

## DISPOSITIONAL EXPLANATIONS OF BEHAVIOR

Rob Vanderbeeken  
Erik Weber  
*Ghent University*

**ABSTRACT:** If dispositions are conceived as properties of systems that refer to possible causal relations, dispositions can be used in singular causal explanations. By means of these dispositional explanations, we can explain behavior B of a system *x* by (i) referring to a situation of type S that triggered B, given that *x* has a disposition D to do B in S, or (ii) by referring to a disposition D of *x* to do B in S, given that *x* is in a situation of type S. Dispositional explanations are adequate and indispensable explanations: they can explain behavior B without explicitly referring to the underlying causal basis in *x* that constitutes a disposition to do B. Radical Behaviorist explanations are a sort of dispositional explanations, but the dispositional model is not restricted to these explanations. The dispositional model is compatible with, or can be applied to, several research programs.

*Key words:* behavior, disposition, causal relation, singular causal explanation, applications, Radical Behaviorism, Folk Psychology

The aim of this paper is to present a general model for the explanation of behavior that focuses on the relation between a system, for example, an object, an organism, a machine, a person, and its present environment. The central feature of this model is the notion “disposition.” In our view, dispositions are stable properties of systems that specify how a system will behave in certain situations. Put differently, dispositions are properties that refer to types of causal relations between a system and situations of a certain type. Once we know that a system has a certain disposition, we know how this system will behave in certain situations, due to the presence of an unspecified underlying set of categorical properties that functions as a set of internal causal factors. Explanations that aim to explain behavior by referring to the absence or presence of external causal factors must presuppose the presence of some internal causal factors (sometimes called “hidden variables” or “mechanisms”). The importance of dispositions for explanations is that they make it possible to explain behavior by means of the relation of a system with its environment alone, that is, without explicit reference to the internal causal

---

**AUTHORS' NOTE:** The research of this paper was supported by the Fund for Scientific Research Flanders—Belgium. We thank the editor and the referees of this journal, as well as Anita Craig, Jeroen Van Bouwel, Petri Ylikoski, and Maarten Van Dyck for their valuable commentaries on previous versions of this paper. We also thank the attendants of the 2nd Flemish-Dutch congress in Leusden, The Netherlands in June 2001, and of the annual ESPP meeting in Fribourg, Switzerland, August 2001 for their fruitful comments.

Please address all correspondence to Rob Vanderbeeken or Erik Weber, Centre for Logic and Philosophy of Science, Ghent University, Blandijnberg 2, 9000 Gent, Belgium. Email: Robrecht.Vanderbeeken@rug.ac.be, Erik.Weber@rug.ac.be.

factors. Nevertheless, in several cases dispositional explanations fall short. Some requests for explanations cannot be fulfilled by means of information about present environmental factors. These requests aim for information about the causal factors that structured the behavioral abilities of a system. In these cases we have two options: we can try to specify the categorical basis of the respective disposition, for example, in mental or neurophysiological terms, or we can try to explain the disposition itself by means of the personal, social, or biological history that shaped the system.

This paper has two parts. In the first we discuss how dispositions and dispositional explanations should be conceived. Although scientists and philosophers use dispositional explanations frequently, the use of dispositions in explanations is generally considered problematic from a theoretical point of view. There are two specific reasons for the skepticism. Firstly, it is not clear what dispositions stand for and how they can be causally relevant for explanations. For this reason some philosophers, for example, Quine (1974), have argued that the use of dispositions is proto-scientific: when science progresses, dispositions will and should be completely eliminated from explanation and be replaced by “respectable” categorical explanations. We will argue, however, that dispositional explanations are adequate and indispensable. Secondly, it is not clear how the format of dispositional explanations should be understood. Hempel (1965) argued that dispositional explanations are a kind of Deductive-Nomological (D-N) explanations, which turned out to be a defective conception since the law-statements in these D-N explanations are purely analytical and not empirical. We will show that dispositional explanations are not D-N explanations but singular causal explanations.

In the second part we discuss the applicability of the dispositional model. What does the dispositional model prove? The dispositional model does not provide a new way to do behavioral science, one that is different from the current mode. The importance of the dispositional model is that it allows for the explanation of behavior by referring to the relation between a system and its present environment. It makes explicit the strategy that is used by different sorts of research programs that aim to explain behavior by referring to present external factors, and it also shows how different sorts of research programs can be utilized in order to fulfill this task. By way of example: Radical Behaviorist explanations fit the dispositional model. Radical Behaviorist explanations aim to explain respondent or operant behavior by referring to external causal factors. These external causal factors can only be conceived as stimuli for a certain response R, if a system has a disposition D to do R when it is in a situation S. According to Radical Behaviorist explanations, a system has a disposition D to do R, due to the presence of a set of unspecified internal causal factors that are the result of a history of reinforcement. The dispositional model is not restricted to Radical Behavioristic explanations, however. There are several strategies available to attribute a disposition to a system. One strategy is induction from previously similar behavior; another is to refer to historical influences of which it can be said that they established a disposition; a third is ahistorical. Although Skinner’s

## DISPOSITIONAL EXPLANATIONS

Radical Behaviorism emphasizes a system's history as the determinant of its behavior, it focuses on a history of reinforcement neglecting a whole area of important historical influences. Unlike Radical Behaviorist explanations, dispositional explanations are not restricted to the reinforcement-thesis. This makes it possible to use dispositions in explanations that are not focused on a particular history of reinforcement. The ahistorical strategy means that we can attribute a disposition to a system based on information about internal causal factors, that is, "inside-the-head" information. The historical strategy relies on an explanation of the disposition itself, while the ahistorical strategy relies on a specification of a disposition. In short, the fact that dispositions, as we conceive them, can be specified by making explicit their internal underlying causal bases or can be further explained by referring to historical facts that structured these dispositions, makes the dispositional model compatible with several research programs, for example, Radical Behaviorism, Theoretical Behaviorism, Cognitive Science, Neuroscience, Evolutionary Psychology, Folk Psychology. Of course, it might turn out that we could give several specifications or further explanations of a disposition. That is to say, the dispositional model is meant to provide a theoretical tool, a means to apply several sorts of research programs, not a means to unify research programs or choose between competing specifications or explanations.

