Causality: Epistemological and Cognitive Considerations

2

2.1 Introduction

In philosophy intuition is used in reasoning as a test-bed for the conclusions of philosophical arguments. Logic, rhetoric and intuition are the main conceptual tools in philosophical reasoning. Intuition often acts as a sort of empirical verification of the acceptability of a particular thesis. Rather like a sort of empirical test or an experimental control, to use an analogy with what happens in natural science. The basis for this method is that intuition is generalisable, or in other words, broadly speaking, it can be shared at a universal level. Moreover, intuition must have foundational validity, a primary capacity for justification that is greater than any other alternative information. It should be greater than the reference to data from the cultural and religious tradition, for example, or the recourse to the theses of classical authors. Likewise it should be able to withstand the hypotheses and empirical confirmations of scientific and technical knowledge.

Experimental philosophy appears to question intuition's alleged foundational and universal nature. Intuition is a psychological phenomenon linked to what is conventionally known, according to some authors (Stanovich 1999; see Chap. 9 of Viale 2012), but not to others (Gigerenzer 2007), as System 1 of mind. Contrary to System 2, which is rational and explicit, this system is implicit and highly context-dependent. It is permeable to the influences of emotional variables derived from the cultural and environmental context. Seen in this way, it would seem difficult to affirm the thesis of the universality of human intuition. The underlying hypothesis derived from the findings of cognitive science argues the contrary: namely that intuition is local and contingent, changing in relation not only to cultural context but also to individual psychological variables, like personality traits or emotional and affective contingencies. Experimental philosophy has explored the universality

The present chapter is a modified version of Viale R. (1997). Causality: epistemological questions and cognitive answers, in G. Costa, G. Calucci and M. Giorgi (Eds.), *Conceptual Tools for Understanding Nature,* Singapore: World Scientific, With kind permission from the Publishers.

R. Viale, *Methodological Cognitivism*, DOI 10.1007/978-3-642-40216-6_2,

or otherwise of human intuition at an empirical level (Alexander 2012). In the first place it has debunked the myth of a form of universal intuition typical of the domain of philosophers. Like all experts philosophers present the same variability and context-dependency as ordinary people. Experimental philosophy uses the methods of cognitive and social science to understand the phenomenology of intuition: how we construct theories around concepts of external reality, how we construct conceptual categories around objects from the same reality, or how the mind elaborates the meaning we give to concepts. An important chapter of experimental philosophy relates to moral philosophy. Are moral rules based more on reason or on emotion? Does universality exist in moral judgment or do the situation and the cultural and social context determine that judgment? Or, even more radically, does an individual possess stable moral judgment or does it change depending on the emotional and pragmatic circumstances affecting the individual when the decision is made? The *situationism o*f Harman (1999) had already given a negative reply to those who supported a *character-based virtue ethics*. Moral judgments depend on the situation in which they are given. Therefore, they are local and not universal. Referring to David Hume's *sentimentalism*, researchers like Nichols (2004) and above all Prinz (2007) relaunched the thesis of a strong link between moral judgments, emotions and sentiments. According to the latter's *strong emotionism*, emotions are not only responsible for judgments, but they are also components of the moral norms themselves. Neuroethical studies seem to provide interesting answers to the ways in which we respond to the trolley problem.[1] People respond differently to this test depending on who is on the track. In many cases, we are willing to sacrifice the person if he is ugly and fat, while in other cases, if it is a child, for instance, and if women are responding, then the answer tends to be negative. fMRI studies (Greene 2008) appear to show that two different brain areas are activated depending on whether moral judgments are made using deontological rules or by analysing the practical consequences. The first type of judgment is linked to the area of emotion while the other relies on reasoning.

Experimental philosophy is based on the relationship between philosophy and psychology. While it has always been present in the history of philosophy, experimental philosophy has experienced varying fortunes. David Hume is a classical example of the pervasive use of the psychology of his time. Moreover, his work on the nature of the intellect was also an important source of inspiration for the development of associationist psychology. He can be said to have made one of the earliest attempts to construct a philosophical argument on an experimental

[1] "Suppose that you are a driver of a trolley. The trolley rounds a bend, and there come into view ahead five workmen, who have been repairing the track. The track goes through a bit of a valley at that point, and the sides are steep, so you must stop the trolley if you are to avoid running the five men down. You step on the brakes, but alas they don't work. Now you suddenly see a spur of track leading off to the right. You can turn the trolley onto it, and thus save the five men on the straight track ahead. Unfortunately. . .there is one track workman on that spur of track. He can no more get off the track in time than the five can, so you will kill him if you turn the trolley onto him." (Thomson 1985, p. 1395).

basis. How can this relationship be outlined? The content of this and other chapters in this book are examples of how scientific research on the human mind can help to define a number of philosophical problems and the relative solutions. In addition to ethics, experimental philosophy can be applied to a number of fields, in particular epistemology, metaphysics, ontology and aesthetics. This book will examine a number of problems linked above all to epistemology and to the philosophy of science. This first part will start by tackling a problem that is often seen as straddling metaphysics, ontology and epistemology, namely causality. It will be argued that by analysing causal reasoning from early infancy to adulthood it is possible to attempt to give an answer to the law of causality in nature and to causal laws and explanations.

This paper has two main goals.

1) To describe what cognitive science may suggest to philosophy concerning the reality of Causal relations (see also Chap. 3 of this volume);
2) To highlight the convergence between epistemology and the psychology of causation concerning tentative models of causal attribution and their anomalies.

2.2 Epistemological Questions

Some of the main issues that arise in the philosophy of causation concern the following questions:
Which are more basic, Causal relations or causal laws?
Are both or neither related to the non-causal state of affairs?
If the latter answer is negative, does the Causal relation derive immediately from experience or is a theoretical relation not directly observable?

There are three main answers to these questions.

a) According to the Humean interpretation, causal laws are more basic than Causal relations since the latter are logically '*supervenient*'[2] on the former, together with the non-causal properties of, and relations between, events. As regards the relation between the causal and non-causal state of affairs, this point of view holds that all causal facts are logically 'supervenient' on the totality of all non-causal facts. We cannot experience Causal relations directly, but only following one another between non-causal phenomena *a* and *b*. The mind will infer a Causal relation between *a* and *b* after having attended a certain number of repetitions of the same relation. But what we believe to be a singular Causal relation is only an application of the mental causal law that our probabilistic reasoning has inductively established. This position may be labelled conventionalist according to causal laws and reductionist with regard to the relation with non-causal facts (*Reductionist Conventionalism of Causation—RCC*).

