Explaining without the Real

Samuel Schindler
Division of History and Philosophy of Science
School of Philosophy
University of Leeds
s.schindler04@leeds.ac.uk

Abstract

The causal-mechanistic account of explanation (CM) is a very intuitive account of scientific explanations. It guarantees objective explanations because it claims that we explain some set of phenomena by referring to the cause, which *produces* the phenomena in question. Yet, this intuitive appeal of the CM account comes at a high prize. Not only does one need to presuppose the reality of the cause, but also is one committed to some form of reductionism. The explanans is more fundamental because it produces the phenomena. In this paper I want to propose an alternative to the CM account which does without the metaphysical baggage the CM account has to shoulder, without compromising on the intuitive appeal of the CM account as an account of distinctively *scientific* explanations.

Keywords: explanation, causal-mechanistic account, kinetic theory of gases, Imaginary-Constitutives, simulation.

Table of contents

1. Introduction

The causal-mechanistic account of scientific explanation (CM) is somewhat intuitively appealing. Scientists explain the world by uncovering its underlying causes and mechanisms which cause or produce the phenomena we observe. The CM account sits well with our notions of objectivity of scientific explanation. A scientific explanation is objective if it uncovers the *true* cause of the phenomena to be explained[1]. This, of course, is contestable. An antirealist like Bas van Fraassen, for instance, denies that scientific explanations have to be true. Explanations are good if they merely provide answers to contrastive why-questions (why did the plane crash instead of landing safely). Explanations are therefore context-dependent and satisfy mere pragmatic desires. A drawback of such a relativist notion of explanation is of course that it loses the intuitive appeal of what it is that makes scientific explanations so particularly successful and objective. In this essay, I want to propose an account of explanation that on the one hand does not compromise on this intuition but on the other hand is not dependent on the truth of the explanans either.

In the next section I want to introduce the philosophical positions about explanation that go under the heading "causal-mechanistic explanations". There are marked differences between the particular positions and some of those positions argue exclusively either for a mechanistic or for a causal account. However, they all share three common assumptions, which I want to challenge in this essay. In Section 3 I shall discuss and criticise the causal-mechanistic account with the paradigmatic example of the kinetic theory, the theory, which has perhaps been most widely used in the discourse about scientific explanation. In Section 4 I want to introduce my account of *Imaginary-Constitutive* (IC) explanation, which does without the heavy metaphysical baggage of the causal-mechanistic account. In Section 5 I shall introduce Einstein's distinction of constructive and principle theories and argue that this distinction is best understood in terms of my IC account of explanation.

2. Heavy Baggage for the Causal-Mechanistic Account

Perhaps the most prominent proponent of the CM account of scientific explanation is Wesley Salmon. Salmon (1984) distinguishes between epistemic, ontic and modal forms of explanations and identifies his own account as an ontic account, explicitly committing himself to a realist account of explanation. According to Salmon, causal mechanisms are characterized by their ability to leave a mark. A baseball, being thrown at a car, *causes* a dent (the mark)[2]. This (rather odd) criterion of mark transmission has been found wanting (cf. Hitcock 1995, Kitcher 1989, van Fraassen 1985). More intuitively, for Salmon, scientific understanding involves "laying bare the mechanisms […] that bring about the fact-to-be-explained" (Salmon 1998, 328) and "[t]o understand the world and *what goes on in it*, we

---

[1] It is not so clear whether the other major accounts of scientific explanation imply strict objectivity. In the DN model any sort of generality can figure (as becomes particularly clear from the examples chosen by philosophers discussing the DN model). Also a unificationary explanation could be any explanation that somehow manages to relate and reduce one class of phenomena to a smaller one. It does not need to be the true explanation. Moreover, it is not at all clear what the necessary link is supposed to be between explanation and unification. A simple classification (like a library classificatory system), in a way, unifies (particular kinds of books under one particular subject label) but does not explain anything.

[2] More recently Salmon has modified his account by substituting the "transmission of a mark" criterion for a criterion of conservation. According to the latter criterion, "a causal process is a spatiotemporal continuous entity which manifests a conserved quantity (such as, for example, linear momentum or angular momentum)" (see Woodward 1979).

must expose the *inner workings*" (Salmon 1984, 133) [3]. Likewise, Machamer, Darden and Craver (2000) define mechanistic explanation thus:

> To give a description of a mechanism for a phenomenon is to explain that phenomenon, i.e., to explain *how it was produced*. (Machamer et al. 2000, 2p.; my emphasis)[4]

Carl Craver (2007) has refined the mechanistic account in his recent book *Explaining the Brain*. In a chapter entitled "The Norms of Mechanistic Explanation", Craver distinguishes between *how-possibly models* and *how actually-models*. The former are merely "heuristically useful in constructing and exploring the space of possible mechanisms, but not adequate explanations". The latter, however, "describe *real* components, activities, and organizational features of the mechanism *that in fact produces the phenomenon"* (Craver 2007, 112). The entities and activities forming a mechanism must "jointly *exhibit the phenomenon* to be explained" (122). The claim that the phenomena must be produced by a mechanism consisting of entities and their interaction can already be found in Ernan McMullin's account of "structural" explanation (1978), which he later dubbed "retroductive" explanation (1984|). To give a structural or retroductive explanation is to explain the "properties" or "behaviour" of an observed system, by referring to the "*constituent* entities or processes and the relationships between them" (McMullin 1978)[5]. Put differently,

> Theory explains by suggesting what might bring about the explananda. It postulates entities, properties, processes, relations, themselves unobserved, that are held to be *causally responsible* for the empirical regularities to be explained. (McMullin 1984, 210; my emphasis)

According to McMullin, we can reveal those "structures" by applying "retroductive explanation":

> This kind of explanation may be called retroductive because it *leads backwards from observed effect to postulated causes.* (ibid.) [6]

Also Cartwright (1983), who never mentioned mechanisms in her work, contended that she believes in the reality of causes, which bring about the phenomena we observe. Also Cartwright believes that it is from the latter that we can infer the former:

> Causes make their effects happen. We begin with a phenomenon which, relative to our other general beliefs, we think *would not occur unless something peculiar brought it about* […] *An effect needs something to bring it about,* and the peculiar features of the effect depend on the particular nature of the cause, so that – in so far as

---

[3] J. D. Trout (2002) criticised Salmon and all the other proponents of the major accounts of explanation for using the notion of understanding as an epistemic virtue and for supposing that a felt understanding is an unequivocal indicator of a scientific explanations proper. Trout quotes episodes of the history of science, where an apparent understanding of a phenomenon turned was blatantly mistaken (e.g. Galen's "understanding" of diseases in terms of his system of four humours). However, Trout claims that explanations have to be true in order to be "good": "*there is no substitute for simply being (approximately) right" (ibid. 230).* Since it is one major aim of this essay to reject the latter view, I take it that Trout's epistemic distinction between understanding and explanation is rather irrelevant.

[4] A main part of Machamer et al.'s paper is devoted to arguing for the irreducibility of both entities and activities in mechanisms: "Entities and activities are correlatives. They are interdependent. An ontically adequate description of a mechanism includes both." (Machamer et al. 2000, 6)

[5] An obvious peculiarity of McMullin's account is that, for him, structures are equivalent with models. This puts him closest to the account that will be introduced below. However, McMullin does join in the CM talk of the "production" of the phenomena and the *inference* of the structure *from* the phenomena etc.