In what follows, we first provide a definition of dispositions. Briefly, we view a disposition as more than a mere description, a shorthand characterization of an empirical generalization: dispositions are stable properties that tell us how systems will behave. Secondly, we discuss why dispositions are often explanatory. We argue that dispositions, as properties that refer to possible causal relations, are not causally significant themselves. They are causally relevant for explanations since they tell us where the causes can be found: they inform us about the causal powers of a system and they inform us about situations that will trigger a typical manifestation. Thirdly, we discuss the format of dispositional explanations. After we showed why a Deductive-Nomological format of explanations using dispositions as initial conditions is problematic, we mention an alternative, singular format of dispositional explanations that can be used in two opposite directions. If we know that a system has a disposition *D* we can explain or predict a manifestation *M* by means of a triggering situation *S*. On the other hand, if a manifestation *M* occurred and we know that a system had (or still has) a disposition *D* we can infer which type of situation was the case. In the final section of this paper we discuss the applicability of dispositional explanations. We show that Radical Behaviorist explanations are dispositional explanations, and discuss three strategies for attributing dispositions: induction from similar behavior in other situations, inference based on information of the social, biological or personal history, and inference of dispositions based on information of the mental (intentional/cognitive) or natural constitution of the internal causal basis. To end with, we discuss two restrictions of the dispositional model.

## Dispositions

### *Defining Dispositions*

Dispositions are capacities, or call them powers, abilities, tendencies, or propensities of systems. In our view, dispositions are not mere descriptions, stating empirical generalizations (e.g., Carnap, 1939; Hempel, 1965). Dispositions are stable properties of a system that describe how a system acts or interacts, given a situation of certain triggering conditions. *Contra* Armstrong (1969, 1996) dispositions cannot be identified with the underlying causal basis to which they refer. Systems simply have two equally important sorts of properties: dispositional properties and categorical properties. Dispositional properties of a system state that a certain type of causal relation will take place when the system is in situations of type S. Such a relation covers three elements: a categorical basis (internal causal factors), a situation of type S (containing triggering causes), and a typical manifestation (the expected effect). Categorical properties, on the other hand, describe the discernible features of a system. Categorical properties describe how things are, for example, the chemical structure, the shape, volume, or color of a thing, while dispositional properties describe how things behave, for example, they bend, bounce, break, shatter or split when dropped. We define dispositions as follows:

- [DD] D is a disposition of system  $x$ , if:
- (i)  $x$  has an underlying property or property-complex U,
  - (ii) U together with a situation of type S is a sufficient cause for manifestation M.

Concerning (ii): when U together with a situation of type S is a sufficient cause for M, we know that if U and S then necessarily M. This means that antidotes or interventions are ruled out: system  $x$  should lack other properties that obstruct or compensate U (unless  $x$  contains antidotes to antidotes or properties that intervene on intervening properties). The same goes for every situation of type S. If there is only one S that can trigger M, then S is a necessary part of a sufficient cause of M. (In these cases a disposition is called a single-track disposition.) If U together with a situation of type S is the only possible sufficient cause for M, U together with S is a sufficient and necessary cause for M.

Concerning (i): we give three reasons why dispositions are not identical with their categorical bases. (1) Swamping: the effect of the underlying causal basis of a disposition can be compensated by other categorical properties. In these cases the disposition remains absent while the respective underlying causal basis is present. (2) Variable or multiple realizability: dispositions can be constituted by different sorts of categorical bases. (3) Multiple constitutability: a single basis can cause different manifestations in different situations, so a single base can constitute several dispositions. This shows that a disposition depends on a causal basis, but is not identical with it. A disposition has two unique traits.

## DISPOSITIONAL EXPLANATIONS

- (a) Dispositions state that an underlying set of categorical properties U, that is, a particular set of internal causal factors, is present without making this cause explicit.

The fact that U is not specified does not mean that it is not specifiable. For some dispositions, U can even be specified in general, that is, one can give an exhaustive list of all possible sets of internal causal factors of different sorts of systems that can constitute this disposition. For some dispositions, on the other hand, the underlying causal basis remains unknown. This does not imply, however, that there is no such thing as a causal basis in these cases: dispositional properties describe action tendencies of systems and this requires the existence of a responsible causal basis. Since dispositions are not identical with their causal basis they cannot be considered as causes themselves. Dispositions are causally inert. Since dispositions presuppose a causal basis that can provoke a manifestation M and since they are distinct from their categorical basis, there must be some sort of causal basis whenever we speak of dispositions.<sup>1</sup>

- (b) Dispositions give information about possible situations that will trigger M.