[2] A set of properties *A* supervenes upon another set *B* in order to ensure that no two things can differ with respect to *A*-properties without also differing with respect to their *B*-properties. In slogan form, 'there cannot be an *A*-difference without a *B*-difference'.

b) According to the *Theoretical Realism of Causation (TRC)* Causal relations are real, but we cannot experience them directly. Causal concepts are theoretical concepts so that Causal relations can only be characterized, indirectly, as those relations that satisfy some appropriate theory (Tooley 1990, pp. 215–36).
c) According to the *Empirical Realism of Causation (ERC)*, Causal relations are more basic than causal laws and do not depend on the non-causal state of affairs. We can observe the Causal relations, not only in the everyday sense of that term but also in a much stronger sense which entails that the concepts of Causal relations are analytically basic. As Armstrong (1968) and Fales (1990) have pointed out, knowledge is strictly perceptual and has nothing to do with inference. It is like the perception of something pressing against one's body.

What is the contribution of cognitive science to these questions of the philosophy of causation?

It is not the aim of this chapter to justify the contribution that cognitive science may make to philosophical arguments. What I wish to point out are two similar positions concerning the relations between cognition and epistemology.

One is the 'naturalizing epistemology' programme. Cognitive science contributes to this by supplying the models of how the mental machine functions, how it processes information, how it produces the 'torrential output' of knowledge from the 'meagre input' of information from the world, to use Quine's terminology (Quine 1985). But isn't there the risk that epistemology becomes only a descriptive enterprise, that we lose sight of the other side of the moon? A possible answer that is now popular in the theory of rationality and ethics is that we may extrapolate the intuitive or the cognitive models of our justification of what is rational or irrational, or in the case of ethics our intuitive or cognitive models which we use to justify what is right or bad. Intuitive models are realized through one of the various different approaches to *reflective equilibrium*.[3] Cognitive models are realized through the methodology of cognitive sciences.[4]

[3] Goodman's (1965) proposal of reflective equilibrium tries to answer the question of how we justify a principle of inference. According to Goodman this is justified by its conformity to accepted inferential practice, and by its accordance with the singular inferences of everyday life. The crucial test of this claim is to check the intuitive acceptability of the inferential rules generated by reflective equilibrium. Stich and Nisbett (1980) have, experimentally, shown how some irrational rules, like the 'gambler fallacy' and other probabilistic biases, got through the reflective equilibrium test even when researchers asked the subjects to reflect on the principles that support their conclusions. The first volume (Viale, Chapter 13, 2012) contains a more detailed analysis of the *internalist* approach to justification.

[4] A different approach is to put aside the common sense criterion and to try to discover the cognitive mechanisms that are responsible for the justificatory processes of our inferences, and in this case for causal attribution. We may extrapolate the cognitive models of our justification of what is rational or irrational. This approach tries to individuate an ideal cognitive mechanism that is responsible for justification and that might allow us to establish a demarcation criterion between what is accepted and what is not. This claim too has many flaws. The notion of justification varies between individuals and therefore it is hard to provide an adequate characterization of a common concept. There is also a great variability in the cognitive procedures responsible for justification.

A weaker alternative position is put forward by philosophers like Alvin Goldman (1986; 1993). When confronted by criteria of rationality, such as the logical consistency of our set of beliefs—namely that rational beliefs must not jointly entail any contradiction—we are faced with an infeasible norm. Knowing a little about how the database of human memory is structured, full of contradictions and temporal structures, we must conclude that the whole of the human race is irrational. Therefore cognitive science may be useful to address questions concerning the feasibility of our epistemological desiderata in relation to the constraints of the human mind. Cognitive science may be relevant in setting standards for epistemology that fit the test of cognitive feasibility.

2.3 The Perception of Causal Relations

One of the first problems that found an interesting answer in experimental psychology dealt with the direct perception of Causal relations, a possibility denied by the Humean tradition and by theoretical realism (TRC).

Research on the perception of causality based on visual information began with attempts to apply the methods of Gestalt theory. According to these methods we perceive a pattern like that of causality not by learning but because our brain processes are configured to respond to key recurrent patterns. The most brilliant results in this field were obtained by Michotte (1952). He proposed that forms of mechanical or physical causation may be directly perceived through the patterns of motion of two objects and not necessarily derived from our experience of their succession, as in Humean position (1888). Michotte experimented with hundreds of patterns and concluded that two evoke universal and immediate impressions of causality: 'entraining' in which object A collides with the stationary object B and they both move off together, and 'launching' in which A collides with the stationary B and B alone moves off. Adults can have strong stable intuitions about the causal nature of connections between the relative movements of dots on a screen, depending on precise variables in their relative movements. Therefore, while these discoveries seem to weaken the Humean and theoretical realist positions, on the contrary they support the claim of the third point of view, that of empirical realism. Perceptual cues alone, without the help of inferences, appear to contain a great amount of information about Causal relations.

Even in the social domain, this perception seems very crucial. Heider and Simmel (1944) investigated the perception of social causality in patterns of motion. Like Michotte, they identified patterns which evoked causal perception: 'Simultaneous movements with prolonged contact (like entraining)...Successive movements with momentary contact (like launching)...Simultaneous movements without contact...Successive movements without contact' (1944, pp. 252–5).

Our notion of justification seems to be based on different topic-specific default concepts that change according to the individual and the field of justification.

Faced with these results, one reaction from a Humean philosopher might be that these performances are based on a particular style of perception, namely one that is learned indirectly through analogous visual experiences with the same causal content and specific to our western culture. In this case, the perception of inference would play a crucial role in the attribution of causality. How can this kind of objection be rebutted? One possibility is to prove that the perception of Causal relation is 'age and culture independent', in other words that it can be found in babies and in tribes.

2.4 Developmental and Cross-Cultural Findings

Recent psychological research seems to have increased our knowledge in this field. Evidence that humans are equipped with a module for perceiving physical causality from motion comes from findings that young children, and even infants, distinguish trajectories that are consistent with movements driven by physical force (e.g. one object moves when another collides into it) from highly similar but anomalous trajectories. Experiments have established that infants make this distinction before they can induce knowledge of causality from experience (Leslie 1982, 1987). Cross-cultural studies have shown that the perception of physical Causal relations are alike among infants and adults in very different cultures from ours, like Hindu and Chinese; that the perception of social causality is also very alike in different cultures during childhood, but change radically among adults. The individualist American attributes more social causes to individual internal dispositions, while, on the contrary, the collectivist Chinese attributes more social causes to the external social context. This different attributional model is reflected in many cultural expressions, such as painting (Fig. 2.1) (Morris et al. 1995, pp. 577–612).

