

Original source of publication: *Studies in History and Philosophy of Science* (2011) 42: 602-609
The final publication is available at www.sciencedirect.com
<http://www.sciencedirect.com/science/article/pii/S0039368111000732>
DOI 10.1016/j.shpsa.2011.09.009

A Pragmatic Account of Mechanistic Artifact Explanation

Jan De Winter

Centre for Logic and Philosophy of Science, Department of Philosophy and moral sciences, Ghent University, Blandijnberg 2, Room 226, 9000 Ghent, Belgium

e-mail: Jan.DeWinter@UGent.be

telephone: ++32 9 264 39 79

fax: ++32 9 264 41 87

Abstract

A mechanistic artifact explanation is an explanation that accounts for an artifact behavior by describing the underlying mechanism. The article shows that there are different kinds of mechanistic artifact explanation: top-down and bottom-up explanation, and I also distinguish between less and more inclusive top-down explanations. To illustrate these different kinds of explanation, the behavior of a simple, fictional artifact is explained in different ways. I defend that which explanation is ideal, depends on pragmatic factors (e.g., the background knowledge of the explainee and the specific goal for which the explanation will be used). For each kind of explanation, the situations, goals and interests for which it is most appropriate are specified, resulting in a pragmatic theory of mechanistic artifact explanation. This theory is compared to Jeroen de Ridder's account of the pragmatics of mechanistic artifact explanation.

Keywords: artifact explanation, mechanistic explanation, technology, top-down, bottom-up, explanatory pluralism

1. Introduction

According to Kroes (1998), a technological design is more than a complete description of a physical system: it also contains what he calls a technological explanation, i.e. an explanation of the function of an artifact in terms of its physical structure. A technological explanation accounts for the function of an artifact by referring to the structure (design) of the artifact, the relevant physical phenomena, and the actions necessary for the artifact to perform its function. Because the function of an artifact cannot be deduced from the explanans of a technological explanation, Kroes concludes that a technological explanation does not connect structure and function on the basis of a logical deduction, but “on the basis of causal relations and pragmatic rules of actions based on these causal relations” (p. 34).

De Ridder (2006a) considers this notion of technological explanation problematic because he does not agree that causal relations and pragmatic rules of action can actually bridge the gap between structure and proper function. He thinks the problem with Kroes' account is that he runs

together two projects that should be distinguished: the project of developing an account of proper function ascriptions, and the project of explaining how the physicochemical make-up of an artifact enables it to perform its function. With respect to the latter project, de Ridder refers to his article on mechanistic artifact explanation (de Ridder, 2006b). In a mechanistic artifact explanation, one does not explain the function of an artifact, but an artifact behavior. De Ridder (2006b) discusses two strategies, top-down and bottom-up, to produce mechanistic artifact explanations, and explicates the kinds of contexts in which each strategy is appropriate. These contexts are, however, quite general, and need to be specified. We should also identify the goals for which top-down and bottom-up explanations are useful. In his Ph.D. dissertation (2007), de Ridder offers an account of the different kinds of information that the different strategies produce, and uses it to argue for the autonomy of top-down explanation. This account is, however, not entirely unproblematic, as we will see below. So a lot of work remains to be done.

This article offers a theory of mechanistic artifact explanation that deals with these issues, and that is inspired by and supplements Weber and Van Bouwel's work on explanatory pluralism (e.g., Van Bouwel & Weber 2002, 2008; Weber & Van Bouwel 2007). A mechanistic artifact explanation will here be defined as an explanation that accounts for an artifact behavior by describing the underlying mechanism. Following Glennan (2002, p. S344), a mechanism can be defined as follows:

A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations.¹

In this article, I show that different mechanistic artifact explanations can be constructed for one and the same artifact behavior, and which explanation is most appropriate, depends on pragmatic factors such as the goals and background knowledge of the explainee (= the person who asks for an explanation).² The relevant contexts, goals and interests are specified.

First, let me give an example of a simple mechanism that I will use to illustrate and clarify my claims throughout the article. The mechanism is depicted in Figure 1. When one puts a wooden box on the left block of the artifact (**1**), then it slides down the slope until it is on the foot of the elevator (**2**). When one pulls the handle, the box moves up (from **2** to **3**), after which it slides down the second slope, to finally drop in the cup on the right hand side of the box-moving device (**4**). The explainee does not know what happens within the dotted line. The explainee knows that when one puts a box on the left block, and pulls the handle after a while, the box will turn up on the right hand side of the artifact, but he does not know the inner workings of the box-moving device. However, he is interested in these inner workings, and therefore, he asks for an explanation that explains the behavior of the artifact by describing the underlying mechanism of this behavior. In other words, he asks for a mechanistic artifact explanation of the transportation of the box.

Section 2 of this article shows that there are at least three different explanations that serve the interest of the explainee. In section 3, it is argued that which explanation is most appropriate differs from context to context, and depends on pragmatic factors such as the background knowledge of the explainee and the specific goal for which the explanation will be used. Based on my findings, I develop a general pragmatic account of mechanistic artifact explanation. This account is compared to de Ridder's claims concerning the pragmatics of mechanistic artifact explanation in section 4, and I conclude in section 5.

2. Top-down and bottom-up

De Ridder (2006b) distinguishes two strategies to create mechanistic artifact explanations: the top-down strategy and the bottom-up strategy. He offers the following definitions:

Top-down strategy: take the behavior to be explained and decompose it into more basic sub-behaviors, reiterate this step if possible, it should become clear how the complex behavior being explained is realized by simpler behaviors in a specific spatiotemporal configuration, and for all the sub-behaviors, indicate which component(s) take(s) care of them.