Dispositions do not tell us which causal basis is to be found in a system, it tells us that there is some sort of causal basis that, together with certain triggering situations, will bring about typical behavior. This makes dispositions an excellent instrument for the study of the relation between behavior and the environment of a system. Namely, dispositions enable us to take some sort of internal structure of a system for granted and, based on this assumption, to predict the actions of a system when confronted with a specific type of situation. Furthermore, dispositions enable us to retrodict: if there has been a manifestation M, and we know that the respective system has (or had) a disposition D, we know that a situation of type S occurred. The fact that we do not know all the instances of situations of type S is no objection: knowledge of one particular situation is enough. In sum, the upshot of a disposition is that it enables us to assume the presence of a sort of causal basis—which makes dispositions *prima facie* causes. So, although dispositions are not causally significant, they can be causally relevant: they indicate where the causes are to be found and what effect is to be expected.<sup>2</sup>

---

<sup>1</sup> In the case of so-called fundamental dispositions (e.g., mass, charge), it is argued (e.g., Ellis, 1999; Ellis & Lierse, 1994; Harré, 1970; Harré & Madden, 1975) that no causal basis can ever be found. It remains to be seen, however, if these properties are genuine dispositions or rather categorical properties that can only be described in dispositional terms. Moreover, explanations of animal or human behavior will not appeal to such dispositions. This issue only concerns a restricted set of cases. For our purposes, this issue can be left aside.

<sup>2</sup> Conceiving dispositions as powers, like Harré (1970), Harré and Madden (1975), Ellis (1999), and Ellis and Lierse (1994) do, is interesting since it stresses the causal aspects of dispositions. Nevertheless, these views take dispositions as primitives and are meant to argue for their causal significance. We believe it is worthwhile not to take dispositions as primitives but to assume they are properties that refer to types of possible causal relations and that they have an underlying causal basis: it avoids problems of the power-ontology (cf. Armstrong, 1999), it enables us to specify some

*The Virtue of Virtutes Dormitivae*

Dispositions are causally relevant: they are not causes but indicate where the causes are to be found. Being properties that refer to types of possible causal relations, descriptions of dispositions are in a sense connected by definition to descriptions of their manifestations. When we say that glass is fragile, we say that it breaks when struck or twisted. When we say that opium has a *virtus dormitiva*, we say that people who ingest it will get sleepy. So, in cases where we answer that a glass breaks because it is fragile or that opium causes sleep because opium has a *virtus dormitiva*, in response to questions regarding why a glass broke when struck or why opium caused sleep, our answers do not always provide any new information. If we already know what glass and opium stand for, these answers can be derived from the questions without any extra empirical information. This is the reason why dispositions often are considered useless for explanations and that some philosophers, following Quine (1974, pp. 8-15) hope that, when science progresses, *virtutes dormitivae* eventually will be completely eliminated from explanation and be replaced with “respectable” categorical explanations.

We believe, on the contrary, that dispositions are indispensable and adequate instruments for explanation. Concerning the indispensability: we often know how systems behave in certain situations, without knowing the internal mechanisms that provoke this behavior. In these cases, we can only use dispositional explanations. Moreover, in several cases we are only interested in the behavioral abilities of a system, not in the internal mechanisms that constitute these abilities. In these cases, the dispositional model is more economical than categorical explanations: it is simple, straightforward, and tells us all we need. Concerning the adequacy: explanations that refer to dispositions are not vacuous or false, but they can be explanatory empty relative to the interest and the background knowledge of the explanation-seeking person. Explanation is a goal-directed human activity: explanations are answers to why-questions (cf. Lipton, 1991; Van Fraassen, 1980). Answers to explanation-seeking why-questions, invoking only dispositions, are explanatory in several cases. For instance, if one does not know that glass is brittle and, thus, has no idea what glass is really like, and if one wonders why a glass object is broken, an answer stating that glass is brittle; that it breaks when dropped, is clearly explanatory. Although this explanation does not specify an independent event that stands to the breaking of the glass as a cause to an effect, the explanation asserts a general hypothesis by means of the dispositional property, roughly stating that at any time when an object of this class is dropped, or when it is twisted by a person or struck hard enough by any physical object, it will, due to some unspecified underlying properties, fly into fragments rather than bend, bounce, etc. So, the disposition tells us that every object made of glass will break, unlike metal or plastic objects of the same size and shape. Furthermore, when we know that an object is fragile, we know that it need not be the case that this object, when broken, had to have some invisible fractures already. Moreover, after obtaining information

---

dispositions in terms of categorical properties, and it explains why dispositions can be causally relevant while being causally inert.

## DISPOSITIONAL EXPLANATIONS

about a disposition of an object, we know whether or not the object was in a situation of S, and we know what kind of situations can trigger the disposition. For example, if a glass is broken, we know that it was struck, twisted, or dropped and not just touched or blown at. Likewise, if we do not know that opium causes sleep, and if we wonder why a person, who ingested opium among other things, fell asleep rather than stay awake, answering that opium has a *virtus dormitiva*, that is, that opium causes sleep when ingested, is also clearly explanatory. We now know that every person at any time will fall asleep soon after ingesting opium. We know that it not just counts for this person, that it will happen again when someone ingests opium (rather than touching it or looking at it, etc.) and that it was the opium and not some other substance that caused the sleeping.