[6] I take McMullin's term "retroductive inference" to be equivalent with the Inference to the Best Explanation. See below.

we think we got it right – *we are entitled to infer the character of the cause from the character of the effect.* (Cartwright 1983, 76; my emphasis)

So although there are very marked differences between the respective approaches of Salmon, Machamer et al., McMullin and Cartwright, what unites them is the belief that the phenomena are *produced* or *caused* by the *real mechanisms* or *real causes* which underlie the phenomena in question. What can also be said about causal-mechanistic account of explanation is that usually, the causes and mechanisms in question cannot directly be observed. According to McMullin,

It is the *penetration of the "invisible realm"* (as Newton called it) that constitutes the real triumph of contemporary scientific explanation. (McMullin 1978., 145; emphasis added)

Also Salmon in a chapter entitled "theoretical explanation" of his book *Scientific Explanation and the Causal Structure of the World* alleges that

One obvious fact about scientific explanations […] is that they frequently appeal to unobserved or unobservable objects […] [and that] entities, processes, and events that are not observed – and that in many cases are not directly observable by the unaided human senses – are invoked for the purpose of explaining some phenomenon. (Salmon 1984, 206)

I shall therefore presume from now on that the standard CM account comprises at least these three assumptions:

i. there is a "hidden" (unobserved or unobservable) cause or mechanism *cm* that
ii. causes or produces the observable phenomena *p;*
iii. we can infer *cm* from *p*.

I shall furthermore presume that causes and mechanisms are equivalent and that therefore their respective predicates of causation and production are equivalent as well. One might want to say that causes are only parts of mechanisms or, to the contrary, that they are epistemically stronger than mechanisms. However, causes and mechanisms are usually treated on a par (hence the term "causal-mechanistic account"). In any case, the necessary clarification is not for me but for the proponents of the CM account. Independently of this, the conclusions reached here are going to be *as* valid because I am not concerned with the inherent details of the CM account but rather with its broad relations with the phenomenological world. See Fig. 1.

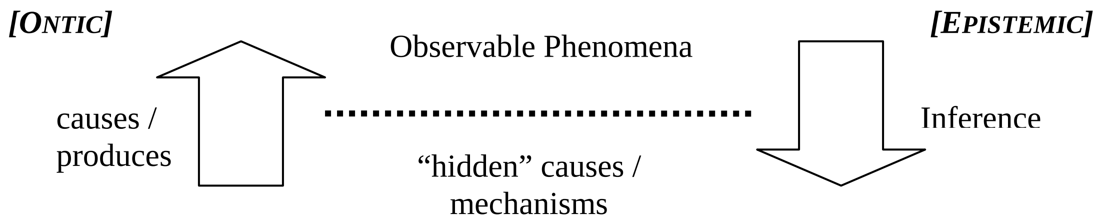

Fig. 1: The causal-mechanistic account

The CM account as defined above seems to be dependent on some form of reduction. We explain a macrophenomenon by appealing to the microphenomenon which we say "produces" the macrophenomenon. Microphenomena therefore must be somehow more fundamental than macrophenomena. There even appears to be some sort of identification being made when we, for instance, talk about other states of matter. The answer we usually provide to the question of "what makes a solid body?" is that in a solid body the molecules are fixed in their position in respect to each other. In a gaseous state these molecules are in motion and interact freely. Short-handedly, we will tend to say that the gas is *nothing but* molecules in motion. Reductionist models have been put forth in different guises by several authors (Nagel 1949, Hooker 1981, Schaffner 1969, Bickle 1998). One should first of all distinguish between intertheoretic reduction (reduction of theory T to theory T*) and what Craver (2005) calls microreduction, i.e. the reduction of phenomena of higher levels to phenomena of lower levels (see also Aerts and Rohrlich 1998). Although these two sorts of reductionism sometimes are not clearly separable (think of the reduction of thermodynamics to statistical mechanics and the implied reduction of macrophenomena like temperature to microphenomena like molecular motion) it is of course microreduction rather than intertheory reduction, which is most relevant in the context of CM. Notice that this reduction of macro- to microphenomena is non-eliminative: the reality of the macrophenomena is not denied, however, microphenomena are clearly more important and fundamental than their macrophenomenological counterparts, since the former *cause*, i.e. bring into existence the latter.

So as it appears CM carries heavy baggage. Not only is CM committed to realism about the causes and mechanisms (including all the entities and relations holding between them) but also to some sort of reductionism. In contrast, I want to propose an alternative to the CM model of explanation which is both non-committal to the truth of explanations and non-committal to reductionism without having to compromise on the intuitive appeal of the CM account. In the next section I shall discuss the most widely used example in philosophical discourse about scientific explanations – the kinetic theory of gases. With this example I shall exemplify the CM account and point towards its shortcomings.

3. Exemplifying Explanation: the Kinetic Theory of Gases

Ever so often, philosophers of science, when theorizing about scientific explanation, have used non-scientific examples (the flagpole, the barometer, and John’s failure to get pregnant are cases in point)[7]. One might ask what the relevance of such contrived examples is supposed to be when we want to understand *scientific* explanations as exercised in actual scientific practice. However, philosophers of science have failed to provide proof or at least reasons as to why our intuitive and common sensical notions of explanation should bear any relevance on *scientific* explanations. Notwithstanding this observation, DeRegt (2000, 152) noticed that the kinetic theory of gases is often quoted as a paradigm case of scientific explanation[8]. Salmon (1984) has used the kinetic theory to support his causality account, Friedman (1974) to illustrate his unification account and Cushing (1991) to make a point for his claim that visualizability is a necessary condition for understanding[9]. So without a doubt, philosophers of science seem to ascribe some import to the kinetic theory for our understanding of scientific explanation in the context of explanation. Of course, philosophers of science have looked at the kinetic theory and its explanatory potential from very different angles depending on their theoretical predispositions. Just briefly, Salmon (1984, 229p.) has claimed that molecules “qualify as causal processes” and that each molecule – in line with his

[7] Woodward (1979) made the same point already 38 years ago. It is still valid. DeRegt and Dieks (2005, 162) noticed recently that philosophers theorizing about scientific explanation rely on their intuitive judgments and employ their intuitions as “basic facts”, which I regard as equally misleading.

[8] In fact, the kinetic theory of gases is used in a multitude of contexts of philosophical analysis. Others contexts apart from explanation are reduction, models (see Schindler 2006), analogies (Hesse 1966; Schindler 2006), and even the positivists used it as an illustration of their idea of correspondence rules (see Hempel xxx; Carnap xxx). In analogy, one could say that the kinetic theory for the philosophy of science is something like E. coli for microbiology.