Bottom-up strategy: name the structural components of the artifact and give information about their physicochemical make-up and spatial configuration, show how their physicochemical features and configuration result in various behaviors and then describe how these behaviors, in their spatiotemporal configuration, together make up the behavior to be explained. (de Ridder, 2006b, p. 87)³

Before I illustrate these strategies, it should be noted that whether or not the first step of the top-down strategy (decompose the behavior into more basic sub-behaviors) can, and thus should, be reiterated, is relative to context in de Ridder's view. We can infer this from the following quote.

It would be absurd if the [...] guideline ['break the function down as fine as possible'] were interpreted as requiring that every functional decomposition end in fundamental physics. The relevant sense of 'as fine as possible' must be more pragmatic. It seems plausible that certain sub-behaviors will count as basic or atomic for particular engineering disciplines. Which will vary across different fields; where mechanical engineering will accept, say, the strength of a particular alloy as given, materials engineering will be interested in how this strength is realized by the behaviors of elements in the alloy. (de Ridder, 2007, p. 83)

This implies that which sub-behaviors are not further decomposed in a top-down explanation can differ from context to context, and thus, that a top-down explanation can include less or more information.

Let me construct two different top-down explanations of the transportation of the box in the box-mover. The first makes abstraction of the forces at work, the second does not. The first top-down explanation decomposes the transportation of the box into three sub-behaviors: the box's movement from **1** to **2** (see Figure 1) during time interval $[t_1, t_2[$, the upward movement of the elevator with the transportation of the box to **3** during $[t_2, t_3[$, and the box's movement from **3** to **4** during $[t_3, t_4[$. The first sub-behavior is taken care of by slope **A** (see Figure 2); the second sub-behavior is taken care of by elevator system **B**; and the third sub-behavior is taken care of by slope **C**. This is the first top-down explanation.

The second top-down explanation includes more information. Just as the first one, it decomposes the transportation of the box into three sub-behaviors: the box's movement from **1** to **2** during time interval $[t_1, t_2[$, the upward movement of the elevator with the transportation of the box to **3** during $[t_2, t_3[$, and the box's movement from **3** to **4** during $[t_3, t_4[$. But it also decomposes each of these behaviors. Slope **A** causes the box to move from **1** to **2** by reacting to gravitational force F_G (see Figure 3) with normal force F_N , and by reacting to the resultant $F_R (= F_G \cdot \sin \alpha)$ of these two forces with a frictional force F_F that is smaller than F_R , so that the total force F , which is parallel to the surface of the ramp and equal to $F_R - F_F$, is larger than 0. Because $F > 0$, the box slides down slope **A**, from **1** to **2**. The box's movement from **2** to **3**, and from **3** to **4**, can be

decomposed in similar ways. To reduce the complexity of this article, I do, however, not explicate such decompositions here, but they are assumed to be included in the second top-down explanation.

Some may claim that only an explanation such as the second, more extensive top-down explanation is a full explanation of the transportation of the box from **1** to **4**, and that an explanation that does not reveal the forces at work is not a real explanation. I return to this point in section 3, showing that the first, more concise top-down explanation can be a complete and fully satisfying explanation as well.

Now consider a bottom-up explanation of the transportation of the box. Such an explanation might look as follows. The box is a 0,5 kg oak box. Components **D**, **E** and **G** (see Figure 4) are oak blocks that each have a lubricated 30-degree slope. Components **F** and **H** are also made of oak. The foot of **F** has a 30-degree slope that is lubricated, and is strongly screwed down the rod, which is held upright by the surrounding box. The top of **H** is a lubricated 30-degree slope with a barricade at its end. (Of course, several other features should be added, such as the length, width, and height of each component, but I will stop here because summing up all these features is not necessary to clarify how a bottom-up explanation might look.) In which behaviors do the physicochemical and spatial features of the components of the box-mover result?

This is revealed in the second part of the bottom-up explanation. Because the ramps in the box-mover are constituted by oak components, they do not collapse when a 0,5 kg box is on top of them (if the constituting components were made of, say, whipped cream, then the ramps would collapse). Instead, they react to the weight F_G of the box with normal force F_N (see Figure 3). F_R is the resultant of F_G and F_N , and is equal to $F_G \cdot \sin \alpha$. Since $F_G = 4,9$ N and $\alpha = 30^\circ$, F_R is equal to 2,45 N. The box slides down the slope if F_R exceeds the friction between the slope and the box (F_F). In the box-mover, it is guaranteed that F_F is smaller than 2,45 N because the slopes and the box are lubricated (lubrication reduces friction). As a consequence, the box moves down the slopes in the box-mover, from **1** to **2**, and from **3** to **4**. Because **F** is made of oak and because the foot is strongly screwed down the rod, it is solid enough to carry the weight of the box if the handle is pulled. As **G** is made of oak, it is solid enough to prevent the box from sliding down the foot of **F** as long as it is next to **G**. Because **F** can carry the weight of the box, and because **G** prevents the box from sliding down **F**'s foot, we know that when the box is on the foot of **F** and the handle is pulled, the box moves from **2** to **3**.