### Dispositional Explanations

#### *Problems With the D-N Account*

According to the Deductive-Nomological model of explanation (Hempel & Oppenheim, 1948/1965)—the “received” view—explanations are arguments. We have a D-N explanation if the following conditions of adequacy are fulfilled (1965, p. 247):

- (C1) the explanandum must be a logical consequence of the explanans, i.e., the explanandum must be logically deducible from the statements that form the explanans;
- (C2) the explanans must contain at least one statement that expresses a general empirical law;
- (C3) the statements constituting the explanans must have empirical content, i.e., they must be capable, at least in principle, of being tested by experiment or through observation;
- (C4) the statements constituting the explanans must be true.

Hempel had high expectations of behaviorism. Unlike some behaviorists, however, Hempel recognized the importance of explanations in psychological terms.<sup>3</sup> But still, in his view, S-R Laws play a significant role in scientific explanations of behavior. Most behaviorists, on the other hand, think that scientific explanations are and need to be covering law explanations. In his (1965, p. 462), Hempel gives the following format of dispositional explanations:

- C<sub>1</sub>: *a* is in a situation of type S at *t*.
- C<sub>2</sub>: *a* has property D
- L: *Any x with property D will, in a situation of type S, behave in manner R*
- E: *a* behaves in manner R at *t*.

---

<sup>3</sup> Hempel corrected his view during the years. Hempel (1949, p. 18): “All psychological statements that are meaningful [. . .] are translatable into statements that do not involve psychological concepts.” Hempel (1966, p. 110): “In order to characterize [. . .] behavioral patterns, propensities, or capacities [. . .] we need not only suitable behavioristic vocabulary, but psychological terms as well.”

This account, however, faces similar difficulties as Hempel's account of Rational Explanations (note that this account is also based on a disposition, namely that of being a "rational agent"). Hempel tries to shape both forms of explanations in such a way that they fit the covering law model, but in both cases the stated law is analytical rather than empirical: the L-statement is an analytical true statement that can be derived from the antecedent conditions, that is, from  $C_2$  alone. This means that the law-statement has a purely formal, inferential function and has no explanatory import at all. Hence, dispositional explanations are not proper D-N explanations since conditions (C2) and (C3) remain unfulfilled.<sup>4</sup> This is very interesting. How is it possible that the explanandum can be deduced without an empirical law? The answer is obvious: the antecedents  $C_1$  and  $C_2$  are sufficient since the disposition takes over the role of an empirical law. Even Hempel (1965, p. 458) notes that dispositional statements are lawlike statements that differ from general laws because they also mention a particular individual that has the dispositional tendency. In other words, dispositions are general tendencies conceptualized as properties of individual systems. When an explanation refers to a disposition, this explanation no longer requires a law-statement.

The defective D-N format of dispositional explanations turns out to be a perfect example of how an alternative, singular model of explanations is possible. Singular explanations are not arguments but singular causal sentences. In the previous section, we discussed why dispositions, being properties that refer to types of possible causal relations of a system, are causally relevant. We have also seen that dispositions can be described in a singular causal sentence: (cf. [DD(ii)]) "U together with a situation of type S is a sufficient cause for the manifestation M." This shows how dispositions can be used in causal explanations using singular sentences instead of arguments. Such an alternative singular approach will be important for at least the following reasons:

- (i) The fact that the D-N format of dispositional explanations is defective, makes a singular, dispositional approach indispensable.
- (ii) A singular approach using sentences can avoid several classical problems of the D-N account (cf. Cartwright, 1983; Salmon, 1989).
- (iii) A singular approach using dispositional terms in sentences can be a successful alternative to the D-N account, especially when it comes to *ceteris paribus* laws.

Concerning (iii), Lipton (1999) sums up the problems of *cp* laws and concludes, rightly we believe, that "dispositions can come to the rescue." We need *cp* laws to bridge the gap between lawlike and accidental generalizations. Dispositions can ground *cp* laws and can explain why *cp* laws sometimes fail, for example, in cases

---

<sup>4</sup> Hempel (1965, pp. 159-162) argues that the L-statement is significant in cases of so-called broad dispositions since it points out one way of the variety of symptomatic ways a broadly disposition can manifest itself in. Nonetheless, we do not need an L-statement for such a specification. If we know what the broad disposition stands for, and given the fact that the antecedent circumstances mention a particular situation, we know which corresponding particular manifestation can or must be inferred.



## DISPOSITIONAL EXPLANATIONS

where categorical properties are obstructed or compensated. Dispositions provide an alternative to *cp* laws in explanations because dispositions imply a necessity (or support counterfactuals), and they have the same flexibility. Woodward (2000) shows us how this can be done. He argues that the important feature of laws concerning explanation is their invariance. A generalization describing a relationship between two or more variables is invariant if it remains stable as various other conditions change. Note that such a generalization can only be explanatory if it is not invariant under all circumstances. Once we are aware of this central notion, Woodward argues, we can try to distinguish degrees of invariance. Depending on the sensibility for changes or interventions, a tendency will be closer to an accidental or a lawlike regularity. We believe dispositions can come in handy here, because dispositions are located in particular systems, unlike laws that are supposed to rule in the whole universe. Dispositions are invariant tendencies of systems relative to particular situations: some systems can have a disposition to do B in S, while they do B' in S'. Other systems have the disposition to do B only in S' and still others lack the disposition to B completely.