[9] Visualizability is an almost neglected topic in the literature on scientific explanation and understanding. Henk deRegt’s work (e.g. 1997, 1999, 2001) is an exception. See also footnote 24.

marking criterion for causal processes – can be marked (e.g. radiation putting molecules into exciting states). Cushing (1991, 344p.) contrasts the derivation of Boyle's law from statistical mechanics using abstract notions like energy, entropy, equilibrium conditions of the system, with the derivation from kinetic theory, *picturing* molecules as perfectly elastic point particles. Both derivations are perfectly valid and yet, only the second one – according to Cushing – gives us an understanding. Friedman (1974, 14) has directed his attention to yet another aspect of the kinetic theory. Friedman claims that it is not the "reduction" of Boyle's law to molecular motion, which gives us understanding (as the proponents of CM and reductionists would claim). Rather, it is the fact that the kinetic theory relates previously unrelated gas laws with each other. In other words, it brings about a *unification* of empirical laws. I do not disagree with those characterisations. Yes, the kinetic theory does bring about unification. Yes, the kinetic theory does give us an understanding thermodynamics doesn't[10]. What about pictures, which Cushing holds to be necessary constituents of explanation? We shall get to that question later. Finally, I – as I think would the majority of philosophers of science – do feel an intuitive sympathy with the CM account of how the kinetic theory manages to provide explanations. The CM account – as defined above – would have it that the kinetic theory explains empirical laws because it offers a *mechanism* or *cause* in the guise of molecular motion that *produces* or *causes* the phenomena figuring in those laws. But then, if we don't want to doubt the reality of the phenomena – which we of course don't – the CM account *requires* us to likewise accept molecular motion as being real. If it wasn't, it surely couldn't *produce* the phenomena we take to be real.

Let me go into some detail of how the kinetic theory explains Boyle's law. Boyle's law states that whenever we compress a gas in some sort of container, the pressure of that gas will increase (and vice versa). This is explained in the kinetic theory thus: the kinetic theory *presumes* that heat is equivalent to molecular motion. It also *presumes* that gas pressure (the macrophenomenon) is caused/produced by molecules (or merely extensionless particles for that matter) hitting the wall of the container. Now Boyle's law is explained by an increase of the *frequency of particles hitting the wall of the container*. As I said above according to the CM account the kinetic theory explains Boyle's law because it refers to the *true* components of gases, namely molecules. However, there are two problems with this picture. First, nobody will contend that the kinetic theory (or the billiard ball model, as it is widely known today) is very idealised and simplified. Molecules with a finite extension, quantum mechanical properties, and intermolecular forces holding between them, i.e. what we take to be *real* molecules, do not figure in the kinetic theory. Rather, what we are dealing with in the kinetic theory are extensionless, perfectly elastic particles. That is all what's needed in the kinetic theory for explaining Boyle's law. And even though we do know about those gross oversimplifications, this is still the standard explanation of Boyle's law in our textbooks. So here CM runs into the first difficulties. Why do we still use this utterly simplified model if, according to the CM account, a good explanation is an explanation which is true? How does the billiard ball model manage to explain anything *at all*, given all the gross oversimplifications making it literally wrong? Even more worrisome, as Michael Strevens (forthcoming, 2) points out, this explanation which makes so many idealising assumptions, is in fact "*explanatorily optimal*", i.e. "it cannot be improved". Strevens (forthcoming) compared the idealised explanation of Boyle's law – what he calls the textbook explanation – with he calls the canonical explanation of Boyle's law. The latter would include all the characteristics of a real gas, which in the textbook explanation of Boyle's law are left out. It does take into account intermolecular forces, it does assign a finite extension to molecules, etc. (see Strevens forthcoming, 21pp., for more details). After considering every single one of

[10] In fact, this insight is of course nothing but what underlies Einstein's distinction between constructive and principle theories. I shall introduce this distinction below in more detail.

the assumptions of the canonical explanation of Boyle's law and removing those, which have to be deemed irrelevant for the explaining Boyle's law, he quite surprisingly ends up with the textbook account of Boyle's law. Strevens concludes that

> [O]n the subject of collisions [for instance], the canonical model and the idealized model say *precisely the same thing*: collisions make no difference to Boylean behavior, and thus are irrelevant to the explanation of Boyle's law. *The idealized model, then, adopts the optimal explanatory policy on collisions.* (25)

In other words, even though we *could* give a full-fledged canonical explanation of Boyle's law with all the real properties of gases that are known today, it would not make any sense to do so, because the idealised version is everything you need for explaining Boyle's law. All the assumptions contained in the canonical explanation but not in the textbook explanation just don't make any difference for explaining Boyle's law.

The second point I would like to make about the explanation provided by the kinetic theory is that the link it establishes to the world has to be deemed arbitrary. As mentioned above, it is an *assumption* of the kinetic theory not only that temperature is equivalent with molecular motion, but also that the underlying phenomenon of pressure is caused by molecules hitting the wall of the container. This assumption is by no means necessary. This becomes most apparent if we compare the explanation of Boyle's law provided by the kinetic theory with the explanation given by its predecessor, namely the caloric theory of heat.[11] The caloric theory of heat, quite differently from the kinetic theory of heat, makes the following relevant assumptions. First, temperature is not conceived of as molecular *motion* but rather as a separate substance. This substance, again, consists of tiny particles which are mutually repulsive but attracted by material particles. This results in the process depicted in Fig. 2.

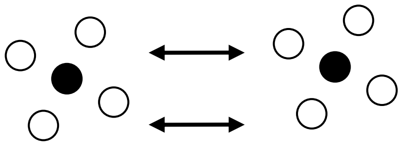

Fig. 2 The caloric theory. Caloric particles (white) repel each other but are attracted by matter particles (black).

Now, Boyle's law is explained thus: by decreasing the volume of a gas, the repulsive forces between the caloric particles – given the reduced space between them – will gain efficiency, i.e. the repulsion between the particles will increase, causing an increase in macrophenomenological pressure. Therefore pressure, in the caloric theory, is taken to be equivalent to the mutual repulsion of caloric particles. So we see that the link that was made in the kinetic theory between pressure and molecular collisions with the wall of the container is by no means the only possible nor a necessary link (see Tab. 1 for a comparison).

Theory	"Mechanism"	Empirical regularity	Explanation
Kinetic Theory	Collisions of molecules with the container wall	V↓ ⇒ P↑ (Boyle's law)	V↓ ⇒ increase of collisions
Caloric Theory	Caloric particles	see above	V↓ ⇒ increase of "effectiveness" of caloric repulsions

[11] The caloric theory of heat has figured most predominantly in discussions surrounding the Pessimistic Meta Induction in recent years (see Psillos 1999). Psillos, for instance, alleges that the term 'caloric' never referred and did not do any predictive nor explanatory work in the theory of heat. However, this picture has been strongly contested by e.g. Chang (2003) and Elsamahi (2004), who have pointed out that the caloric theory was indeed a rather empirically and explanatorily successful, i.e. 'mature' theory See also the standard works about the caloric theory by Fox (1971) and Lilley (1948).

Tab. 1: Comparison of the explanations of Boyle's law provided by the kinetic theory and the caloric theory respectively.