These findings seem to give strong support to the anti-Humean third position of empirical realism. Except for social causality, which is permeable to public representations (Sperber 1985, 1991) or cultural values, the mental representations of physical causality are not affected by learning and culture. Moreover, these data find theoretical support in Fodor's modular theory of the mind (1983). According to Fodor, the perceptual systems are modular, that is they are separated by a barrier that separates information from other parts of the larger cognitive system, especially from what he calls the 'central system'. Modular input analysers have privileged inputs and are not subject to information from other parts of the system. The Müller-Lyer illusion is the paradigmatic example of this perceptual modularity. Even if Fodor admits a background theory and some inferential content in perception, these play a very limited role.

But what the recent research has shown are also other characteristics of causal perception that do not agree with the ERC position and with Fodor's theory of mind (see Viale 2012). Data from developmental studies and a certain universality in the characterisation of causal perception in cross-cultural studies seem to support the hypothesis that we are endowed with early-developed cognitive structures, which correspond to maturational properties of the mind–brain. These orient the subject's

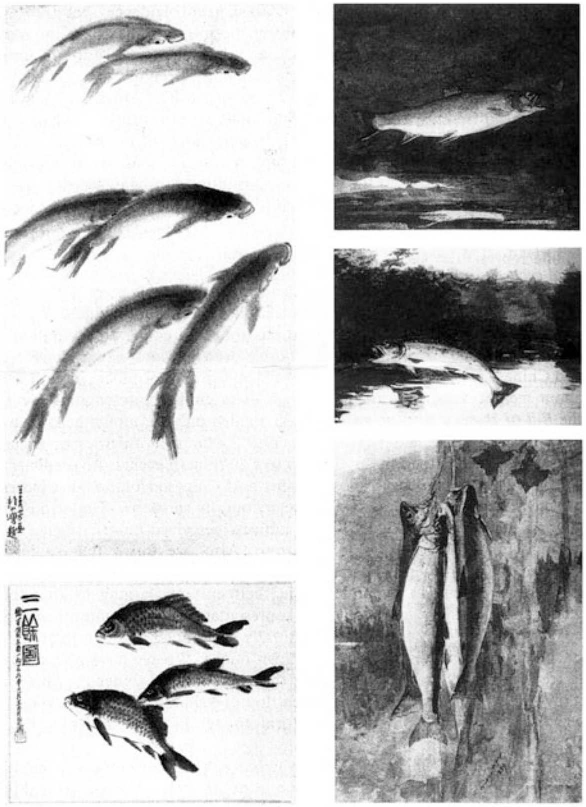

Fig. 2.1 On the *left* are *Fish* and *Three Fish* by Wang Ch'ing-fang (1900–1956), a Chinese painter best known for depictions of fish, whose watercolours are admired for capturing the group's' rhythms of movement' (Hejzlar 1978). On the *right* are *Leaping Trout*, *Trout*, and *Adirondack Catch* by Winslow Homer (1836–1910), perhaps the most prominent American painter of fish, whose watercolours are noted for capturing the fish's 'magnificent struggle' against nature, man, and 'impending death'—and only in death are fish portrayed in a group (Cooper 1986)

attention towards certain types of cues, but also constitute definite presumptions about the existence of various ontological categories, as well as what can be expected from objects belonging to those different categories. Moreover, they provide the subjects with 'modes of construal' (Keil 1995), different ways of recognizing similarities in the environment and making inferences from them. More surprisingly, contrary to Piagetian theory—according to which the notion of causality is domain-general and gradually modified by experience—'different conceptual domains are structured by different principles which (1) carry information about the types of stimuli that are likely to correspond to particular ontological categories, (2) convey expectations about non-obvious properties of objects in different domains, (3) constrain the manner in which spontaneous inductive inferences are made about objects from different domains' (Boyer 1995, p. 623). The previous Piagetian notion of formally defined stages, characterized by principles which apply across conceptual domains, has been replaced by a series of domain-specific developmental schedules, constrained by corresponding domain-specific principles. These principles constitute a core of, probably innate, 'intuitive theories', which are implicit and constrain the later development of the explicit representations of the various domains. As Gelman highlights, 'different sets of principles guide the generation of different plans of action, as well as the assimilation and structuring of experiences' (1990, p. 80). They establish the boundaries for each domain and single out stimuli that are relevant to the conceptual development of the domain.

The three main intuitive theories individuated by developmental psychology are the theory of physical objects, the theory of biology and that of psychology. These theories allow infants to individuate some theory-specific causal mechanisms that explain interactions among the entities in the domain. The child has intuition of what characterizes a living being from an artefact or an object. Between the ages of 2 and 5 the child assumes that external states of affairs may cause mental states and that there is a causal chain from perception to beliefs, and from intentions to actions.

The intuitive theory of physical causality is the least controversial and a rich source of empirical data. Intuitive physical principles orient the child's understanding of the physical environment from infancy. Principles specifying that solid objects are cohesive and continuous and are not susceptible to action at distance (Fig. 2.2) seem to emerge before the age of 4 months (Leslie 1988; Baillargeon and Hanko-Summers 1990; Spelke 1990). At around 6 months the infant is able to apply a principle of support—namely that an object will fall if it is not supported (Spelke 1990). The specific patterns of movements allow infants to make ontological distinctions between self-generated and non-self-generated movement (Massey and Gelman 1988). This distinction gives an initial skeleton to a differentiation between animate and inanimate objects, which has important consequences for causal reasoning in the biological and psychological domain (Fig. 2.3).

2.5 Epistemological Reflections and Implications

What are the implications of these data for the epistemology of causation? It appears that these studies provide greatest support for the second position of theoretical realism. In order to recognize a relation between objects as a Causal relation we appeal, automatically, to an implicit, innate theory that is domain specific. The perceptions of causality are theory-laden and it is impossible to outline a purely empirical perception of causality. The presence of these innate theories may also account for the explanatory ability of the perceiver. As many experiments have shown, the child can explain and predict the behaviour of the effect on the basis of the cause. They do not perceive the causality in the relation between two objects but are able to use inferential reasoning according to the top-down intuitive theory.

Will this answer satisfy the philosophers? Are the many domain-specific intuitive causal concepts satisfactory representations of causality in the real world? This decentralized, piecemeal approach to causality takes the opposite line to the age-old philosophical enterprise of establishing a general framework, an intentional model to define causality.