### *A Singular Two-Way Model*

Hempel believed that D-N explanations using dispositions are not vacuous, even though dispositions are considered to be *virtutes dormitivae*. The reason for this is that the explanans contains more than just a premise stating the disposition. It also must include a law-statement and a premise mentioning the presence of a situation S. As we have seen, we can do without the law-statement but not without one of both these premises. This means that we can explain the behavior of a system, for example, the shattering of a window panel, by pointing out that it was in a situation S, for example, it was struck by a stone. But as Ryle (1949, p. 88) already noticed, we often use dispositions in explanations in a different sense also: "We ask why the glass shattered when struck by the stone and we get the answer it was because the glass was brittle." This shows, we believe, two important aspects of dispositional explanations: in order to explain a manifestation M, we can only refer to a situation S if we know that a system *x* has a disposition D to do M in S, that is, it need be the case that *x* contains some "hidden variables" or "mechanisms" that make it possible that S can trigger M (cf. Staddon on the importance of "internal states," 2001, chap. 7). Furthermore, we can only refer to a disposition D in an explanation if we know that *x* is in a situation of type S. If not, referring to a disposition D is useless. This means that we can use dispositional explanations in two opposite directions: (i) when we know that a system *x* was in a situation of type S, we can explain a manifestation M by stating that this system *x* has a disposition D to do M and that D gets triggered when *x* is in a situation of type S, and (ii) when we know that a system *x* has (or had) a disposition D to do M in a situation of type S, we can explain and even predict a manifestation M by referring to the occurrence of a situation of type S. If, however, the explanation-seeking why-question does not make it clear whether or not the questioner already knows that system *x* is in situation of type S, or that system *x* has a disposition D,

our explanation should contain both (i) and (ii). Consider the following types of explanation-seeking why-questions:

- (a) Why is it that  $x$  does B, rather than B' at  $t$ ?  
(*Property-contrast*)
- (b) Why is it that  $x$  does B at time  $t$  but not at  $t'$ ?  
(*Time-contrast*)
- (c) Why is it that  $x$  does B, while  $y$  does B'?  
(*Object-contrast*)

Question (a) can be answered by referring to the fact that  $x$  was in a situation of type S at  $t$ , given the fact that  $x$  has a disposition D to do B in S (and not B'), or by referring to the fact that  $x$  has a disposition D to do B in S, given the fact that  $x$  was in a situation of type S at  $t$ . Question (b) can be answered by referring to the fact that  $x$  was in a situation of type S at  $t$  and not at  $t'$ , given the fact that  $x$  has a disposition D to do B in S, or by referring to the fact that  $x$  has a disposition D to do B in S, given the fact that  $x$  was in a situation of type S at  $t$  and not at  $t'$ . Question (c) can be answered by referring to the fact that  $x$  and not  $y$  was in a situation of type S, given the fact that both  $x$  and  $y$  have a disposition D to do B in S, or by referring to the fact that  $x$  and not  $y$  has a disposition D to do B in S, given the fact that  $x$  and  $y$  are in a situation of type S. Each of these cases show that the dispositional model of explanations is a two-way model that can be applied to explanation seeking questions that include Property-, Object-, or Time-contrasts. For reasons of economy, we will only fully spell out our model using a P-contrast:

- (i) Why is it that  $x$  does B, rather than B' when confronted with a situation S?

*Description of the actual world*

- (1)  $x$  was confronted with situation S.
- (2) Systems with disposition D do B when confronted with situation S.
- (3) Systems with disposition  $D_1', \dots, D_n'$  do  $B_1', \dots, B_n'$  in S.
- (4) Dispositions  $D_1', \dots, D_n'$  are incompatible with D.

*The Differences*

- (E)  $x$  has a disposition D (rather than one of the dispositions  $D_1', \dots, D_n'$ )

- (ii) Why is it that  $x$ , with disposition D, does B, rather than B'?

*Description of the actual world*

- (1)  $x$  has disposition D.
- (2) Systems with disposition D do B when confronted with situation S.
- (3) Systems with disposition D do B' when confronted with S'.
- (4) Situations S and S' are incompatible.

## DISPOSITIONAL EXPLANATIONS

### *The Differences*

(E)  $x$  was confronted with situation S (rather than S').

(E) is explanatory in both cases, since we know that D presupposes an underlying causal basis U and that U together with S is a sufficient cause for M (cf. [DD]). Note that, in both cases, (E) cannot be derived from the description of the actual world alone. Namely, in (i) it can be that  $x$  does B rather than B' without having the disposition D, that is, without it being the case that  $x$  will do B whenever in a situation S (cf. indeterminate behavior). In (ii), it is possible that  $x$  was in a situation S'', and  $x$  can do B in S'', for example, because of disposition D'', or even by accident. This means that (E) is explanatory in both cases if, of course, the background knowledge of the person who requests these explanations includes (1)-(4) and lacks (E).

This brings us to a possible second step in the process of explanation: if the questioner already knows (E), the provided explanation will not satisfy him or her. What the questioner really wants to know, then, can be specified by a derived question of the original why-question: "Why does  $x$  have D (rather than another disposition or rather than plainly not-D)?" An answer to such a question must invoke either a specification of the disposition or a further explanation, that is, an explanation of the disposition itself. Here, we can use different strategies for attributing dispositions (cf. *infra*) in the inverse direction: we can try to specify the disposition by making explicit the internal basis that constitutes the disposition in intentional, cognitive, theoretical, or naturalistic terms, or we can try to explain the disposition by pointing at remote causes, that is, the historical influences that structured the disposition. The following figure gives an outline of the structure of dispositional explanations.