So far, I have explicated physicochemical and spatial features of the different components of the box-moving device, and I have shown how these features result in various behaviors. One thing remains to be done for our bottom-up explanation to be complete: it should be clarified how these various behaviors make up the behavior to be explained, that is, the transportation of the box. This can easily be done by specifying the temporal organization of the different behaviors. The box's movement from **1** to **2** occurs during time interval $[t_1, t_2[$, its movement from **2** to **3** occurs during $[t_2, t_3[$, and its movement from **3** to **4** occurs during $[t_3, t_4[$. This organization constitutes the transportation of the box from **1** to **4** during $[t_1, t_4[$.

What are the differences between top-down and bottom-up explanation of the transportation of the box? A salient difference is the fact that the bottom-up explanation does not only refer to sub-behaviors of the behavior to be explained and the components taking care of these behaviors, as the two top-down explanations do; it also mentions physicochemical and spatial properties of the artifact components, while the top-down explanations do not.

A second difference is that the artifact components to which the top-down explanation refers, differ from the artifact components to which the bottom-up explanation refers (compare Figure 2 to Figure 4; the components are marked out differently). This is due to a difference in

perspective. When one uses the top-down strategy, one takes a behavioral or functional perspective: one marks out the artifact components on the basis of the sub-behaviors of the behavior to be explained. If several pieces contribute to the same behavior, then there is, from a functional perspective, a tendency to take them together as one more-piece component. When one uses the bottom-up strategy, one takes a structural perspective, which means that one identifies the artifact components before one knows what they do. From this perspective, one will typically pick out one-piece components; these components cannot be taken together on the basis of the behavior to which they contribute because one does not know the relevant behaviors yet. Let me call artifact components that are identified from a functional perspective, functional components or F-components, and artifact components that are identified from a structural perspective, structural components or S-components. Since top-down explanations correspond with a functional perspective, they refer to F-components, and since bottom-up explanations correspond with a structural perspective, they refer to S-components.⁴

3. Pragmatics

So there are at least three mechanistic artifact explanations of the transportation of the box in the box-mover. Does the fact that the bottom-up explanation reveals more about the artifact than the top-down explanations imply that the bottom-up explanation is always the best explanation? I do not think so. I think that for each explanation, there are contexts in which it is ideal, and that which explanation is ideal depends on pragmatic factors such as the background knowledge of the explainee and the specific goal for which the explanation will be used. Mechanistic artifact explanations are useful in at least two kinds of contexts: the context of use, and the context of design. First, let us consider the context of use.

3.1. *Context of use*

If we know the underlying mechanism of the box transportation, then we can draw several conclusions about how the box-mover will react to certain manipulations. For instance, we know that if we hold the box-mover upside down, it will not display the behavior it is supposed to display; and if we pull the handle before the box is on the foot of the elevator, then the box will end up under the elevator. We can draw these conclusions from all the mechanistic explanations of the box transportation discussed. But some conclusions can only be drawn from the bottom-up explanation. For instance, we know that we can drive (metal) screws in the parts of the box-mover (e.g., for attaching the box-mover to a wooden bottom plate) because we know that these parts are made of oak, and not of metal or some other hard material. The fact that the parts are made of oak, and not of metal or some other hard material, is only revealed by the bottom-up explanation.

We can draw more conclusions about an artifact from a bottom-up explanation of the behavior of this artifact, but this comes with a cost. Suppose we want to draw conclusions that are not only relevant to the artifact under consideration, but also to artifacts with other physicochemical/spatial properties. We should then make abstraction of the physicochemical and spatial properties, and only focus on the factors that the different artifacts have in common. Top-down explanations do that, contrary to bottom-up explanations, and therefore, the conclusions we draw from them are more likely to be relevant for artifacts with other physicochemical/spatial properties as well. So while bottom-up explanations are more useful if our goal is to draw as

much conclusions as possible about a certain specific artifact, top-down explanations are more useful for drawing conclusions that are relevant for as much artifacts as possible.

But how much information should exactly be included in a top-down explanation? This depends on the background knowledge of the explainee. Consider the first top-down explanation of the box transportation. Once we know this explanation, we can predict what will happen if we turn the box-mover upside down, or if we pull the handle before the box is on the foot of the elevator; we do not need the additional information that the second top-down explanation provides to make such predictions. This is because we are familiar with falling objects and with objects sliding down slopes.

But sometimes, we want to make predictions concerning situations with which we are not familiar. Suppose that, before the first man traveled in space, one wanted to know how the box-mover would behave in an orbital space station. The explainee could then not be familiar with the situation about which he wanted to make a prediction. As a result, the first top-down explanation could not help him out: knowing the sub-behaviors of the box transportation on earth, and the components taking care of these behaviors, would not be sufficient to predict what would happen in space. The explainee would need more information, and such information is provided by the second top-down explanation. The second top-down explanation reveals that slope **A** causes the box to move from **1** to **2** by reacting to F_G with normal force F_N . Since slope **A** does not react to F_G with normal force F_N in an orbital space station, our explainee could conclude, on the basis of the second top-down explanation, that the box would not move from **1** to **2** in the space station.

After several decades, most people have seen videos, pictures, etc. of objects in orbital space stations, and they are familiar with the dispositions of such objects (they float). Given this background knowledge, they can predict that the box-mover will not transport boxes in an orbital space station as it does on earth, even if they only have the top-down explanation of the box transportation; they do not need to know the forces that cause the sub-behaviors of the box transportation to have a sense about how the box-mover will behave. This demonstrates that how much information is needed to make predictions about how the box-mover will behave in certain situations, depends on the explainee's familiarity with such situations. If he is familiar with such situations, the first top-down explanation is most appropriate, as it provides all the information needed, and nothing more. But if the explainee is not familiar with such situations, he needs additional information, and that information is provided by the second top-down explanation. So in that case, the second top-down explanation is most appropriate.