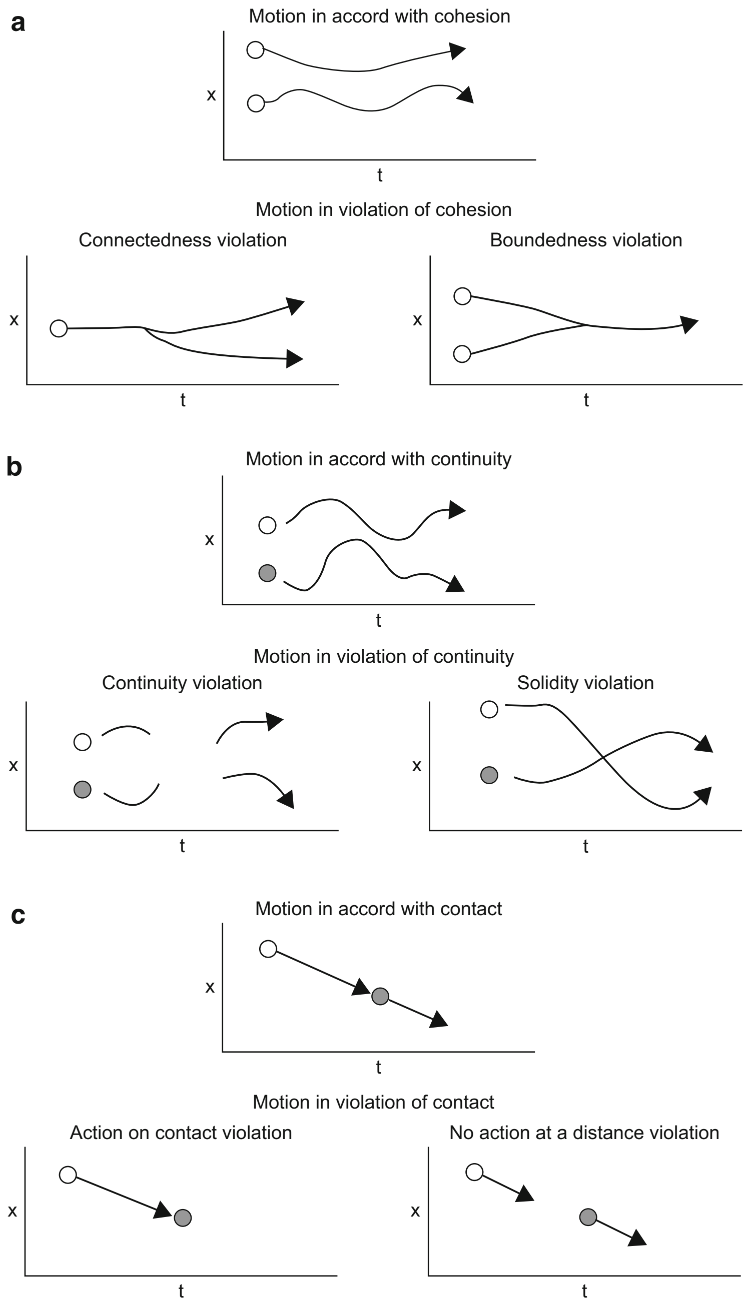

Fig. 2.2 Principles guiding infants' physical reasoning. (**a**) The principle of cohesion: a moving object maintains in connectedness and boundaries. (**b**) The principle of continuity: a moving object traces exactly one connected path over space and time. (**c**) The principle of contact: objects move if and only if they touch (Spelke et al. 1995)

According to the 'feasibility criterion' we ought to assess our prescriptive models of causality according to human causal thinking. While it is easy to apply the feasibility criterion to deductive or probabilistic reasoning, because the target prescriptive models are very clear, the situation is very different in the case of causality.

Therefore, what can we say about the relation between causal cognition vs causal epistemology?

a) **Intuitive domain-specific theories vs learning mechanism**: the hypothesis regarding intuitive domain-specific theories appears to be underdetermined by data on causal perception in infants. The same data can support an alternative hypothesis about the presence not of intuitive domain-specific theories but of an innate learning mechanism plus a restricted core of innate beliefs. According to this alternative perspective, 'the infants first form a preliminary all-or-nothing concept that captures the essence of the phenomena but few of its details. In time this initial concept is progressively elaborated. They identify discrete and continuous variables that are relevant to the phenomena and incorporate this accrued knowledge into their reasoning, resulting in increasingly accurate interpretations and predictions over time' (Baillargeon et al. 1995, p. 80). According to this model infants are born with a highly constrained mechanism that guides their acquisition of knowledge about the objects. Data supporting this model come from recent psychological research on the development of infants' intuitions about phenomena like support, collision, unveiling, arrested-motion, occlusion and containment. If this hypothesis is true, the default causal learning weakens the support for theoretical realism and strengthens the Humean position.
b) **Wide and domain-specific causality vs restricted and domain-general causality**: the cognitive use of the causal label looks, from an epistemological point of view, too wide with regard to the type of relation among phenomena and too specific with regard to the ontological context. The epistemological concept of causality tries to define the logical properties of causality that can be applied in all natural domains. Generally they are formal a priori criteria, modal or statistical notions that don't pay too much attention to the division into domains of the natural world. Moreover, most philosophers consider it unacceptable to label the perception of phenomena as causal: for example, the cohesion of an object *as a consequence* of its not crumbling, or its solidity *because* it does not crush, or containment *as the result of* the smaller object contained in a bigger container, or the unveiling *as the result of* perceiving the existence of the object even if it is covered. Most of these phenomena are related to the properties of the objects while causality refers to the relationship between two different events. It is true that the properties of an object have implications in terms of explaining and predicting its behaviour. But in the main epistemological models of causality, the object called cause must be separated by the object called effect. Stating that the apple is coloured *as a result of* being green is generally not accepted as a Causal relation. This view is challenged by some philosophers who implicitly justify the wider concept of causal cognition. For example, according to Sosa