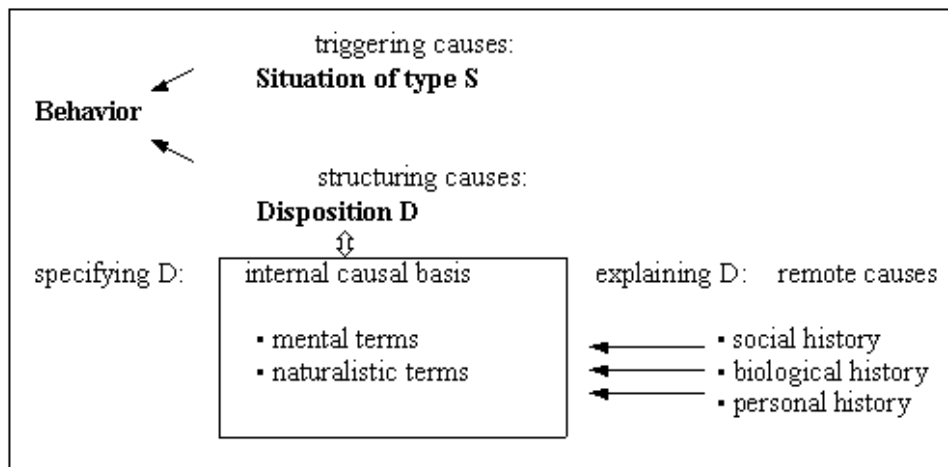


Figure 1: An outline of the structure of dispositional explanations.

## Applicability

### *Radical Behaviorism & Dispositional Explanations*

Radical Behaviorist explanations are dispositional explanations: they explain behavior B of a system  $x$  (i) by referring to external causal factors, that is, a situation of type S including certain stimuli that triggered this behavior, given that  $x$  has or had a disposition D to do B in S, or (ii) by referring to a disposition D of  $x$  that explains why  $x$  behaves as it does in a situation of type S which includes certain stimuli, given that  $x$  is in a situation of S. The central feature of Radical Behaviorist explanations is that they can provide an answer to the question: “Why does  $x$  have D (rather than another disposition or rather than plainly not-D)?” without explicitly appealing to the internal causal factors (i.e., in mental, theoretical, or neurophysiological terms) of systems.<sup>5</sup> Radical Behaviorists prefer to ignore or shortcut the internal causal factors and instead explain the behavior B by referring to the history of reinforcement. B.F. Skinner convincingly argued that some behavior can be explained by referring to a “three-term contingency”: Stimuli, Response and Reinforcement. This strategy consists in an explanatory jump, citing a causal influence that leaps over a temporal gap with no reference to intermediate links bridging cause and effect.

According to Skinner (1969), we can make a distinction between two sorts of conditioning: respondent (or Pavlovian) and operant (or Instrumental) conditioning. In the case of Pavlovian conditioning, behavior (e.g., the salivation of a dog) is shaped by means of a repeated confrontation of the system with paired unconditioned stimuli (e.g., food) and conditioned stimuli (e.g., the ringing of a bell). There must be short delays between the unconditioned and the conditioned stimuli, and unconditioned stimuli unpaired with conditional stimuli must be absent. Pavlovian conditioning is based on stimuli reinforcement: the behavior that is the result of Pavlovian conditioning, called respondent behavior, is said to be elicited by a stimulus. In the case of operant conditioning, behavior (e.g., a rat pressing a lever) is shaped by means of the positive or negative reinforcement of the consequences of this behavior (e.g., food forthcoming after pressing the lever). Operant conditioning is based on response reinforcement: the behavior that is the result of operant conditioning is said to be emitted. This behavior is called operant behavior because it operates on the environment and is guided by its consequences. Dispositional explanations are not restricted to Radical Behaviorist explanations, however. The dispositional model is a three-term model, that is, Situation-Disposition-Manifestation, in which the middle term fills in the explanatory missing link between situations and manifestations. A disposition of a system can be established in several ways and reinforcement is one of them. Moreover, in order to use dispositions in an explanation, we do not necessarily need to explain

---

<sup>5</sup> There are exceptions however. Nontraditional Radical Behaviorists like theoretical behaviorists, for example, Staddon J. E. R. (1973, 2001), try to capture the hidden variables within a system in theoretical terms, by means of parsimonious black-box models.

## DISPOSITIONAL EXPLANATIONS

how a disposition is established. As we will see in the next session, there are several good strategies to attribute a disposition to a system.

### *Radical Behaviorism and Beyond*

One possible strategy to attribute a disposition to a system is induction from past behavior: when a system always seems to behave in a certain way when confronted with a situation of a certain type we can generate empirical generalizations. The disadvantages of this strategy are, for instance, that dispositions that only manifest themselves once, or dispositions that did not manifest themselves before, cannot be captured. A second type of strategy concerns hypotheses about the establishing of a disposition. These strategies are historical in character: a disposition is attributed to a system based on beliefs about historical influences that shaped this disposition. Such strategies focus on the past, looking for remote causes that could have structured this system in one or another way. Although Radical Behaviorism emphasizes a system's history as the determinant of its behavior, it unilaterally focuses on a history of reinforcement, neglecting a whole area of important historical influences, albeit biological, social, or personal. The reinforcement-thesis enables us to explain why a system has learned to behave in a certain way, how a system was able to adapt itself to its environment. Put differently, the reinforcement-thesis is a tool to explain how the personal history of a particular system shaped this system's (compulsive or volitional) behavioral abilities or tendencies. Nevertheless, some historical influences in the personal history of a system have nothing to do with reinforcement (or at least, it is not clear how to get reinforcement into the picture here). For instance, some historical facts can have a major impact on the psychological constitution of a system, for example, traumatic experiences, or on the physical constitution of a system, for example, a viral infection. Moreover we can also attribute dispositions to a system based on information about its social or biological history. In these cases of social or biological inheritance, we cannot, or do not need to, presuppose the reinforcement-thesis. Concerning the biological history we can refer to inborn physical or psychological traits of a system that are the result of natural or sexual selection. Concerning the social history, we can refer to the constraints of the social surroundings, for example, social rules, habits, resources that prevented the system to develop other skills, practices, attitudes, etc.