Of course, we take the link made within the caloric theory to be wrong. However, it must be realised that despite this wrong link, the caloric theory still manages to provide an explanation of Boyle's law. And not just any explanation. Rather, it is an explanation that has to be considered to be on *a par* with the one provided by the kinetic theory of heat. Although the causal mechanisms of the kinetic and the caloric theory are very different, they seem to be equal *in kind*. It is hard to say in what sense they should differ (see again Tab. 1 for illustration). Both explanations postulate unobservable entities and unobservable relations (or rather interactions) holding between them. Both sets of entities and relations in the respective accounts can be said to *cause* or *produce* the observable phenomena. So if that is the case, the CM account is in trouble. The CM account assumes that a particular phenomenon can only be caused or produced by the *real* mechanism or cause. If we got two equally good explanations for the same phenomenon, the CM account cannot tell us why they both are good explanations of the phenomena. After all, they cannot be both true. One might want to say that the CM account is not challenged at all because the kinetic theory provides the only true explanation. The explanation of the caloric theory is just false and therefore has not to be considered. This would not only be outspokenly Whiggish but it also would be turning a blind eye on the quite obviously equally good explanation by the caloric theory (see again Tab. 1). Furthermore, as I have argued with Strevens (forthcoming) in this section, the kinetic theory does *not* provide the true explanation (which would involve all the true properties of molecules). However, even if one were to do such a move of claiming (approximate) truth for the kinetic theory, never the less, this would be merely begging the question. After all, we are trying to find out what constitutes a scientific explanation. Merely pointing out that one explanation might be true will not explain to us why another explanation of the same phenomena is not only possible but even equally good.

Of course, the kinetic theory should be deemed the better theory after all because it can account for *more* phenomena than the caloric theory, like for instance the production of heat through friction. However notice that in the latter case the explanation of the kinetic theory is not "better". Rather, the caloric theory cannot account for it at all. On the other hand – as I argued above – the explanation of Boyle's law provided by the caloric is *on a par with* the one by the kinetic theory. So the kinetic theory being "better" does not seem to have anything to do with the *quality* of the explanation it provides but rather with its *scope*. But since it is the nature of explanation rather than their scope we are concerned with in this essay, I take it that the latter does not need to be regarded in this essay.

Given the conclusions reached in this section that the explanans needs not be the true cause (as in the kinetic theory), that Boyle's law can be explained equally well by the kinetic, and the caloric theory and that both explanations do not seem to be different in kind, the causal-mechanistic account seems to be utterly inappropriate. In the next section I therefore want to suggest that the explanans-explanadum relation should be conceived of in rather different terms from the usual talk of causation or production.

4. Imaginary-Constitutives, Correlation, and Simulation

Let me summarise the conclusions reached in the last section: rather than saying that the phenomena are *produced* or *caused* by the *underlying* (true) mechanism, we should speak of a *postulated* system which is *arbitrarily* linked to the observable world. Now, these two conclusions on their own result in either relativism or pragmatism about explanation a la van Fraassen. Yet, either I deem to be rather undesirable. Furthermore, based on these two assumptions it remains a mystery why we have a strong *intuition* about molecular motion *causing* or *producing* macrophenomena. How can we possibly reconcile our intuitions with the conclusions reached in the last section? Since we have ruled out causation or production to be *actual* it has to be something else, which somehow gives us the impression of causation/production. Let me make the following suggestion: rather than a mechanism or a cause *producing* the phenomena or empirical regularities, we should rather speak of *re-production* or *simulation*. It then becomes inappropriate to cling to the traditional notion of causes and mechanisms, which is why I want to introduce the notion of *Imaginary-Constitutives* (*IC*s).

ICs are defined as entities which possess two paradoxical properties: they are both unobservable and visual at the same time (see Schindler 2006). The first property is an assigned ontic property, the latter characterises the way we cognize about them. That is, the first property refers to the fact that, in the real world we cannot possibly observe those entities. The latter property refers to the fact that despite the first property, we cognize about exactly those entities by visualising them. It is useful to distinguish between IC elements and IC events. Applied to our cases of the kinetic and caloric theory ICs *elements* refer to the postulated entities like particles in motion and caloric particles, respectively. IC *events* refer to molecules hitting the wall of the container and the repelling caloric particles respectively. Both, IC elements and IC events are *correlated* with real world entities. The IC elements of particles in motion and caloric particles are correlated with the observable phenomenon of temperature, and the IC events of molecules hitting the wall of the container and repulsive caloric particles are correlated with the observable phenomenon of pressure. In other words, IC systems are closed systems which make contact to the world only by arbitrary correlations. Phenomena are not part of IC systems and can thus not be construed as consequences of IC elements. Let me therefore introduce the following two assumptions of my IC account of scientific explanation:

i. Observable phenomena are *'correlated'* with an postulated IC system;[12]
ii. The IC system *explains* a real world regularity holding between the observable phenomena, if it is capable of *simulating* it;

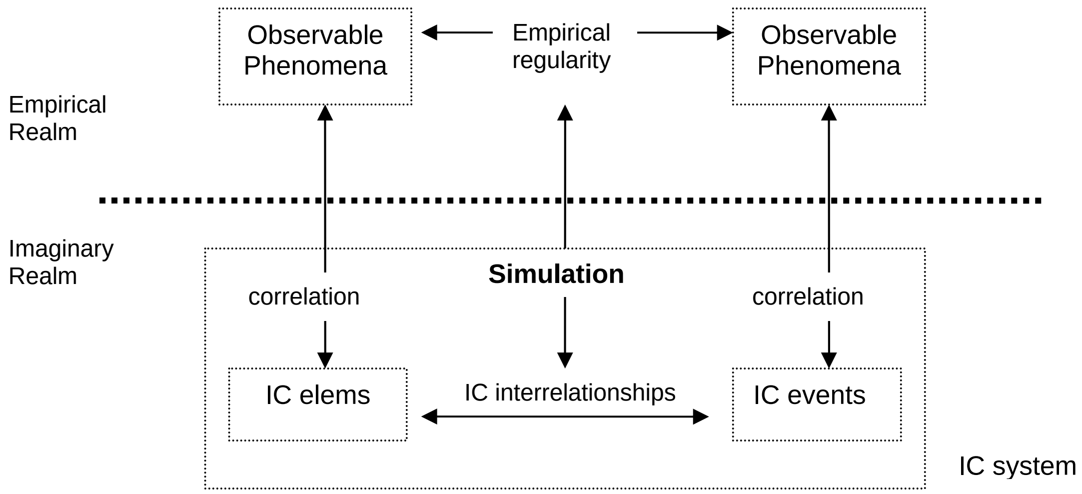

Fig. 3: IC simulation

[12] Correlations bring to mind one of the central tenets of logical positivism. However, correspondence rules have nothing to do with what I'm proposing. Whereas the logical positivists' correspondence rules hold between observational statements and a *formal* system, my correspondence relationships hold between observable phenomena and non-formal entities (ICs). Furthermore, in contrast to the positivist notion, which conceives meaning to be conveyed to the calculus *by observables*, the converse asymmetrical relationships holds between ICs and observable phenomena: unobservable ICs give meaning to observables.

Tab. 2 lists the details of IC simulation of Boyle's and Gay-Lussac's law by the kinetic and the caloric theory respectively. From this table it becomes clear that the empirical regularities are *explanatorily underdetermined*: they can be simulated by different IC systems equally well[13].

Correlation		**Simulation**	
Phenomenon	***IC***	***Empirical regularity***	***IC***
P	Collisions of molecules with the container wall	V↓ ⇒ P↑ (Boyle's law)	"IC volume"↓ ⇒ collisions ↑
see above	Caloric particles repelling each other	see above	"IC volume" ↓ ⇒ "effectiveness" of caloric repulsions ↑
T	Molecular motion	T↑⇒ P↑ (Gay-Lussac law)	molecular motion ↑ ⇒ frequency of collisions ↑
see above	Caloric particles	see above	caloric particles ↑ ⇒ mutual repulsion of caloric particles ↑ [14]

Tab. 2 Simulation of empirical regularities by means of IC systems. P=pressure, V=volume, T=temperature, ↑ = increase, ↓ = decrease. Two empirical gas laws as explained by the kinetic theory and the caloric theory.