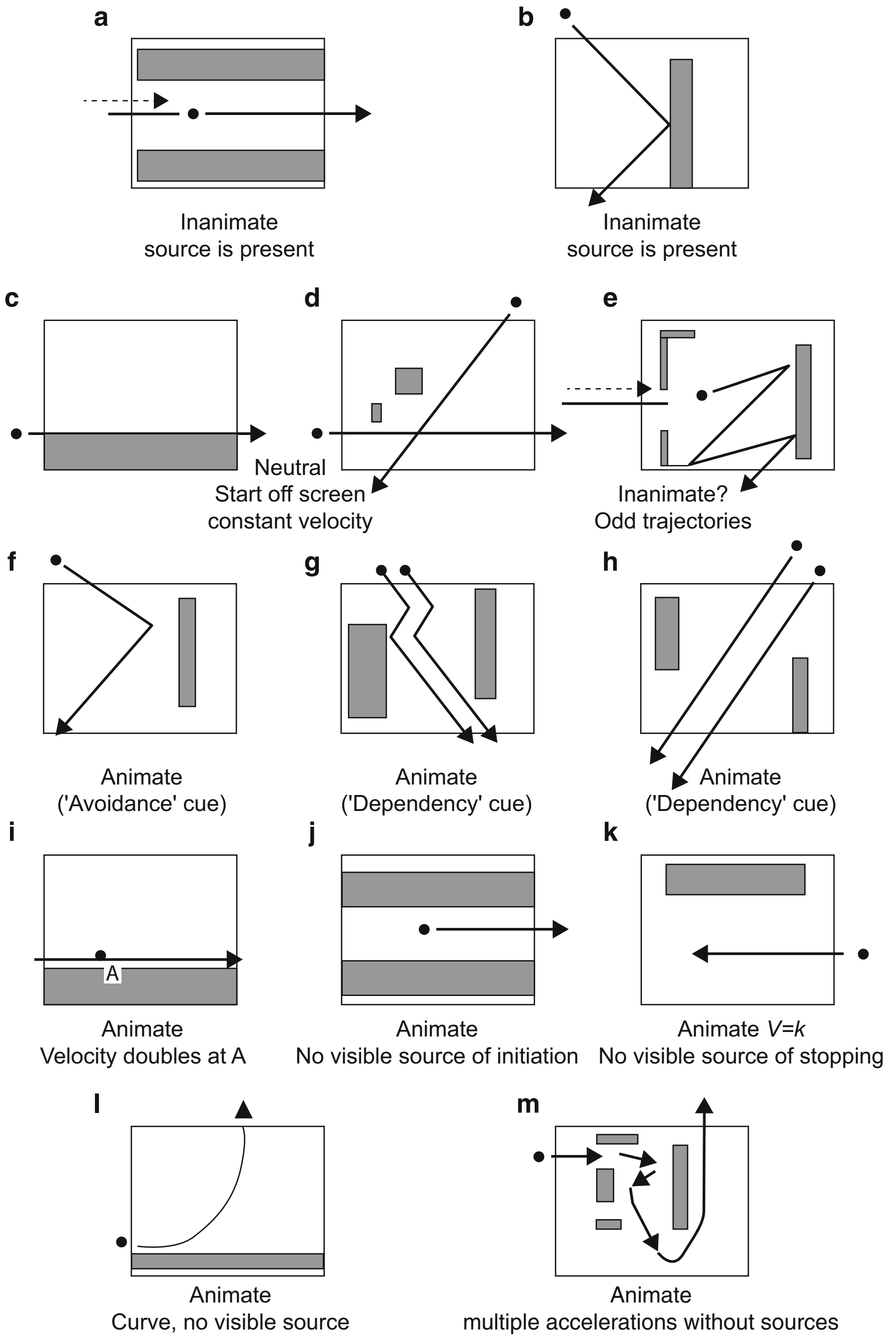

Fig. 2.3 Differentiation between reasoning about animate and inanimate objects (Gelman et al. 1995)

(1980) nomological causality should be added to other types of causality that satisfy the necessary condition. In his view, there are three kinds of causality (material, consequentialist and inclusive) and they are represented by sentences that use the terms '*because of*', '*a consequence of*' and '*as a result of*'. All these types of causality are relations between a source and a consequence or result and each of them is a case of necessitation, like nomological causality. This point of view works very well with causal cognition, but it is far from the philosophical main stream.

c) **Causal mechanism vs causal law**: there is some confusion among psychologists of causation when using causal concepts. To summarize, there are two kinds of philosophical traditions on causality. The first focuses on the causal laws, while the second focuses on the Causal relations or mechanisms. According to the former, causal reasoning relies on the generation and application to the reality of general causal laws or law-like statements. According to the latter, causal reasoning relies on singling out local Causal relations or mechanisms. The psychologists tend to confuse and mix them. Some, like Carey (1995, p. 268), write that the "mechanism tradition" is based on explanation depending from general laws, while the logical tradition has to do with the modal and statistical conditions of the relations between cause and effect. The former is domain-specific while the latter domain-general. On the contrary, it is well known that the theory-based explanation, like the nomological deductive model, tries to fix general criteria of causal understanding and generally all the philosophy of causation is domain-general. Instead the most sensible tradition, in terms of approaches to the problem of context-specificity, is that of philosophers like Mackie, working on modal notions like necessity and sufficiency, who introduces the concept of '*causal field*' to separate the causal factors from the mere conditions in each causal context (see the fifth paragraph of this chapter). But this is a general model applicable in every domain. Moreover, even a supporter, like Salmon, of the mechanism tradition does not allow any domain-specific interpretation of his models. Besides it is not clear in the psychology of causation if the intuitive domain-specific theories are general principles that apply to the interpretation of the real world or are concepts that allow the local identification of singular mechanisms and Causal relations. The various principles of cohesion, continuity and so on seem to correspond to the former case, while the primitive mechanical notion FORCE, outlined by Leslie (1995), seems to adapt to the second case.

d) **Mechanical, functional and intentional causality**[5]: some misunderstanding of the notion of domain-specific causalities derives from the interpretation of the specificity as related to the formal properties of the causal concepts. The

[5] The term causality applied to the biological and psychological domains corresponds to the recent reformulation of the functional and intentional models in terms of causal concepts made by philosophy of biology and psychology. If the intuitive conceptual grasping of biology and psychology of infants and children is really in terms of causal representations we may say that is more up-to-date than previous methodological and philosophical models.

psychology of causality seems, sometimes, to hold different formal types of causality according to the domain. Instead, the empirical data that they obtain show only different kinds of explanation for different domains based on different causal factors: the behaviour of inanimate objects is explained in terms of force, thrust, obstacles and resistance; that of animate objects in terms of beliefs, intentions and so on. The specificity only concerns the different events that are considered causal in producing an effect, depending on the different parts of its nature under investigation. This is the same use of causality made by philosophy and science. Moreover, domain-specificity that presupposes a partial modularity at the conceptual level is not supported by the fundamental Fodor's theory of mind, which is one of the main theoretical bases of the cognitive research tradition. The modularity may be only at the perceptual level, while the conceptual level is holistic. Besides, there are other models that can account for causal reasoning and are not domain-specific. For example, the mental model theory of Johnson-Laird (1983) proposes a theory of sub-concepts that can account for a domain-general causal reasoning. It relies on three kinds of general sub-concepts: those for temporal relations, those for negation and those for the epistemic state.