In sum, concerning human beings, we can learn a lot of the social, the biological as well as the personal history of a person, without relying on Radical Behaviorism. Research programs that seek to uncover the relation between the past social background (e.g., social class of parents, social position of an individual) and actual behavior are, for example, Generic Structuralism (i.e., Bourdieu, 1990) and Neofunctionalism (e.g., Alexander, 1998). Evolutionary Psychology (e.g., Barkow, Cosmides, & Tooby, 1992), on the other hand, studies the relation between behavior and the biological background of person in order to capture, for instance, sex or age related dispositions. Dispositions related to the personal

history (e.g., childhood, education, life experience) is the subject of Psychoanalysis or Personality Psychologists (e.g., Hogan, Johnson, & Briggs, 1997).

A third type of strategy to attribute a disposition to a system is to focus on the information about the internal mechanisms of a system, in order to obtain information concerning the absence or presence of the underlying causal basis of a disposition. These strategies are ahistorical: we focus on information of the physical or mental constitution of a system itself. Four possible options can be taken here. (i) We can try to infer some dispositional properties from the mental/intentional knowledge of a system, concerning beliefs, desires, intentions, used decision rules and the like, obtained by observation, communication, or simulation. Research programs that use this strategy are, for example, Symbolic Interactionism (e.g., Mead, 1934/1970), Ethnomethodology (Garfinkel, 1967), Simulation Theory (e.g., Gordon, 1986), and Rational Choice Theory (e.g., Elster, 1986). (ii) Another strategy involves focusing on the presupposed mental/cognitive functioning of a system in terms of “memory,” “representation,” “expectation,” “causal inference,” etc. Research programs that use this strategy are for example, the model-approach of Cognitive Psychology (e.g., Anderson, 1985) and Theory-Theory (e.g., Gopnik & Wellman, 1994). (iii) We can try to capture the internal states of a system based on hypotheses about the hidden variables within a system by means of the invention and testing of parsimonious black-box models. Theoretical Behaviorism (i.e., Staddon, 1973, 2000, 2001) applies this option. (iv) We can focus univocally on the physical side of the internal basis, relying on neuroscientific research on the relation between brain capacity or brain pathology and mental behavior or bodily movements. For most behavior, however, it is likely that the neurophysiological basis is extremely complex and may never be completely unravelled, which makes it hard to infer the existence of dispositions related to behavior.

From the above it is clear that there are three available strategies for the attribution of a disposition to a system: induction from similar behavior in other situations, inference based on information of the social, biological or personal history, and inference of dispositions based on information of the mental or natural constitution of the internal causal basis. One objection might be that using all sorts of research programs is problematic because the claims of each are incompatible. However, using relevant and reliable insights of a research program need not imply a commitment to every theoretical assumption and allied philosophical claims. We believe that, in these times of explanatory pluralism, a theory of explanation should not commit itself to one particular point of view. A variety of strategies, using all sorts of information, enable us to attribute more dispositions, to attribute them more easily, and to obtain specific information about how dispositions are brought about and what they stand for. Allowing for more strategies improves the reliability and the justification of the attributions as well, since strategies can be used to verify one another.

*The Limits of the Dispositional Model*

The dispositional model is restricted in two ways. A first restriction concerns the nature of the subject matter. A dispositional explanation requires that we can point out stable relations between a systems behavior and its environment. If we do not have such information, we simply cannot use dispositional explanations. Note that, in several cases there are good reasons to believe that there are no such stable relations. Some systems react completely different when confronted with the same situation. Such behavior can be called indeterminate concerning present external factors: we cannot mark a strict relation between a system and the present environment. In order to deal with these cases, we can consider at least two options. One option might be to introduce probabilities, another searching for weak patterns.<sup>6</sup> However, we believe that an externalist approach reaches its limits here: it can be very successful for explaining and predicting, if only there is a clear relation with the environment. If not, like in cases of external indeterminacy, the behavioral output depends highly on internal determinants. This is where an internalist approach should take over. To say something more about the problematic cases: we can distinguish at least two possible sorts of indeterminate behavior. (i) Some indeterminate behavior is largely internally stimulated behavior. Here, the relation with the present environment is minimal. As an example, take writing a letter. Although an environment can influence how and what one writes and even if one continues writing or not, it is clear, we believe, that in most cases the impact of the environment is irrelevant to the output. (ii) For some indeterminate behavior, the relation with the present environment is very complex and plural. Here, behavior is not a response to a single set of stimuli but is rather constituted by means of a complex interaction over a period of time. As an example, take playing the piano. We can try to represent such behavior by means of serials of external causal factors and the respective responses, but it will be hard to capture the relation between such behavior and the environment (cf. Lashley, 1951). For instance: how to divide such an action in parts? A keystroke, a bar, a phrase, a movement, etc.? If these parts are responses to external stimuli, what are they? The sight of a note, phrase, . . . ? The sound of a tone, . . . ?