Let it be emphasised again that both IC elements and IC events are *strictly* unobservable. This is not necessarily the case in the CM account. Causes can be observable. On the other hand, in order for ICs to do any explanatory work, they must be visualised. In a non-visual form, the container of a gas would reduce to a black box. As such there couldn't be any talk about its imagined constituents let alone IC simulation. Furthermore, particularly with the caloric theory it becomes clear that ICs are not visualisations *about* anything. Pictures like the one in figure Fig. 2 do not visualise any real world entities. This, of course, will readily be granted. However, if one concedes that the explanations of the caloric theory are on a par with the ones provided by the kinetic theory, then – I want to suggest – the visual nature of the explanans of the kinetic theory should not be conceived as a visualisation *of* anything either. ICs are autonomous systems, with their visualisability being an intrinsic property.

It is worth pointing out that IC simulation is not to be subsumed under those sorts of simulations that merely aim to get the data right. In a computer simulation of, say, playing chess no claims are being made that the used programme and algorithms have anything whatsoever to do with the actual "mechanisms" in human brains needed for a human being to play chess. Notice that, in contrast to IC systems, in these sorts of models no entities are postulated. All what counts in non-IC-simulations is that successful predictions and/or

[13] Explanatory underdetermination (EU) is distinct from the traditional Underdermination of Theories by Evidence (UTE). UTE is a thesis about the *predictive* empirical consequences of theories. EU is much harder to achieve. It requires the *qualitative* simulation of empirical regularities, not just the *quantitative* prediction of the phenomena.

[14] Gay-Lussac's law (sometimes also referred to as Charles' law) is explained by the caloric theory thus: the pressure of a gas increases with an increase in temperature because an increase in temperature, according to the caloric theory, is equivalent to the influx of caloric particles. Since caloric particles repel each other, they will increase the interstices between the gas molecules.

behaviour can be generated *quantitatively*. In IC systems, on the other hand, empirical regularities are simulated in a purely and in an immediately comprehensibly *qualitative* sense (see above). In the CM account, on the other hand, it is not at all clear whether the production of the phenomena should be understood in a qualitative or quantitative sense. Can we speak of causation, when we can give an account of the phenomena in qualitative terms? Or do the phenomena have to be produced quantitatively by their cause? How are those forms of causation to be understood anyway? In the IC account of explanation, these problems do not arise. But how do we get from a IC explanation, which is explicitly qualitative to a quantitative prediction of the phenomena? As a matter of fact, quantitative predictions are parasitic on qualitative explanations. In the kinetic theory the phenomenon of pressure is quantitatively derived as being equivalent to the average kinetic energy per molecule. This is how this derivation is presented in a standard textbook: Imagine a cubic container with the two opposite sides A and B and edges of the length *d*. Due to the postulated elasticity of the particles a particle hitting the wall of the container with velocity v_x (along the x-dimension) will rebound with the momentum $-2mv_x$. The force F imparted on the wall by a single particle on the wall is:

$$F = \frac{mv_x^2}{d}$$

The total force for all particles in the container is:

$$F = \frac{Nm}{d}\overline{v_x^2}$$

where N is the total number of particles and v bar the average velocity of a single particle. Under the assumption that the motion of the particles is completely random, i.e.

$$\overline{v^2} = 3\overline{v_x^2}$$

this yields:

$$F = \frac{N}{3}\left(\frac{m\overline{v^2}}{d}\right).$$

Then, the total pressure (defined as $P = F/d^2$) is:

$$P = \frac{2}{3}\left(\frac{N}{V}\right)\left(\frac{1}{2}m\overline{v^2}\right).$$

That is, the pressure is proportional to the number of molecules per unit volume and to the average translational kinetic energy of the particle[15]. We can re-write the equation as

$$PV = \frac{2}{3}N\overline{KE}$$

and thus relate the phenomenon of pressure to the derivations of our model. If we now remind ourselves of the ideal gas law ($PV = nRT$) and assume the theoretically derived

[15] The translational kinetic energy is the last term on the right hand side of the equation.

equation to be equal to the experimentally derived ideal gas law we can combine these two equations to give

$$T = \frac{2}{3R}\overline{KE}\,.$$

In this manner we have related the phenomenon of temperature to the average molecular kinetic energy. We have thus derived the molecular equivalents of both pressure and temperature. However notice that a fundamental *presupposition* of these derivations is the *qualitative* picture of molecules bouncing against the walls of the container, i.e. an IC event. This picture one can translate into mathematical language and write down the change of momentum of a molecule as *2mv* after the collision with the wall, which is then the *basis* of the *quantitative* derivation of pressure. In other words, quality comes before quantity. The IC is therefore constitutive of the derivation of the quantitative predictions of the theory. De Regt and Dieks (2005) made this point when they recounted Ludwig Boltzmann's introduction of the kinetic theory in his *Lectures on Gas Theory* (1964), where Boltzmann discusses Boyle's law on merely qualitative grounds, as we have done already above. De Regt and Dieks point out that

> [i]t is important to note that the above reasoning [about Boyle's law] does not involve any calculations. It is based on general characteristics of theoretical description of the gas. Its purpose is to give us *understating* of the phenomena, *before* we embark on detailed calculations. Such calculations are subsequently motivated, and given direction, through the understanding we already possess. (de Regt and Dieks 2005, 152p.; original emphasis)[16]

Let me conclude this section. The CM account assumes that there are causes producing the phenomena. I have pointed out in the last section that the CM account is therefore committed to a realism about the causes and some sort of reductionism. The CM account also assumes that the causes producing the phenomena can be *inferred* from the phenomena. Nobody has ever managed to spell out what this is supposed to mean. My account of IC explanation, in contrast, accomplishes what the CM account does not: it provides a clearly spelled out and detailed account of how the explanans of explanations relates to the explanandum in terms of correlations between phenomena and ICs and in particular in terms of IC simulation of empirical regularities. Furthermore, in contrast to the CM account, where it is not at all clear whether the phenomena have to be produced by their causes by a qualitative or quantitative sense, in the IC account empirical regularities are clearly simulated in a purely qualitative sense. I also suggested – agreeing with de Regt and Dieks (2005) – that quantitative prediction is dependent on a qualitative prediction of the phenomena, being equivalent to IC explanations[17].

[16] De Regt (2001) has even put this as a so-called Criterion for the Intelligibility of Theories (CIT), which requires that a theory T must allow the scientist to "recognise *qualitatively* characteristic consequences of T without performing exact calculations" for T to be intelligible (261).
[17] Of course, I do not suggest that *all* predictions are dependent on qualitative predictions. A large amount of modelling does do without qualitative IC predictions. However, if we want to *understand* the predictions our models make, there is no alternative for IC explanation/prediction.