3.2. Context of design

Mechanistic artifact explanations are also useful in the context of design. Suppose we want to design an artifact that performs the same behavior as some pre-existing artifact. A mechanistic artifact explanation of the behavior of the pre-existing artifact can then be very helpful. Suppose we want to build a fireproof box-mover that is as flat as possible. Due to the new requirements, we do not consider the physicochemical and spatial features of the original box-mover's components to be very relevant for our goal (we cannot just use oak because the box-mover should be fireproof, and the slopes should be reconsidered because the box-mover should be as flat as possible), and therefore, the explanation should not explicate these physicochemical and spatial features. So the bottom-up strategy drops out. What is useful in this context, is to know the sub-behaviors of the box transportation and the components taking care of these sub-behaviors. Such information is very useful for our goal because these sub-behaviors and components can be used to realize the box transportation in the new box-mover as well. Although

the components of the new box-mover will have different physicochemical and spatial properties, it can, just as the original box-mover, transport the box by making it slide down two slopes and by using an elevator to transport the box from the first slope to the second. As both top-down explanations refer to these sub-behaviors and components, the question becomes: which top-down explanation is most appropriate?

This depends, again, on contextual factors. To know the minimal height of the box-mover (it should be as flat as possible), one should determine the optimal degree of slope (the lower the degree of slope, the flatter one can make the box-mover, but if the degree of slope is too low, the box does not slide down anymore). To determine this optimum, the generalizations to which the second top-down explanation refers ($F_R = F_G \cdot \sin \alpha$, and $F = F_R - F_F$) are very useful. If the material of the box and of the ramp is given, then we can, on the basis of these generalizations, determine a minimal degree of slope that guarantees the sliding down of the box. So it seems that the second top-down explanation of the box transportation is most useful, and thus most appropriate, as it refers, contrary to the first top-down explanation, to generalizations that can help us in reaching our goal. This is often the case, but there are also contexts in which the first top-down explanation is most appropriate.

Suppose the explainee has a lot of experience with slopes, and knows, for different combinations of materials (e.g., a copper box on a zinc ramp, an aluminum box on a glass ramp, etc.), the lowest degree of slope that guarantees the sliding down of the box. Due to this background knowledge, the explainee does not need the generalizations to which the second top-down explanation refers to reach his goal, that is, to make a fireproof box-mover that is as flat as possible. The additional information that the second top-down explanation provides would then be redundant, and the first top-down explanation would provide all the information needed, and thus be ideal in this context.

It is also possible that the job of the explainee is not to build the new box-mover by himself, but to delegate sub-tasks to experts. Once he knows that a box can be transported by making it slide down two slopes, and by using an elevator to transport it from the first slope to the second, he can give slope experts the assignment of designing the two slopes, and experts in elevator systems the assignment of designing the elevator system. No additional information is needed, which means that the first top-down explanation is fully satisfying.

So top-down explanations are often most appropriate for developing an artifact that performs the same behavior as a pre-existing artifact, and how much information should exactly be included in the top-down explanation can differ from context to context. In still other contexts, the most promising explanatory strategy to create an artifact that displays the same behavior as a pre-existing artifact, is the bottom-up strategy. This is the case when a description of the physicochemical make-up and spatial configuration of the components of the pre-existing artifact is a good starting point for the new design. If the current context resembles the context in which the pre-existing artifact was designed (same requirements, same material costs, etc.), then a lot of physicochemical and spatial features of the components of the original artifact can be used for the new artifact as well. In such cases, knowing these features is useful. Because a bottom-up explanation explicates them, contrary to top-down explanations, it is more appropriate in such cases.

The first part of a bottom-up explanation (description of the physicochemical make-up and spatial configuration of the S-components of the artifact) is a first suggestion on how an artifact performing the desired behavior might look. We may, however, still be critical of the physicochemical make-up and spatial configuration of the components of the original artifact, and ask questions such as ‘Why is the artifact made of the materials it is made of?’ The second

and third part of a bottom-up explanation (demonstration of how the physicochemical make-up and spatial configuration of the artifact components result in various behaviors and a description of how these behaviors result in the behavior to be explained) help one to answer such questions. For instance, the bottom-up explanation of the box transportation reveals that the artifact components are made of oak *in order to be able to carry the weight of the box without collapsing*. Such information is not only useful to legitimize the physicochemical or spatial property under consideration, but it is also relevant if one decides to change it: e.g., given that the components of the original box-mover are made of oak in order to be able to carry the weight of the box, we know that the new material should be solid enough to carry the weight of the box as well.

3.3. *General claims*

Based on my considerations on (the appropriateness of) the different explanations of the transportation of the box in the box-mover, I can develop a more general account of mechanistic artifact explanation. We have seen that there are at least two strategies to produce mechanistic artifact explanations: top-down and bottom-up. Which strategy is most appropriate depends on pragmatic factors. More specifically, it depends on whether or not information about the physicochemical make-up and spatial configuration of the artifact components helps the explainee reach his goal. When such information is useful, for instance, because the explainee wants to draw as much conclusions as possible about the artifact, or because he wants to develop a new artifact and the physicochemical and spatial properties of the components of the existing artifact form an interesting starting point for this, bottom-up explanation is most appropriate. If information about the physicochemical and spatial features is not useful, top-down explanation is most appropriate. This is the case when the explainee wants to draw conclusions that are not only relevant for the artifact under consideration, but also for other artifacts, with other physicochemical and spatial features, or when only the abstract design, but not the physicochemical and spatial features of the artifact components, is relevant for the development of a new artifact.