A second restriction on dispositional explanations concerns the interests that motivate requests for explanations of behavior. Requests for explanations can have theoretical or practical goals. Two important theoretical goals are to obtain information about the expectability of the explanandum or about the realizability (or the possibility) of the explanandum. Two important practical goals are to obtain information that allow for the control of the explanandum, or information that enables us to form our ethical or juridical attitude towards the behavior (i.e., the intentional actions) of a human being. Requests for explanations that are motivated by the latter practical goal, for example, “Was it acceptable or praiseworthy that *x* did B?”, “Is *x* responsible for the consequences of his behavior?” etc. clearly aim

---

<sup>6</sup> Elster (1999) distinguishes two types of mechanisms that help capture weak or general patterns. We know a Type A mechanism if we know a set of incompatible results that can be expected, without knowing which one. Type B mechanisms tell us that the result will be one of two opposite results.

for information about the internal causal factors, that is, the reasons, that provoked the behavior. In these cases, the questioner wants to know about the deliberation process that motivated the agent, his beliefs, desires, intentions, etc. Put differently, such requests for explanations aim for Folk Psychological or intentional explanations. An externalists approach, like the dispositional model, is more convincing, we believe, when it recognizes its limits, respects the indispensability of an internalist approach, and when it is receptive to uncover their mutual relation.

### References

- Alexander, J. C. (1998). *Neofunctionalism and after*. Oxford: Blackwell Publishers.
- Anderson, J. R., (1985). *Cognitive psychology and its implications* (2nd ed.). New York: W. H. Freeman and Co.
- Armstrong, D. M. (1969). Dispositions are causes, *Analysis* 30, 23-26.
- Armstrong, D. M. (1999). Comment on Ellis. In H. Sankey (Ed.), *Causation and the laws of nature* (pp. 35-38). Dordrecht: Kluwer Academic Publishers.
- Armstrong, D. M., Martin, C. B., & Place, U. T. (1996). *Dispositions: A debate* (T. Crane, Ed.). London: Routledge
- Barkow, J., Cosmides, L., & Tooby J. (Eds.). (1992). *The adapted mind*. New York: OUP.
- Bourdieu, P. (1990). *The logic of practice*. Cambridge: Polity Press.
- Carnap, R. (1939). Testability and meaning. *Philosophy of Science* 3, 439-53.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.
- Ellis, B. D. (1999). Causal powers and laws of nature. In H. Sankey (Ed.), *Causation and the laws of nature* (pp. 21-42). Dordrecht: Kluwer Academic Publishers.
- Ellis, B. D., & Lierse, C. E. (1994). Dispositional essentialism. *Australasian Journal of Philosophy* 72, 27-45.
- Elster, J. (Ed.). (1986). *Rational choice*. New York: NYUP.
- Elster, J. (1999). *Alchemies of the mind, rationality and the emotions*. Cambridge: UP.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. NY: Prentice-Hall.
- Gopnik, A., & Wellman, H. (1994). The theory-theory. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257-293). New York: Cambridge University Press.
- Gordon, R. M. (1986). Folk psychology as simulation. *Mind and language*, 1, 158-171.
- Harré, R. (1970). Powers. *British Journal for the Philosophy of Science*, 21, 81-101.
- Harré, R., & Madden, E. H. (1975). *Causal powers: A theory of natural necessity*. Oxford: Blackwell Publishers.
- Hempel, C. G. (1949). The logical analysis of psychology. In H. Feigl & W. Sellars (Eds.), *Readings in philosophical analysis* (pp. 373-384). New York: Appleton-Century-Crofts.
- Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: The Free Press.
- Hempel, C. G. (1966). *Philosophy of natural science*. New York: Prentice-Hall.
- Hempel, C. G., & Oppenheim, P. (1965). *Studies in the logic of explanation*. In C. G. Hempel. (Original work published 1948)



## DISPOSITIONAL EXPLANATIONS

- Hogan, R., Johnson, J., & Briggs S. (1997). *Handbook of personality psychology*. San Diego, CA: Academic Press.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-146). New York: Wiley.
- Lipton, P. (1991). *Inference to the best explanation*. London: Routledge.
- Lipton, P. (1999). All else being equal. *Philosophy*, 74, 155-68.
- Mead, G. H. (1970). *Mind, self and society. From the standpoint of a social behaviorist*. Chicago: University of Chicago Press. (Original work published 1934)
- Quine, W. V. O. (1974). *Roots of reference*. New York: Open Court.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson Press.
- Salmon, W. (1989). Four decades of scientific explanation. In P. Kitcher & W. Salmon (Eds.), *Scientific explanation* (pp. 3-219). Minneapolis, MN: Minnesota Press.
- Skinner B. F. (1969). *Contingencies of reinforcement*. New York: Appleton-Century-Crofts.
- Staddon, J. E. R. (1973). On the notion of cause, with applications to behaviorism. *Behaviorism* 1, 25-63.
- Staddon, J. E. R. (2000). *The new behaviorism. Mind, mechanism and society*. Sussex: Psychology Press.
- Staddon, J. E. R. (2001). *Adaptive dynamics: The theoretical analysis of behavior*. Cambridge, MA.: MIT Press.
- Van Fraassen, B. (1980). *The scientific image*. Oxford: Clarendon.
- Woodward, J. (2000). Explanation and invariance in the special sciences. *British Journal of Philosophy of Science*, 51, 197-254.