5. Two Kinds of Theories: Construction, Explanation and Visualisability

One might wonder how general the account proposed in this paper is intended to be. I want to answer this question by referring to Einstein's distinction between 'principle theories' and 'constructive theories'[18]. Constructive theories, Einstein characterises thus:

> They attempt to build up a picture of the more complex phenomena out of the materials of a relatively simple formal scheme from which they start out. (Einstein 1935, 167)

Principle theories, in contrast,

> […] employ the analytic, not the synthetic, method. The elements which form their basis and starting-point are not hypothetically constructed but *empirically discovered ones***,** general characteristics of natural processes, principles that give rise to mathematically formulated criteria which the separate processes or the theoretical representations of them have to satisfy. (ibid.)

Interestingly, Einstein quoted the kinetic theory as a prime example for a constructive theory, tying in very nicely with our previous discussion. The kinetic theory, as an instance of constructive theories, *postulates* that a gas consists of perfectly elastic particles. From this assumption one *constructs* or derives empirical laws like the ideal gas law. On the other hand, thermodynamics, Einstein's example of a principle theory, does not postulate any entities. It just states, on the basis of observations, that particular events – like perpetual motion – are impossible and from those derives that other observational states are necessarily so. Most importantly Einstein says that

> When we say that we have succeeded in understanding a group of natural processes, we invariably mean that a constructive theory has been found which covers the processes in question. (ibid.)

Principle theories, on the other hand, do not do any explanatory work[19]. I therefore want to constrain IC explanations to Einstein's class of constructive theories. In fact, I'm prepared to claim that the explanations provided by Einstein's constructive theories are nothing but IC explanations. Not only does Einstein's distinction save me from having to account for every single theory out there but also does my IC account of explanation then provide a nice explication of Einstein's class of constructive theories. I take Einstein's distinction to be an epistemological, not an ontological distinction. Constructive theories are not theories, where some causes literally (i.e. ontologically) produce "complex phenomena", like in the CM account. Rather, – and I think this is quite clear from Einstein's own phrasing – Einstein's distinction is an epistemological one. Constructive theories "build up a *picture*" of the phenomena from a "simple formal scheme". In the language of ICs that simply means that in constructive theories, IC systems are constructed that are able to simulate particular empirical regularities.

It is unknown that Einstein's constructive / principle theory distinction was foreshadowed by an essay called "Outlines of the Science of Energetics" by the physicist and engineer William John Macquorn Rankine. According to Rankine, there are two ways of constructing physical theories. According to the "abstractive method"

---

[18] Einstein's distinction has been discussed recently in the context of controversies surrounding substantivalist vs. relationist interpretations of spacetime in general relativity (see Brown and Pooley 2004, Brown 2005, Balashov and Janssen 2003).

[19] Still, principle theories are not to be confused with empirical laws. The conservation of energy, for instance, goes well beyond a mere *descriptive* generalisation, having tremendously broader scope than lower level empirical laws.

a class of objects or phenomena is defined by describing, or otherwise making to be understood, and assigning a name or symbol to, that assemblage of properties which is common to all the objects or phenomena composing the class, *as perceived by the senses, without introducing anything hypothetical.* (Rankine 1881, 210; my italics)

That is, physical theories arrived at with the "abstractive method", very much like Einstein's principle theories, do not introduce "anything hypothetical" but merely classify phenomena "as perceived by the senses" by assigning "names" or "symbols" to a particular set of properties common to a set of phenomena. In contrast, in the "hypothetical method"

a class of objects or phenomena is defined, *according to a conjectural conception of their nature, as being constituted, in a manner not apparent to the senses* […] Should the consequences of such a hypothetical definition be found to be in accordance with the results of observation and experiment, it [i.e., the hypothetical definition] serves as the means of deducing the laws of one class of [observable] objects or phenomena from those of another [class of unobservable objects]. (ibid.)

Hypothetical theories, therefore, very much in contrast to abstractive or principle theories, but very much like Einstein's constructive theories, make hypothetical conjectures about what is beyond what can be observed. From those conjectures – like in Einstein's constructive theories – one can deduce empirical consequences, which should be "in accordance with the results of observation and experiment". Although explanatory power does not figure in Rankine's dichotomy, Pierre Duhem, who was very much influenced by Rankine's dichotomy, referred to Rankine's abstractive theory as "representative" and to his hypothetical theory as *explanatory* theory (cf. Duhem 1962, 52) [20]. A large part of Duhem's classic *The Aim and Structure of Physical Theory* is dedicated to deriding explanatory theories and promoting the benefits of representative theories:

While we regard a physical theory as a hypothetical explanation of material reality, we make it dependent on metaphysics […] A physical theory is not an explanation. It is a system of mathematical propositions, deduced from a small number of principles, which aim to represent as simply, as completely, and as exactly as possible a set of experimental laws. (Duhem 1962, 19) Concerning the very nature of things, or the realities hidden under the phenomena we are studying, a theory conceived on the plan we have just drawn teaches us absolutely nothing, and does not claim to teach us anything. (*ibid*., 21)

Duhem clearly distances himself from those theories, which hypothesize about the 'hidden realities' in order to provide explanations and locates these endeavours in the realm of metaphysics. A proper physical theory, for Duhem, does not attempt to explain the world by some "hidden causes". Physical theories rather are mere formal and abstract systems, whose only function is it to represent experimental laws, by classifying them. Duhem also notably associates non-abstract theories (which do hypothesize about the hidden realities) with (feeble English) "*visualizing* minds". Clearly then, Duhem appears to associate visualisation with explanation and deems both to be malignant. Of course, Duhem's distinction between proper physical (i.e. abstract) theories and deceptive (i.e. concrete) theories including "mechanical models" clearly mirrors Rankine's distinction between hypothetical and abstractive theories[21].

---

[20] Also Ernst Nagel mentions Rankine's distinction in his classic *The Structure of Science* (1961, 125pp.). Rankine's notion of abstractive theories and its influence on William Thomson and James Maxwell is analysed in Moyer (1977).

[21] Nancy Cartwright (1983), in turn, follows Duhem in claiming that theories are abstract entities, which classify and systematize phenomenological laws. However, in contrast to Duhem, Cartwright does attribute explanatory power to theories.

What Duhem derided, steps to the fore in James McAllister's theory dichotomy, which I take to be congruent with the dichotomies discussed above: visualisability. According to McAllister

some theories *postulate visualizable structures or mechanisms that are said to underlie phenomena*. By visualizable structure or mechanism, I mean one of which there is a *mental image*, drawn typically from everyday experience, *that guides our understanding* of the nature or dynamics of the phenomena. (McAllister 1999, 48pp.; added emphasis)[22]

This is only one type of theory.