Furthermore, we have seen that top-down explanations can include more or less information. How much information should exactly be included in a top-down explanation depends, again, on pragmatic factors, such as the background knowledge of the explainee, and the task for which he will use the explanation. The more background knowledge the explainee has, the less explanatory information he may need to be able to achieve his goal, and if his job is to delegate sub-tasks to experts, he may need less information than when his job is to develop a new artifact all by himself.

3.4. *Limitations of mechanistic artifact explanation*

We have seen that there are various goals for which mechanistic artifact explanations are very useful. There are, however, also contexts in which non-mechanistic explanations are preferential. Suppose that the question ‘Why is the box transported from the one end of the box-mover to the other end?’ is not motivated by the interest in the underlying mechanism of this behavior, but by the desire to know why the box-mover did not work properly a minute ago, while it does work properly at this moment. The explanatory interest in the box transportation can then be motivated by the desire to know what has changed. An explanation that perfectly serves this interest is: the box is transported because someone repaired the handle of the box-mover, which was broken a minute ago. This is not a mechanistic artifact explanation as it does not explicate any mechanism,

and it outperforms the aforementioned mechanistic artifact explanations because it reveals what has changed during the time interval under consideration, while the mechanistic artifact explanations do not.

So it is wrong to think that offering a good explanation of an artifact behavior means offering a mechanistic explanation of this behavior. Although mechanistic artifact explanations can serve various goals, there are also goals that they do not serve, such as the desire to know what changed during a certain time interval (also see De Winter, 2010). For these goals, mechanistic artifact explanations are bad explanations.

4. Comparison with de Ridder's account

Now, let me compare my pragmatic account of mechanistic artifact explanation to Jeroen de Ridder's account of the pragmatics of mechanistic artifact explanation. The latter account can be divided in two parts: (1) some remarks de Ridder (2006b) makes on the kinds of contexts in which bottom-up explanations and top-down explanations are useful, and (2) a pragmatic argument for the autonomy of top-down explanations developed in de Ridder (2007). Let us start with the first part.

According to de Ridder (2006b), bottom-up explanations provide structural understanding, which “shows one exactly how a particular mechanism *implements* a piece of behavior” (p. 95). Such understanding is highly useful in engineering contexts because “it differentiates between the detailed implementation choices to be made in figuring out the details of a new design” (p. 95). Top-down explanation, on the other hand, provide behavioral understanding, which “exhibits how a particular complex behavior can be created out of simpler behaviors, independent of the particular realization of these behaviors” (p. 95). De Ridder considers such understanding to be useful in early stages of engineering design, when the engineer is reflecting on how a complex artifact behavior can be constituted by simpler sub-behaviors, without worrying (yet) about how to realize these sub-behaviors. Top-down explanations are, according to de Ridder, also very useful in a lot of everyday contexts, because “[i]n most everyday contexts [...], people are not interested in the gory details, but only need a rough behavioral understanding of how a complex behavior is produced” (p. 89).

There are at least four respects in which my account is more sophisticated. Firstly, it identifies some specific goals for which mechanistic artifact explanations are useful: the goal of drawing conclusions about the use of the artifact, the goal of designing an artifact that performs the same behavior as the artifact of which the behavior is explained, and the goal of delegating sub-tasks of designing such an artifact to the right experts. Secondly, it specifies the kinds of engineering contexts in which bottom-up explanation is most appropriate: design contexts in which the physicochemical and spatial features of the components of the existing artifact form an interesting starting point for the new design. If these features are irrelevant for the new design (because the components of the new artifact should have different physicochemical and spatial features due to new design requirements), top-down explanation is most appropriate. Thirdly, I have shown that bottom-up explanations are not only useful in engineering contexts, but that they can also be useful in contexts of use; bottom-up explanations are most useful if the goal is to draw as much conclusions as possible about the use of the artifact. Fourthly, I have identified two pragmatic factors on which the amount of information that should be included in a top-down explanation depends: the background knowledge of the explainee and his specific task.

Now consider the second part of de Ridder's account of the pragmatics of mechanistic artifact explanation: his pragmatic argument for the autonomy of top-down explanations. According to the pragmatic argument, top-down explanations are more appropriate when one is interested in comparative information. Comparative information shows how the realization of the behavior to be explained *is similar to* other possible realizations of the same behavior. Because the factors to which top-down explanations refer, are not only characteristic of the realization of the behavior to be explained in the artifact under consideration, but also of the realization of this behavior in artifacts with other physicochemical/spatial features (as long as the artifact components perform the same sub-behaviors), top-down explanations are more appropriate if one is interested in comparative information. Because bottom-up explanations do not make abstraction of the physicochemical make-up and spatial configuration of the artifact components, they are, according to de Ridder, less appropriate for the interest in comparative information. But they are more appropriate if one is interested in contrastive information, i.e. information that shows how the realization of the behavior to be explained *differs from* other possible states of affairs. By explicating physicochemical and spatial features, bottom-up explanations focus on factors that are unique for the realization of the behavior to be explained in the artifact under consideration. These factors distinguish the realization of the behavior to be explained in the artifact under consideration from other possible ways to realize this behavior.