e) **Causal realism from the evolutionary point of view** (this point will be analyzed in Chap. 3 of this volume): if, according to naturalizing epistemology or the feasibility criterion, there are some lessons for philosophy that come from research on causal cognition, it is mainly about the reality of Causal relations. There is a lot of data from developmental and cross-cultural studies showing that human beings universally perceive, represent, explain and predict the necessity of given effects after given causes. Is this sufficient to assert that causality is the cement of the universe? In philosophy the debate on causality and in particularly on what characterizes the necessary relation between cause and effect is age-old and not conclusive. As we have seen at the beginning, two of the traditions support causal realism and this position is gaining increased attention from the philosophical community. Moreover, what characterizes many of the philosophical arguments on causation on both sides of the barricade is the frequent appeal to intuition, common sense, ordinary language and other cognitive concepts, as in the case of many other philosophical arguments. From Aristotle and Hume up to today the philosophy of causation founded its arguments heavily on how the mind processes information, how it represents reality, how it establishes folk scientific hypotheses on the natural and social world. But information on mental activity relied mainly on the personal and idiosyncratic intuitions of the philosopher. Cognitive science nowadays contributes to fill this gap of knowledge and allows philosophy to reason, starting from better founded mental notions and in this case from notions that assert, strongly, the causal structure of the world. Can we avoid drawing a conclusion from this pervasive tendency of mental activity? To answer yes to this question would be, by analogy, like asserting that another pervasive tendency of mental activity, namely the perception of objects, colours, shapes, sounds, smells and so on, has nothing to do with the assertion of the reality of the external world. We see, we touch, we hear and we smell and this mental activity is, instead, the main basis for affirming the reality of the external world. Therefore, the argument for analogy might also allow us to assert the truth of causal realism. But the reason for drawing this conclusion might be another. A challenger of causal cognition might reply that the perception of causality is an illusion, like the Müller-Lyer illusion or many others, and therefore it does not represent how the world is made. It corresponds only to some wired-in brain devices that constrain our perception and representation of the real world in a non-realist way. How can we respond to this objection? The answer might come from evolutionary theory. It is well-known that evolution is related to changes in genetic frequency and this phenomenon is a result not only of natural selection but also of the differential rate of mutation, of migration and genetic drift. These different mechanisms may be responsible not only for improving the adaptation but also for the fixing less adaptive genes and, consequently, suboptimal phenotypical characters. There are many examples of this effect, such as *pleiotropy*, *heterozygote superiority* and *meiotic drive*. One example will suffice for all: albinism in Arctic animals is often symptomatic of serious ocular disease, as well as producing white fur. Therefore, from an evolutionary point of view it might be explained that during its evolution the human race has selected some negative characteristics that are responsible for suboptimal elements of mental activity. This explanation may account for local illusions, like the Müller-Lyer illusion, which do not have much effect on human evolutionary fitness. However, it cannot explain the presence of negative characteristics that pervade mental activity and dramatically decrease its ability to represent and predict natural events. The pragmatic impact on the evolutionary fitness of an illusory causal cognition would be too negative to be allowed in the evolution of the human race. This consideration is more probable if we think that causal cognition has been found not only among humans but also in many other species that have similar problems of adaptation. Humean causality based on associative learning of the repetition of contiguous events is even found in brain-less micro-organisms. Another type of causal knowledge not based on close contiguity but in innate interpretation of certain specific events is found in birds, fish and insects. The animal knows that a causal connection between two events is highly probable,—e.g. a certain behaviour during the courtship produces a certain effect in the other animal. Lastly, there are tool-using animals like chimpanzees and orang-utans that have the same perception of physical causality as that perceived by man in Michotte experiments. They know how to hit one object with another and they show good technical abilities in nest construction and tool-use (Sperber et al. 1995). Therefore, if evolution does not allow the selection of a mind that misrepresents important aspects of reality, can we assert that reality is causal? The answer is yes, at least with the same certainty that we have in affirming that the world is made of singular objects, like chairs, apples and dogs. All these perceptions and representations derive from mental modelling and are not completely bottom-up, but our behaviour and the relative positive pragmatic feedback from reality tends to reinforce realist cognitive style. Environmental correcting feed-back is proved by the presence in humans of a small number of spurious, causal attributions compared with the possible enormous number based on temporal and spatial contiguity. The phenomena of *epiphenomenon* and *substitutive causation* are relatively frequent in causal thinking. But it is also well known that there are many processes of correction based on empirical testing and counterfactual reasoning. Besides, there are many findings in cross-cultural research (Morris et al. 1995, pp. 577–613) that show that even magical thinking is based on religious beliefs and not on different causal cognition. In fact, contrary to traditional anthropological theories supporting pre-logical mentality, tribal people like those of Papua show normal causal thinking in many domains not under the theoretical influence of religious beliefs.

2.6 How Epistemology Identifies a Causal Relation

In the philosophy of causation one of the most debated problem has been how to individuate the cause of an effect. Traditionally the answer was the set of factors that together is sufficient to produce the effect. But this solution had many negative consequences. For example, in a house that has burnt down, the sufficient set can contain many factors, including a spark from a short circuit. For example, wooden walls, oxygen, the lack of humidity, the lack of a fire-prevention system and so on, up to cosmic irregularities. How can we restrict the sufficient set in order to avoid a possible regression to the infinite? How can we build a set of relevant causal factors?

The first answer appeals to the modal notions of **necessity and sufficiency.**

The most well-known model is proposed by Mackie and may be summarized as follows:

> '*If C is a cause of E (on a certain occasion) then C is an INUS condition of E, i.e. C is an Insufficient but Non Redundant part of a condition which is itself Unnecessary but exclusively Sufficient for E (on that occasion).*' (1974, p. 62)

I leave aside all the difficulties of this approach. One of these was how to find a criterion that allows the relevant cause to be extrapolated from the many irrelevant ones that are an insufficient but necessary part of the condition, like oxygen or wood, leading to a fire in a house. Mackie proposed the concept of '*causal field*', which was not defined but rather illustrated by examples. A question like 'What caused the house to catch fire' may be expanded into 'What made the difference between those times, or those cases. . .when no such fire occurred, and this particular instance when a fire did occur?' In this case the causal field is the number of normal and stable characteristics of the house, like the building materials, the lack of a fire-prevention system, the presence of oxygen and so on. Therefore what caused the fire must be a difference in relation to the causal field and the short circuit is the obvious candidate. But different causal fields implying different causal explanations may be chosen in different contexts for causal accounts of the same event. In the biological and medical sciences, this situation is frequent. The causal field of normal conditions changes considerably depending on the different disciplinary analyses of the cause of a disease. Moreover it is not specified how the mere conditions can be extrapolated from the non-causal ones according to this model.

The second answer is expressed in terms of **probabilistic and non-necessary relations**. The main concept is that of 'statistical relevance', based on the difference between the probability of an effect given the presence of a potential cause and that probability given its absence (Salmon 1984). The weakness of this concept is that not all statistically relevant relations are causal. For example, although a drop in the barometric reading varies with storms, one would not draw the conclusion that the drop in the reading causes storms. To explain the distinction between genuine and spurious causes one answer within the statistical relevance approach is to base judgements on conditional contrasts (Reichenbach 1956; Suppes 1970; Salmon 1984). A contrast for potential cause C with respect to effect E is computed within subsets of events in which alternative causal factors Kj are kept constant: if

$$P(E/C.K1.K2\ldots Kn) - P(E/C.K1.K2\ldots Kn) > 0$$

then C is inferred to be a facilitating cause of E. Or using the '*screening-off*' method, we may say that if

$$P(E/C.K1) = P(E/C)$$

then we can say that factor $K1$ has been screened off by C—e.g. the lack of boats on the sea *(K1)* has been excluded by the drop in barometric pressure *(C)* in relation with the storm *(E)*. But even following this criterion, we are not certain that there is a direct Causal relation between C and E, but only a general Causal relation. It might be the case that there is a common cause responsible for both C and E, for example the drop in atmospheric pressure *(D)*. How can we establish the presence of a common cause? By applying the *principle of the common cause* stated by Reichenbach in 1956:

$$P(C.K1) > P(C)\ x\ P(K1)$$

that is when two effects happen more frequently jointly than alone then there probably might be a common cause D that explains the scarce reciprocal autonomy. But even with this criterion, we are not sure that what we have found is a direct Causal relation. As Salmon pointed out, only by using probabilistic analysis can we be sure to avoid spurious Causal relations.