Other theories give accounts of empirical data that are not pictorial but abstract. An abstract theory does not evoke a mental image: rather, it *describes* phenomena by means solely of a mathematical or other formal apparatus. (ibid., added emphasis)

Again, like Einstein, Rankine and Duhem, McAllister draws a distinction between theories that postulate unobservable entities and other theories that don't[23]. McAllister – like Duhem – deems visualisablity to be an important feature of theories that are explanatory. In contrast to Duhem, of course, he does not take visualisability and explanation to be a malignant. Abstract theories, on the other hand – like Einstein's principle theories – merely "describe" the observable phenomena. McAllister emphasises that these respective characteristics are not "inessential" or "eliminable". Rather, they appear to be irreducible and genuine properties. In particular, the property of being visualisability McAllister holds to be indispensable not only for explanatory but also for "heuristic power":

[A] theory that refers to a visualizable mechanism cannot generally be reduced to a purely abstract theory without the loss of some *explanatory or heuristic power*: it frequently turns out that the visualization that such a theory puts forth plays a role in *generating the theory's explanations* of phenomena or in showing *how the theory should be applied or further developed*. (McAllister 1999, 50; added emphasis)

I have discussed the heuristic power and their "fertility" of IC theories somewhere else (Schindler forthcoming). However, let us note that, in line with McAllister's emphasis on visualisation, James Cushing (1991) famously asserted that visualisability is pivotal for our understanding and that quantum mechanics can therefore not be understood *in principle*:

The argument presented here actually begins from the intuition, based on experience and on (some) history of physics, that *understanding of physical processes must involve picturable physical mechanisms and processes that can be pictured* (Cushing 1991, 341).

Why should this be the case? Why should our understanding be dependent on visualisations[24]? What in McAllister and Cushing's account lacks justification (apart from

[22] In this quote, McAllister remarks that the "mental image" is "drawn typically from everyday experience". I take this to refer to analogical reasoning. For details about models and analogies see the locus classicus Hesse (1966) and for a criticism of the latter see Schindler (2006).
[23] McAllister does not say that explicitly but I infer that from his remark that they "are said to underlie phenomena".
[24] For the contrary claim see de Regt (2001). Yet, de Regt concedes at the end of his paper that "the historical analysis in the preceeding sections has made it clear that visualisation often plays a role in rendering scientific theories intelligible" (260). The visualizablility or picturability of scientific explanations has been a neglected topic. Among the very few discussions of visualizablility De Regt (1997) has pointed out that Erwin Schrödinger, much influenced by Ludwig Boltzmann's *Bildtheorie*, believed that scientific explanations have to be *anschaulich* (i.e. visualizable). DeRegt argued that both Schrödinger and Boltzmann (deRegt 1999) distinguished between representative and explanatory aspects of those (mental) pictures. Neither Schrödinger nor Boltzmann apparently believed in the reality of these pictures but were very convinced that they were

intuition and personal experience) gains a rationale with ICs: ICs are *intrinsically* visual systems.

So not only do ICs explain what Einstein had in mind with constructive theories – they construct IC systems capable of simulating empirical regularities – but also, my IC account provides an answer as to why the explanans in constructive theories should be visual(-isable). Construction, explanation, visualisability. These features are unified by my IC concept. Without it, they can be only weakly associated on the base of personal flavour and intuition.

6. An impoverished IC explanation: Newtonian Force

Newtonian or any other force seems to contradict my claim that the entities figuring in the explanans of scientific explanation always have to be unobservable *and* visualizable. Forces are unobservable but not visualizable. However, this apparently contradictory case can be turned into a confirmation of my account as I shall outline in this section.
Philosophers of science are divided about the explanatory content of Newton's theory. Whereas some claim that the theory is non-explanatory, others claim the opposite. Furthermore, it is well known that not only was Newton's concept of gravity only very reluctantly taken up by Newton's contemporaries when first proposed in the *Principia*, but also Newton himself was rather unhappy with the idea of "action at a distance".

It is inconceivable that inanimate matter should, without the mediation of something else which is not material, operate upon and affect other matter without mutual contact. (Newton 1961, 253-254)

As Ernan McMullin has pointed out about Newton's theory,

[i]f the new mechanics were accepted, one either had to allow action at a distance or else introduce new and illegitimate-seeming modes of bridging the spaces between planet and sun. (McMullin 1989, 294)

Bridging the spaces between planets and sun meant that one had to posit some form of aether (either material or spiritual), which would allow to explain gravity in terms of contacts between the ether molecules. Because such a theory would have been built from these (molecular) elementary foundations it would have represented a genuine "constructive" theory, in the sense discussed above[25]. However, Newton had shown in book II of his *Principia* that the idea of a material aether was not tenable because it would disturb the motion of the planets and would cause them to spiral inwards. Although Newton did explore the possibility of an immaterial aether, in the *Principia* Newton refrained from speculating about this mystic medium[26]. Therefore Newton has been often quoted as being *the* forerunner of an instrumentalist approach towards science. Newton's "hypothesis non fingo" is the most prominent case in point, to be understood as "I do not speculate about the causes of gravitation". And yet, as McMullin points out, Newton can not be seen as an instrumentalist who was happy to accept that his theory was merely a neat device providing for empirical adequacy but not much more. In other words, Newton was not a modern Ptolemy, whose theory was constructed in order to just "save the phenomena" (of course, van Fraassen 1980

---

necessary in order to evoke understanding. I not only share the belief that visualizablility is a necessary condition of scientific explanation but I also follow Schrödinger and Boltzmann in refraining from identifying our mental pictures with the furniture of the world.

[25] One might here be reminded of Harvey Brown's claim that the theory of general relativity is not a constructive theory and would have been only such, if one were able to work out the behaviour of clocks and rods (length contraction, time dilation) in terms of a micro-theory, namely Quantum-mechanics. See Brown…

[26] Newton did however carry out this speculation in private and most notably did so in the Scholium attached to the Principia. Cf. McMullin 1989, 301p., Iliffe (2004), which contains many relevant references.

notoriously holds this to be the aim of science). McMullin bases this claim on the following quote by Newton:

Hitherto we have explained the phenomena of the heavens and of our sea by the power of gravity, *but have not yet assigned the cause of this power.* This is certain, that *it must proceed from a cause that penetrates to the very center of the sun and planets*, without suffering the least diminution of its force. (Newton; citied in McMullin 1989, 291; my emphasis)

Clearly, Newton's hope to find "a cause that penetrates to the very center of the sun and planets" is irreconcilable with a mere instrumentalist approach. On the other hand, Newton of course never found this cause. Newton therefore considered his theory to be incomplete. If we were to try to classify Newton's theory as either a principle theory or a constructive theory where would we put it? I would like to argue that what we have got in Newton's theory appears to be an impoverished IC, as it were. On the one hand the theory postulates an *unobservable* force, which accounts for the attraction between the bodies of the universe, the swing of the pendulum, the change of the tides and other phenomena. On the other hand, though, the concept of force has to be taken as a primitive. One cannot go one level deeper and explain the attraction of the planets in terms of (ether) molecules propagating motion. Now notice that, *if* one could give such an explanation (within a Newtonian framework), the explanans would be *visualizable*. Therefore, I want to claim that in Newton's theory only the necessary but not the sufficient condition of the IC is present: although an unobservable is postulated, it is not visualizable. Because this is the case, the theory lacks the degree of intelligibility proper IC theories possess. Newton (and other people) would even go as far as saying that his theory was lacking a "cause" (see the quote above). Now, although when first faced with this new sort of theory scientists were highly dissatisfied with this impoverished IC theory, in want of a complete IC theory, they came to accept force as an explanatory primitive. In other words, Newton's theory is a bit of an odd ball. McMullin puts this thus:

In mechanics itself, however, Newton established the credentials of a new and unfamiliar way of proceeding. The language of *force* and *attraction* carried with it the suggestion of agency, of the Sun drawing the planets, of an agent-causal explanation being given, even though this was expressly excluded by Newton himself […] The form of explanation given is weaker than agent-causal, where the agency itself is identified and the mechanisms of its action described. It might be called "dynamic" since it is dependent on the peculiarity of the notion of force as Newton elaborates it. (McMullin 2001, 297p.)