De Ridder also considers top-down explanations more appropriate than bottom-up explanations for the interest in counterfactual information, that is, the interest in factors that make a difference with respect to the occurrence of the behavior to be explained.⁵ Changing one of the physicochemical or spatial properties will often not make a difference with respect to the occurrence of the behavior to be explained (suppose the box-mover is made of plastic instead of being made of oak), while changing one of the sub-behaviors of this behavior will make a difference. Since bottom-up explanations describe the artifact components' physicochemical and spatial features, while top-down explanations only focus on sub-behaviors, de Ridder (2007, p. 187) concludes that the top-down strategy "does a better job at suggesting the right kinds of counterfactuals".

So de Ridder claims (1) that top-down explanation is more appropriate if the explainee is interested in comparative information, while bottom-up explanation is more appropriate if the explainee is interested in contrastive information, and (2) that top-down explanation is more appropriate if the explainee is interested in counterfactual information. I agree with both claims, but I reject de Ridder's argument in favor of (2); I have a different reason for accepting (2).

De Ridder's (2007, p. 185) argument in favor of (2) is that while the top-down strategy "suggests the right kinds of counterfactuals", the bottom-up strategy "incorrectly suggests that the spatial and physicochemical properties in their categorical descriptions make a crucial difference to the artifact's overall behavior." This is not correct. Consider the bottom-up explanation of the transportation of the box. This explanation includes a lot of counterfactual information. Just as the two top-down explanations of the box transportation, it refers to sub-behaviors that are crucial for the overall behavior. It also reveals factors that make a difference with respect to these sub-behaviors. For instance, the generalization "the box slides down a slope in the box-mover if $F_R > F_F$ (see Figure 3)" reveals a factor that makes a difference with respect to the sliding down of the box: " $F_R > F_F$ ". This factor is difference-making because we can prevent the box from sliding down the slope by eliminating this factor (by manipulating F_R or F_F so that $F_R = F_F$). Another generalization included in the bottom-up explanation is " $F_R = F_G \cdot \sin \alpha$ ". This generalization points to factors that make a difference with respect to F_R : F_G and α . F_G and α are difference-making because we can change the value of F_R by changing the value of F_G

or of α . I could go on, but the point is that the bottom-up explanation picks out the factors that make a difference. More generally, we can say that, just as top-down explanations, bottom-up explanations offer a lot of counterfactual information. Of course, bottom-up explanations also refer to physicochemical and spatial features, which usually do not make a crucial difference, but, contrary to what de Ridder claims, bottom-up explanations do not suggest that they do. For instance, the bottom-up explanation of the box transportation does not suggest that the fact that the components of the box-mover are made of oak is *crucial* for the overall behavior of the box-mover, but only that it is *relevant* for this behavior.

Why, then, is top-down explanation more appropriate for the interest in counterfactual information than bottom-up explanation? To answer this question, we can use the three criteria suggested by Van Bouwel and Weber (2008, p. 175) to compare alternative explanations: their accuracy (correctness, that is, their relation with reality), their adequacy (their relation to what the explainee expects from the explanation, to the purpose of the explanation) and their efficiency (the amount of work that is needed to construct the explanation). Suppose that someone seeks a mechanistic artifact explanation of an artifact behavior and that he is interested in counterfactual information. Consider two explanations, a top-down and a bottom-up explanation, which are equally correct, and which both provide all the counterfactual information needed. The two explanations are then equally accurate and equally adequate. But the top-down explanation outperforms the bottom-up explanation with respect to efficiency, that is, its construction requires less work. This is because not only counterfactual information should be included in the bottom-up explanation, but also a lot of information about the artifact components' physicochemical make-up and their spatial configuration. Such information should not be included in the top-down explanation, so that less effort is required to construct it. Because the top-down explanation is more efficient than the bottom-up explanation, while being equally accurate and adequate, it is more appropriate in the context under consideration, that is, a context in which the explainee is interested in counterfactual information.

5. Conclusion

The general outline of my pragmatic account of mechanistic artifact explanation is this. If we want to explain an artifact behavior by the underlying mechanism of this behavior, that is, if we want to construct a mechanistic artifact explanation, then this can often be done in more than one way. Which mechanistic artifact explanation is most appropriate depends on pragmatic factors, such as the goal of the explainee and his background knowledge. The reason is that these factors determine which information is useful for the explainee, and thus, which information should be included in the explanation and which not.

Following Jeroen de Ridder, I have distinguished between bottom-up and top-down explanation, and between less and more inclusive top-down explanations. We should not think of one of these kinds of explanation as the real/best characterization of mechanistic artifact explanation. Each kind is, in my opinion, a genuine characterization of a class of mechanistic artifact explanations, and which kind is ideal can differ from context to context. For each kind of explanation, I have specified contexts, goals and interests for which it is more and less appropriate, resulting in a pragmatic account of mechanistic artifact explanation that is more sophisticated than de Ridder's account of the pragmatics of mechanistic artifact explanation.

Acknowledgements

Research for this paper was supported by subventions from the Research Foundation – Flanders through research project 3G003109. I am very grateful to Erik Weber, Jeroen Van Bouwel, Bert Leuridan and two anonymous reviewers for reviewing earlier versions of this paper, as well as to Gust De Winter, for offering several examples of mechanisms that I used to test my claims.