What suggestions can we draw from the difficulties of these two approaches? First, we must appeal to contextual criteria, like the causal field, to individuate the factors to be analysed according to their probability. Second, we must analyse the relative causal roles of the candidate factors utilizing empirically *subjunctive conditionals* of the following form:

'*if we change a given causal factor then the effect would be....*'

and in cases where the Causal relation cannot be replicated we should use *counter-factual conditionals* like:

> '*if a given causal factor had been changed then the effect would have been*. . ..'

Therefore, in both the epistemological traditions, modal and probabilistic, faced with formal inadequacies in ensuring a correct identification of the Causal relation, there has been a tendency to appeal to a-posteriori criteria based on pragmatic and cognitive factors—in the case of the notion of causal field—or based on empirical methods—in the case of conditionals.

2.7 How Cognitive Science Identifies a Causal Relation

Cognitive science research in the field of adult causal reasoning is inspired by previous models of philosophy of causation. There is a clear debt to the work of thinkers like Mill, Mackie and Hart & Honoré, on the one hand, and Reichenbach, Suppes and Salmon, on the other.

There are three main approaches that are based on different criteria of causal attribution, but which are separated by very fuzzy borders and are affected by more or less the same problems.

a) **Normality criterion**: this approach is more linked to the philosophical tradition and, in particular, to Mackie (1974) and Hart and Honorè (1959). In the context of their '*norm theory*', Kahneman and Miller (1986, p. 148) noted that 'the why question[6] implies that a norm has been violated' and 'requests the explanation of an effect, defined as a contrast between an observation and a more normal alternative'. A cause does not need to be statistically unusual, but it must be abnormal in the sense that is not 'a default value among the elements that the event [to be explained] has evoked' (p. 149). Hilton and Slugoski (1986, p. 77) write that among the set of individually necessary but jointly sufficient conditions, 'The abnormal condition that distinguishes the target case. . .becomes dignified as the cause. Those necessary conditions. . .that are not abnormal. . .are relegated to the status of mere conditions.'" Hilton adds (1995, pp. 495–526) that the contrast cases may be the normal functioning of a state of affairs—e.g. the house before the fire—an ideal model—e.g. a healthy body or a legal system—and a hypothetical case, which never occurred or might have occurred—e.g. the counterfactual scenario evoked by Bush of Iraq dominating the whole Middle East region if America did not intervene.
b) **Conversational criterion**: a cause is always a condition assumed to be unknown to the hypothetical inquirer—e.g. the short circuit in the house fire—and an enabling condition is typically a condition assumed to be already known

[6] A "Why question" is a typical question that needs a causal explanation. For example "Why did the house burn?"; "Why did Gore loose the presidential election?"; "Why did your mother get the pneumonia?"

to the inquirer—e.g. the presence of oxygen during the house fire (Hilton 1990; Turnbull and Slugoski 1988). This distinction is an application of Grice's (1975) conversational maxim of quantity, which prescribes speakers to be as informative as but not more so than is required for the purpose of an exchange. The *informativeness account* is similar to the notion of *relevance* introduced by Sperber and Wilson (1986). According to its criterion of relevant information, the main difference is being able to derive new assumptions. While all unknown conditions are informative, not all are conversationally relevant. A condition that is constantly present and unknown to a particular inquirer would be informative to him, but irrelevant because it would not allow him to predict the effect—e.g. the presence of a Van Gogh painting in the house that went on fire. Hilton (1990) proposed a conversational model of causal explanation that subsumes the normality criterion as a special case. He assumed that '…in explaining an event to a competent adult, we would refer to individuating features of the case which cannot be presupposed from general world knowledge, such as abnormal conditions, and omit to mention… [what] can be presupposed' (p.67).

Both normality and conversational criteria have the most serious problem in separating enabling conditions from non-causal ones. The normality criterion reintroduces the concept of necessity to specify the enabling conditions. But this move brings us to the age-old, intricate question of representing the necessary relations between events. Therefore, it looks like a very weak solution.

However, when the conversational criterion states that a cause is always a condition assumed to be unknown to the hypothetical inquirer, it is not able to distinguish between a short circuit and a Van Gogh painting or some other non-causal conditions. In this case it is not able to separate the cause from a non-causal condition.

c) **Probabilistic contrast model**: the identification of a cause depends on its covariation with effects on a focal set—the set of events implied by the context. Cheng and Novick (1991, p. 94) hold that 'the covariation is hypothesized to be computed over a focal set as specified by our probabilistic contrast model:

$$\triangle Pi = Pi - non\text{-}Pi$$

where i is a factor that describes the target event and Pi is the proportion of cases for which the effect occurs when factor i is present and *non-Pi* is the proportion of cases for which the effect occurs when factor i is absent. When ΔP is greater than some (empirically determined) criterion, then there should be a causal attribution to factor i.' The short circuit is a cause because it covaries with the fire in the focal set—e.g. the house. The oxygen is an enabling condition because is constant, that is it is the same when the house does not catch fire and when it catches fires. But it co-varies with fire in another focal set—e.g. fire in a chemical laboratory. The Van Gogh painting is a non-causal condition because it never co-varies with fire.

2.8 Concluding Remarks on Scientific Causal Reasoning

Many comments can be made on these cognitive criteria for causal attribution. The first is that these models cannot avoid appealing to very problematic philosophical notions, like necessity, and consequently they crash into the same barriers as the modal tradition. Moreover, when they correspond to the real causal cognition, some of them, like normality and conversational criteria, seem too loose: they cannot avoid including the Van Gogh painting among the causal conditions of the fire in the house. If they try to be more precise, as in the case of the probabilistic contrast model, they make the same mistake as many theories of rationality, such as game theory: they seem to place too much emphasis on the ability of human probabilistic computation. Moreover, they are too external and lack theoretical depth in terms of the mechanisms of the human mind. The mental model theory suggested by Johnson-Laird seems much more promising for establishing a deeper cognitive theory of causality that might meet the standard of the feasibility criterion or the more pretentious desiderata of the naturalizing epistemology programme.

