or being recruited/selected for causing E, is the end of a process in which only local causes figure.

Suppose that in effect a heart is there, and its intrinsic properties are instantiated, causing the pumping of blood, because the heart had the disposition to pump blood. Now, surely a number of local processes took place that had the heart's being there as the final effect. Which of these two possible causes are we going to admit? We have the dispositional story and the local causal story: which is correct? I would like to say<sup>8</sup> that dispositions can act as “catalysers” of local causal processes, leading them one way or another, and so that (a) and (b) can be compatible, but it is difficult to develop this intuition into a precise idea. However, what I think can be more easily defended is that you need the functional explanation to understand what is going on. That is, functional explanations respond to this question: why is the whole process taking place in the way it is taking place? Why do hearts stay, while a lot of other things are eliminated? We cannot answer those questions without mentioning dispositions.<sup>9</sup>

Of course, what I am saying about processes of selection, or natural processes of recruitment, applies also to processes of design. Take a watch; suppose you want to know how come this thing is there, doing what it does. You can be told the story of the watchmaker who gathered up the materials out of which the watch was made, placing everything in a certain order and making the watch work. But that would not be the end of the story. Part of what you still have to know is why the watchmaker gathered those materials, why he put them in that particular order, etc. And this is where dispositions enter —“look, this little dented wheel is capable of transmitting its movement to this other piece, something that the watchmaker knew, so he used it for this reason.”

This kind of causal explanations, then, invoke dispositions necessarily, and provide a type of information that would be missing if we were to use only categorial explanations. We would miss the explanation of why a whole process is taking place in the way it is taking place: there are cases where there is a response to that question, and we would miss it if we eschew functional explanations.

---

<sup>8</sup> As I did in Vicente (2002).

<sup>9</sup> Horgan (1991; 84) posed this problem to Dretske's approach to mental causation: “[I]n principle, the recruitment process also is explainable in more fundamental terms, and at lower theoretical levels the correlation will play no explanatory role.” If we stay at the explanatory level, I would take that ‘also’ at face value, and claim that both kinds of explanations can live together. Moreover, I would say that the one that makes use of the “correlation” (i.e. the backward-looking disposition) sheds light on the one couched in the “more fundamental terms”. I can see that my position is weak insofar as I do not explain what kind of processes these “dispositional explanations” track —if they are not causal, what are they?. Unfortunately, I have no answer at the moment. Notions such as *catalyser* or *structuring cause* are by all lights just place-holders. However, I take it it is just a fact that dispositions explain why things are selected.



The program explanation model for functional explanations solves the explanatory exclusion problem by claiming that there is only one kind of complete explanations, categorical explanations, and that functional explanations are pointers to them. But this cannot work in the case of processes of selection. If we put an exclusion problem like (a-b) above in explanatory terms,

- (a') a heart's being there is explained by the pumping blood disposition of the heart,
- (b') a heart's being there is explained by a number of causal local processes,

it seems that (a') makes sense, completes, or explains, (b'), the, let me put it in this way, local "blind" causal story.

As I say, there are some causal processes of the kind of (b) above whose occurrence can be explained, answers to "why questions" that go over and above the chain of local causes and explain why such a chain exists in the way it exists. This is not the usual case, of course: there is no such an explanation for the movement of the planets, or for the collapse of wave-functions. But there are also processes of selection in the world, and the occurrence of those processes can be explained only by mentioning dispositions, since it is in virtue of the dispositions a particular instantiates that it, or its intrinsic properties, is selected. I take it we could not really understand how the evolution of life is taking place if we did not take dispositions into account. We would have an adequate account of the order (even the causal order) of events and a good record of types of things that stay and the types of things that disappear. But my point is that we would miss the explanation of why those things are happening, that is, we would end up with the "noisy and furious" story.

#### *Acknowledgements:*

I have to thank those who attended the workshop *Dispositions and propensities in science* for their various and interesting comments. Many thanks also to Mauricio Suárez for inviting me to the workshop, for his encouragement and his (coming) suggestions, and to Nicholas Maxwell for his lively discussion and friendship. Research for this paper was funded by the BFF2002-03842 project of the Spanish Ministerio de Ciencia y Tecnología.

#### REFERENCES

- Block, N. (1990). "Can the Mind Change the World?", in Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam*. Cambridge: Cambridge University Press.
- Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.
- Ehring, D. (2003). "Part-Whole Physicalism and Mental Causation", *Synthese* 136: 359-388.
- Fodor, J.A. (1990). "Making Mind Matter More", in *A Theory of Content and Other Essays*. Cambridge: MIT Press.
- (1997). "Special Sciences: Still Autonomous after all these Years", *Philosophical Perspectives* 11: 149-163.
- Heil, J. (1999). "Multiple Realizability", *American Philosophical Quarterly* 36: 189-208.
- Horgan, T. (1991). "Actions, Reasons, and the Explanatory Role of Content", in B. McLaughlin, *Dretske and his Critics*. Cambridge, MA: MIT Press.

- Jackson, F. and Pettit, P. (1990). "Program Explanation: A General Perspective", *Analysis* 50: 107-117.
- Kim, J. (1993). *Supervenience and Mind*. Cambridge: Cambridge University Press.
- (1998). *Mind in a Physical World*. Cambridge, MA.: MIT Press.
- Mumford, S. (1998). *Dispositions*. Oxford University Press.
- Prior, E.W. (1985). *Dispositions*. Aberdeen University Press.
- Prior, E.W., Pargetter, R., and Jackson, F. (1982). "Three Theses about Dispositions", *American Philosophical Quarterly*, 19: 251-7.
- Van Gulick, R. (1993). "Who's in Charge Here? And Who's Doing all the Work?" in Heil and Mele (eds.), *Mental Causation*. Oxford: Clarendon Press.
- Vicente, A. (2002). "How Dispositions can be Causally Relevant", *Erkenntnis*, 56: 329-344.
- Wright, L. (1973). "Functions", *The Philosophical Review*, 82: 139-168.

**Agustín VICENTE** is Profesor Asociado (Assistant Professor) in Logic and Philosophy of Science at Valladolid University. He has worked on mental causation, and is currently also working on the relationships between language and thought. He has published papers on these two topics in journals such as *Erkenntnis*, *Mind and Language*, *Theoria* and *Minds and Machines*.

**ADDRESS:** Department of Philosophy, University of Valladolid. Email: agusvic@fyl.uva.es.