References

- Bechtel, W., & Richardson, R. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton: Princeton University Press.
- Cummins, R. (1975). Functional analysis. *Journal of Philosophy*, 72, 741-765.
- De Ridder, J. (2006a). The (alleged) inherent normativity of technological explanations. *Techné*, 10, 97-115.
- De Ridder, J. (2006b). Mechanistic artefact explanation. *Studies in History and Philosophy of Science*, 37, 81-96.
- De Ridder, J. (2007). *Reconstructing design, explaining artifacts*. Ph.D. Dissertation, Delft University of Technology.
- De Winter, J. (2010). Explanations in software engineering: The pragmatic point of view. *Minds and Machines*, 20(2), 277-289.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69, S342-S353.
- Kroes, P. (1998). Technological explanations: The relation between structure and function of technological objects. *Techné*, 3, 18-34.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70, 556-567.
- Pitt, J. C. (2009). Technological explanation. In Meijers, A. W. M., Gabbay, D. M., Thagard, P., & Woods, J. (Eds.), *Philosophy of technology and engineering sciences* (pp. 861-879). Amsterdam: Elsevier.
- Van Bouwel, J., & Weber, E. (2002). Remote causes, bad explanations. *Journal for the Theory of Social Behaviour*, 32, 437-449.
- Van Bouwel, J., & Weber, E. (2008). A pragmatist defense of non-relativistic explanatory pluralism in history and social science. *History & Theory*, 47, 168-182.
- Weber, E., & Van Bouwel, J. (2007). Assessing the explanatory power of causal explanations. In Persson, J., & Ylikoski, P. (Eds.), *Rethinking explanation* (pp. 109-118). Dordrecht: Kluwer Academic Publishers.
- Woodward, J. (2000). Explanation and invariance in the special sciences. *British Journal for the Philosophy of Science*, 51, 197-254.
- Woodward, J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science*, 69, S366-S377.

Endnotes

¹ I use Glennan's (2002) definition of mechanisms because I agree with de Ridder (2006b, p. 92) that it captures a number of important aspects of mechanisms, and because I see no reason to reject this definition. De Ridder does formulate some criticisms on Glennan (2002), but these focus on Glennan's characterization of a *mechanical model*, and not on his notion of a *mechanism*.

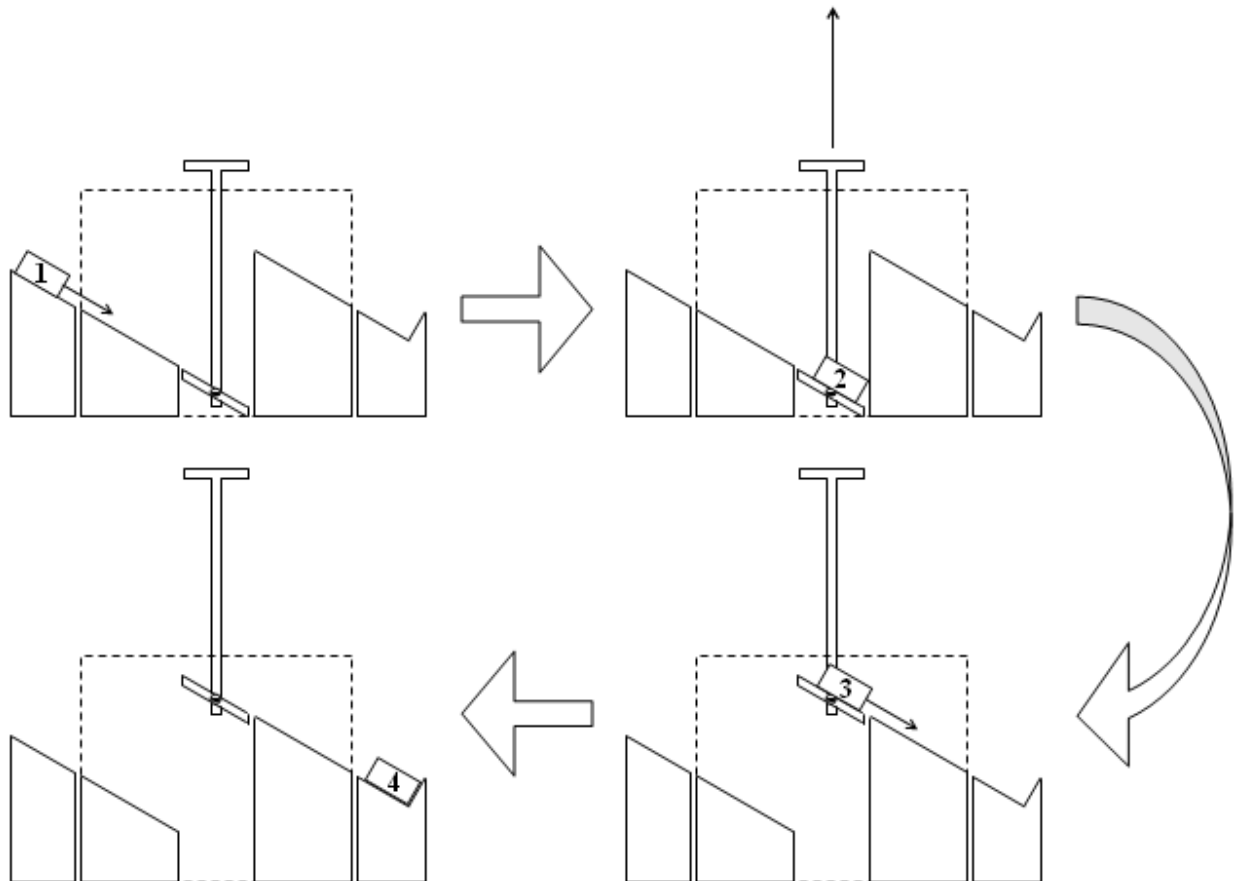
² This corresponds with Pitt's (2009) account of technological explanation. According to Pitt, which explanation should be offered, depends on the audience for whom the explanation is meant, on what would satisfy them, and on the specific why- and/or how-questions the explanation is supposed to answer.