7. Conclusions

In this paper I proposed an IC account of explanation, which is to be understood as an alternative to the causal-mechanistic account of explanation (CM). Whereas the CM account is committed to a realism about causes "producing" the phenomena and to some form of reductionism of the explanandum to the explanans, my account of *Imaginary-Constitutives* (ICs) is not. The IC account acknowledges that the relation between the explanans and the explanandum is to some degree arbitrary. However, in contrast to pragmatist proposals of explanation, which have been suggested as alternatives to realist accounts of explanation (most notably by van Fraassen), the IC account of explanation deems scientific explanations to be objective by having to fulfil the requirement of being able to simulate real world regularities in a qualitative way[27]. The IC account thus does not compromise on the intuition that the CM account somehow provides a plausible account of *scientific* explanation but manages without the metaphysical baggage of the CM account. Furthermore, through the

[27] This distinguishes IC simulations from computer simulations which usually just aim to be quantitatively accurate.

concept of ICs, not only visualizability, but also the constructive element of explanations (Einstein), is taken beyond a mere intuitive justification and is put on a rational footing.

8. References

Aerts, D. and F. Rohrlich (1998), "Reduction", *Foundations of Science*, 1, 27-35.

Balashov, Y. and M. Janssen (2003), "Presentism and Relativity", *British Journal for the Philosophy of Science*, 54: 327-346.

Brown, H. (2005), Einstein's misgivings about his 1905 formulation of special relativity, *European Journal of Physics*, 26, S85-S90

Brown, H. and Pooley, O. (2006), "Minkowski Space-Time: a Glorious Non-Entity", in: Petkov (ed.) *The Ontology of Spacetime* (in preparation).

Cushing, J. (1991), "Quantum theory and explanatory discourse: endgame for understanding?", *Philosophy of Science*, 58, pp.337-358

Chang, H. (2003), "Preservative Realism and Its Discontents: Revisiting Caloric", *Philosophy of Science*, volume 70, p. 902–912.

Craver, C. (2005), 'Beyond reduction: Mechanisms, multifield integration and the unity of neuroscience', *Studies in History and Philosophy of Biological and Biomedical Sciences*, v. 31, No. 2, June 2005, Special Issue: Mechanisms in Biology.

de Regt, H. (1999), "Ludwig Botzmann's Bildtheorie and Scientific Understanding", *Synthese*, 119, pp. 113-134.

de Regt, H. (1997), "Erwin Schrödinger, Anschaulichkeit, and Quantum Theory", *Studies in History and Philosophy of Modern Physics*, no. 4, pp. 461-481.

de Regt, H. (2001), "Spacetime Visualisation and the Intelligibility of Physical Theories", ", *Studies in History and Philosophy of Modern Physics,*, Vol. 32, No. 2, pp. 243–265.

de Regt, H. and Dieks (2005), "A contextual approach to scientific understanding", *Synthese*, 144, 137-170.

Einstein, A. (1935), *The world as I see it*, trans. by Alan Harris, London: J. Lane.

Einstein, A. (1995), 'Letter to Arnold Sommerfeld, January 14, 1908. Document 73', *The Collected Papers of Albert Einstein, Vol 5, The Swiss Years: Correspondence, 1902-1914 (English Translation Supplement)* ed. M J Klein, A J Kox and R Schulmann (Princeton, NJ: Princeton University Press) (translated by A Beck)

Elsamahi, M. (2004), "A Critique of Localized Realism", in *Proceedings Philosophy of Science Assoc. 19th Biennial Meeting - PSA2004: PSA 2004 Contributed Papers*.

Friedman, M. (1974), "Explanation and Scientific Understanding", *The Journal of* Philosophy, vol. 71, no.1 (Jan. 17), 5-19

Fox, R. (1971), The *caloric theory of gases: from Lavoisier to Regnault*, Oxford: Clarendon Press.

Hempel, C. and P. Oppenheim, (1948), "Studies in the Logic of Explanation", *Philosophy of Science*, volume 15, page 135.

Hesse, M. (1961), *Forces and Fields: The Concept of Action at a Distance in the history of physics*, London: Thomas Nelson and Sons Ltd..

Hitchcock, C., (1995), 'Discussion: Salmon on Explanatory Relevance', *Philosophy of Science* 62: 304-20.

Kitcher, P. (1981), "Explanatory Unification", *Philosophy of Science*, volume 48, page 507.

Kitcher, P., (1989), 'Explanatory Unification and the Causal Structure of the World', in *Scientific Explanation*, P. Kitcher and W. Salmon, 410-505. Minneapolis: University of Minnesota Press.

Lilley, S. (1948), "Attitudes to the Nature of Heat about the Beginning of the Nineteenth Century", *Archives Internationales d'Histoire des Sciences*, 1, 630-639.

Machamer, P., L. Darden, C. Craver (2000), 'Thinking about Mechanisms', *Philosophy of Science*, Vol. 67, No. 1 (March), pp. 1-25.

McAllister, J. (1999), *Beauty & revolution in science*, Ithaca: Cornell University Press.
McMullin, E. (1978), “Structural Explanation”, *American Philosophical Quarterly*, vol. 15, no. 2, April.
McMullin, E. (1984), “Two Ideals of Explanation in Natural Science”, *Midwest Studies in Philosophy,* IX, 205-220.
Nagel, E. (1961), *The structure of science: problems in the logic of scientific explanation*, London: Routledge & K. Paul.
Newton-Smith, W.H. (2000), ‘Explanation’, in: W.H. Newton-Smith (ed), A Companion to the Philosophy of Science. Oxford: Blackwell, pp.127-133.
Salmon, W. (1984), *Scientific explanation and the causal structure of the world*, Princeton, N.J.: Princeton University Press.
Scerri, E. and L. McIntyre (1997), ‘The Case for the Philosophy of Chemistry’, *Synthese,* 111, 213–232.
Schindler, S. (2006), “Imaginary-Constitutives: The Ontology of Scientific Models”, presented at the Biannual Meeting of the *Philosophy of Science Association (PSA)*, 2-5 November, Vancouver. Downloadable at http://philsci-archive.pitt.edu/archive/00002970/
Schindler, S. (2007), "Rehabilitating Theory. The Refusal of the bottom-up construction of Scientific Phenomena", to appear in *Studies in the History and Philosophy of Science*.
Schindler, S. (preprint), “Model, Theory and Evidence in the Discovery of the DNA Structure".
Strevens, M. (forthcoming), “Why explanations Lie: Idealization in Explanation”, to be published in Strevens, M. (forthcoming), *The Kairetic Account of Explanation*.
Trout, J. D. (2002), ‘‘Scientific Explanation and the Sense of Understanding’’, *Philosophy of Science* 69: 212–233.
Van Fraassen, B. (1985), “Salmon on Explanation”, *The Journal of Philosophy*, Vol. 82, No. 11, Eighty-Second Annual Meeting American Philosophical Association, Eastern Division, pp. 639-651.
Woodward, J. (1979), “Scientific Explanation”, *British Journal for the Philosophy of Science,* 30, 41-67.