Nevertheless, they provide interesting interpretative cues and suggestions for further research in causal reasoning and, in particular, among other forms of scientific causal reasoning.

Previous approaches (with the partial exception of Cheng and Novick's position) hold that scientific causal reasoning is different and cannot be analysed using the same models as everyday reasoning. The latter deals with particular events that can be explained using abnormal causes. The former is related to general events that are explained by normal causes. I think that this conclusion is mistaken because it relies on a non-realistic model of scientific reasoning.

If we want to use a cognitive language we can divide scientific reasoning into three types: bottom-up inductive modelling; top-down hypothesis testing; deductive coherence seeking.

If we exclude the latter, the former two often deal with causal reasoning.

The first *inductive modelling scientist*, by observing many particular events, tries to individuate the general Causal relations between potential causes and a given effect. He tries to compare a contrast case where there is no change in any variable with a target case in which there is an abnormal change in some of them. Then he will be able to analyse the relative causal roles of the potential candidates and to discover other factors, which are as yet unknown. For example, if he wants to give a causal explanation of the heart-beat and he does not know the conditions involved, he will compare a contrast case—e.g. the normal heart functioning—with a target case—e.g. an increase or decrease in the heart-beat. He will find that among the potential causes, there is an abnormal increase or decrease in the electric activity of the heart pacemaker and an abnormal change in the concentration of adrenaline. Instead the peripheral blood flow will be treated as an enabling condition because it remains, more or less, constant in both cases.

The main difficulty encountered in this kind of approach is to isolate the enabling conditions from the non-causal conditions in both scientific and everyday reasoning. Even the solution proposed by Cheng & Novick does not solve the problem. Indeed, using their approach, the enabling conditions are those that remain constant in the focal set—e.g. the peripheral blood flow is constant in our empirical set in which there is a normal heart-beat and a change in heart-beat—but co-varies with the effect of other focal sets—e.g. when the peripheral blood flow stops, so does the heart-beat. But this criterion does not allow the scientist to rule out from the enabling conditions those that are not causal, like the production of melanin or hay fever, which are constant in the original focal set but co-vary in other focal sets—e.g. when the heart-beat stops so do melanin production and hay fever. How can we cope with these difficulties in science as well as in everyday life? As I showed before, this can be achieved using counterfactual reasoning, when we have gained sufficient empirical knowledge about the functioning of factors involved in the Causal relation or in subjunctive reasoning—namely by empirically analysing the relative role of the different variables.

The second type of scientist, the *top-down hypothesis tester*, mainly follows the same kind of causal reasoning. For example, if he wants to test the hypothesis that the heart-beat depends causally on the electrical activity of the heart pace-maker whose synapses are mediated by noradrenalin, he will compare a contrast case—e.g. normal heart-beat—with a target case—e.g. increased and decreased heart-beat. If he finds an abnormal change in the electrical activity of the pace-maker and a change in adrenaline levels in the blood(relative to noradrenalin levels), he will be able to corroborate his causal hypothesis.

There is an interesting phenomenon that was discovered by Einhorn and Hogarth (1986) with regard to everyday reasoning that has important implications in relation to the understanding of scientific reasoning. They found how adding contextual information leads us to change an explanation. An alteration in the causal field of background presuppositions changes one target explanation in favour of another. The new contextual information suggests an alternative mental model which could explain the effect in question. In an experiment they observed how the preferred explanation for an employee's cancer, which was previously thought to be due to working in a factory where there was a high incidence of cancer, could change when we learn of his heavy smoking and family history of cancer. Another example is still more meaningful. In this case there is no alternative causal scenario, but only a refocusing on the same elements of the scenario and then a new mental model based on the same elements. For example, when we learn that a hammer strike, which ended by shattering a watch, happened in the context of a factory control procedure, we change our explanation of the destruction of the watch from 'because the hammer hit it' to 'because of a fault in the glass'.

In science the causal field is shaped by the constraints of the background presuppositions containing the disciplinary knowledge and the main principles of one's own research tradition. The disciplinary knowledge supplies the material for the reasoning and the conceptual tools of one's own discipline shape the representation of the premises. Given the same inputs of information, the premises of the mental model change according to these different, but neighbouring disciplines, and consequently the conclusion is different. But even the information retrieval and the selection of information from outside changes according to different research traditions. As in the previous examples, the causal scenario will sometimes be made up of different elements and at other times the same elements will have different causal roles.

The biomedical sciences are a good example of this variability in causal attribution because of the many neighbouring disciplines that often work on the same natural phenomenon. Each discipline is specialized to give a causal representation of an event.

For example, let us consider renal calculus. The disciplines involved in the etiopathogenetic explanations of its causes are many: genetics, biochemistry, microbiology, endocrinology, nephrology, urology and so on. Each discipline tends to attribute a causal role to the conditions that are members of its own causal field.

This complexity in the causal explanation gives origin to that variability in the causal attribution of the disease that is a typical and common experience for everyone of us. How many times have we received a different causal attribution from different specialists for the same amount of information? Often the cause for this does not depend on the negligence or bad faith of the doctor (confirming the 'fundamental attribution error' of everyday causal attribution in overestimating the role of dispositional factors in controlling behaviour), but rather on the constraints and peculiar view of his particular causal field that obliges him to select only a limited set of potential causal conditions, leaving the others in the background set of enabling or non-causal conditions.

If this difficulty is serious among disciplines that are, in principle, commensurable, the situation becomes dramatic when there is incommensurability between the different causal fields. Think of the causal explanations and the relative therapies given for diseases like maniac-depressive psychosis. Psychosurgery regards it as an abnormal transmission between some parts of the brain; genetics as a hereditary transmission; neurochemistry as abnormal changes in some mediators; social psychiatry as an hostile environment; phenomenological psychiatry as a mistaken existential project; psychoanalysis as the outcome of many possible internal psychodynamic factors, and so on.

To conclude, the research tradition on causal cognition offers interesting suggestions to the philosophy of science to explain the variability of causal explanations and not to allow their explanation to be reduced to irrationality and social interests. It is clear that social and pragmatic factors help to maintain the various scientific research traditions and disciplines separate. But it is also clear that different causal models of the same kind of events do not depend only on social negotiation and bargaining in the scientific community, as some relativist theories of the sociology of science might claim, but instead mainly on the constraints of the human mind and of disciplinary knowledge.

Springer

http://www.springer.com/978-3-642-40215-9

Methodological Cognitivism
Vol. 2: Cognition, Science, and Innovation
Viale, R.
2013, XIX, 447 p. 7 illus., 1 illus. in color., Hardcover
ISBN: 978-3-642-40215-9