³ Similar explanatory strategies are presented in Bechtel & Richardson (1993, p. 18). The top-down strategy is Bechtel & Richardson's synthetic strategy applied to an artifact. The bottom-up strategy, on the other hand, is not entirely the same as Bechtel & Richardson's synthetic strategy applied to an artifact. The difference is that the bottom-up strategy includes giving information about the physicochemical make-up of the components, while the synthetic strategy does not. Cummins (1975) also presents two explanatory strategies, the analytical strategy and the subsumption strategy, that are similar to, respectively, the top-down and the bottom-up strategy. For the differences between Cummins's explanatory strategies on the one hand, and the top-down and bottom-up strategy on the other hand, see de Ridder (2006b, p. 87n).

⁴ I do not assume that F- and S-components are necessarily different, because often, they are not. My point is that in a top-down explanation, we can be sure that the components referred to are F-components, while it is not guaranteed that they are also S-components. In a bottom-up explanation, the components referred to are necessarily S-components, but they are not necessarily F-components.

⁵ For counterfactual accounts of causation, see, e.g., Lewis (1973) and Woodward (2000, 2002).

Figures



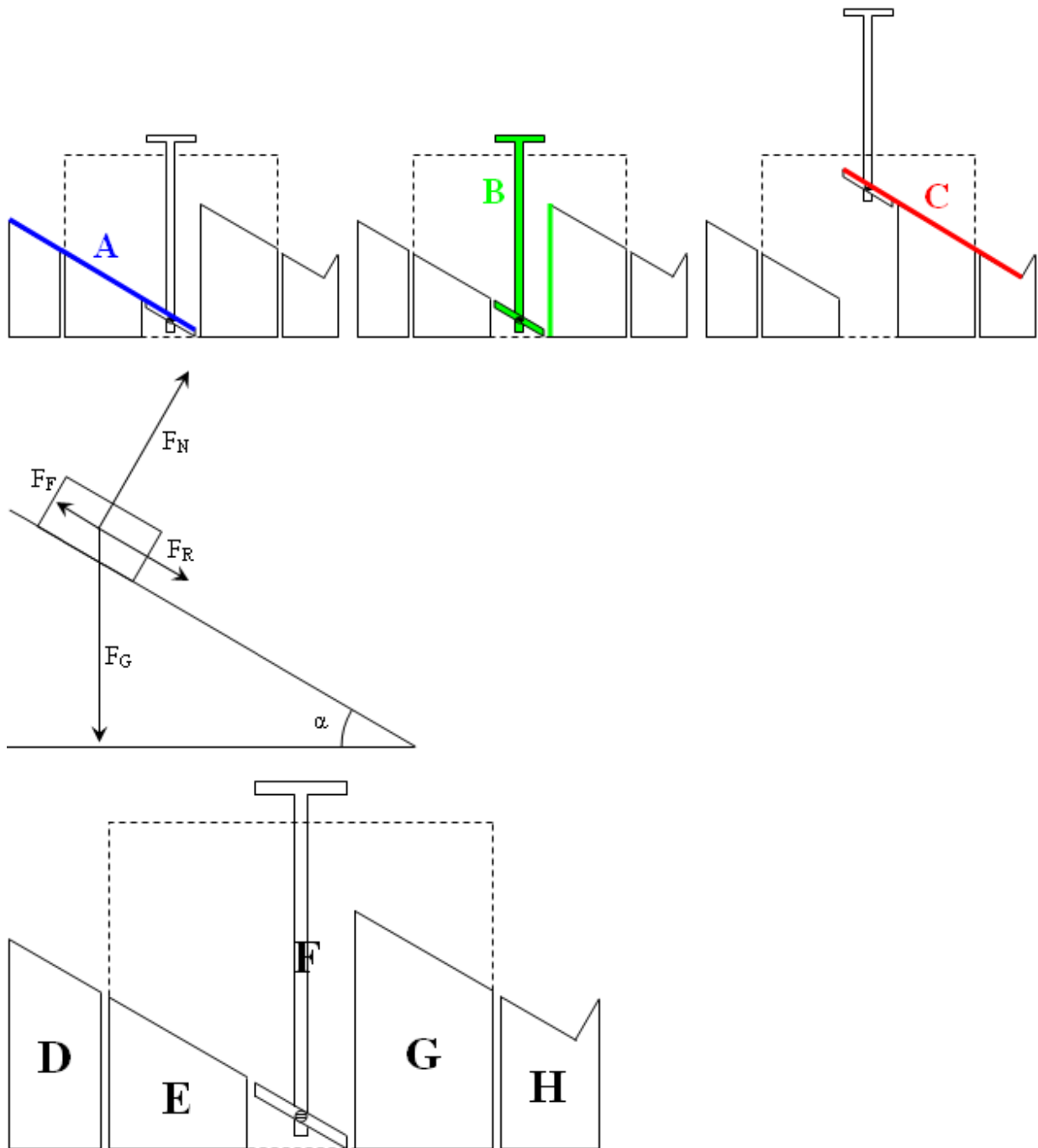


Figure captions

Figure 1. Simple mechanism.

Figure 2. Artifact components to which the top-down explanation refers.

Figure 3. Forces exerted on a box on a slope.

Figure 4. Artifact components to which the bottom-up explanation refers.