Productive Idealizations for Scientific Understanding: A Case Study in Effective Theories

Karla Weingarten

RCNP & ISiS, Radboud University, Nijmegen.
Email: karla.weingarten@ru.nl

January 20, 2026

Abstract

Scientific models often rely on idealizations to obtain insights into target phenomena, which distort and simplify their targets and are therefore commonly regarded as false. This conflicts with the idea that if we take the goal of modeling to be the ability to make inferences about a target, idealized models should, prima facie, be on equal footing with models that maintain a closer correspondence to the target. In this article, I argue that we can best capture the contributions of idealizations by assigning to them, besides an extrinsic content associated with the relation between model and target, also an intrinsic informational content. The latter is generally not dependent on the target but grounded in independent scientific conclusions, and hence true. This way idealizations can actively contribute to understanding and be called productive. I support this claim with a case study of effective field theories (EFTs) in quantum field theory, which are non-fundamental theories with a limited domain of applicability. Their intrinsic cutoff scale, I argue, functions as a productive idealization that carries information contributing crucially to understanding. An EFT can thus be better at providing understanding of the specific phenomenon studied than a UV-complete theory. Understanding is not correlated to the degree of idealization, or in an EFT context, the fundamentality of the theory.

Contents

1 Introduction

Within the plethora of models that scientists employ in their daily research activities, one would be hard-pressed to find a model that is an exact copy of its target. Most models distort their targets, and they do so deliberately: they omit unnecessary details, change scale and size of the system, or isolate a specific mechanism. This 'mis-match' has also been recognized by those philosophers interested in the nature of the representative relation between a model and its target. Similarity or more sophisticated notions for the mapping between a model and target, are rejected as grounds for representation, which is instead based on the model's capacity to facilitate inferences about the target (Hughes 1997; Frigg and Nguyen 2021b; Suárez 2024). These inferences can then lead, for example, to scientific understanding. What is more important than the depiction of target elements in the model is what can be learned with the model about the target. In other words, those who endorse an inference-based account of representation are more interested in the model-to-target pipeline, rather than its reverse, the mapping of features target-to-model.

Important classes of model distortions are idealizations, simplifying distortions of the target, and approximations, model-internal simplifications. As they are distortions, modifications of the original target, idealizations are generally considered to be falsehoods. Hence, they are often credited only a peripheral role in the inferences made with a model. For example, they are thought to be markers of target irrelevancies only, or to not appear in the model's explanatory or deontic core (see, for example, Elgin 2010; Rice 2016; Frigg and Nguyen 2025). Because idealizations, when focusing on their derivation via the target-to-model pipeline, are considered falsehoods, they cannot actively contribute to a model's success. This would automatically render idealized models epistemologically inferior to more complete alternatives.

In this paper, I push back against this view and argue that we should focus on the model-to-target inferences also when we ask whether idealized models can provide understanding. Building on work in the literalism vs. factivism debate (Lawler 2021; Frigg and Nguyen 2021a), I argue that idealizations carry information in two distinct ways: in addition to the extrinsic, target-related aspect that is considered a falsehood, idealizations also have an intrinsic informational content that can generally be considered to be independent of the specific target system studied, and is hence true. This makes idealizations not only distortions of the target, but also carriers of true information.

This also has consequences for the understanding a model can provide: As carriers of true information, idealizations and approximations contribute actively to scientific understanding, explicitly improving it without reference to something else. They can thus be labelled 'productive'. Consequently, more complete models are not immediately superior; idealized models can provide understanding just as well.

This framework is in line with many idealized models or theories used in scientific practice. In this paper, I will specifically discuss the example of effective field theories (EFTs). These non-fundamental theories are designed to lose validity at specific energies. They operate with a cutoff Λ, below which they provide an accurate description of the relevant phenomena, but above which they lose validity. This cutoff needs to be introduced in order to apply a renormalization procedure through which the theory is rendered predictive.

EFTs are our current best theories in high-energy physics. Their prevalence is in tension with an important guiding principle that has long inspired physics research: to find a final theory of everything that can describe all physics at all scales. Over the last decade, due to lack of empirical findings to substantiate proposed unifying theories, the idea that high-energy physics may be inherently effective itself has become more and more recognized. Consequently, the question arises what this effectiveness implies for our scientific understanding of the physical world. This practical situation mirrors the theoretical considerations made in this paper: Can we only obtain understanding with fundamental theories, or do the idealizations and approximations actively contribute to understanding obtained with EFTs? Following my productive idealizations approach, I conclude that the latter is the case. The idealizations and approximations present in EFTs can be characterized as productive, too. Understanding is not dependent on the degree of fundamentality ascribed to a theory.

This paper proceeds as follows: Section 2 begins by introducing the importance of idealizations in scientific modeling as well as in scientific understanding, and concludes with a characterization of idealizations as carriers of information which contribute productively to understanding. Section 3 provides an overview over effective field theories, their special effective nature, and their status within physics and philosophy. Finally, section 4 argues that effective field theories contain such productive idealizations as their intrinsic cutoffs, and that EFTs therefore provide scientific understanding independent of their degree of fundamentality.

2 The Role of Falsehoods for Scientific Understanding

In this section, we will discuss the role of falsehoods for scientific understanding. A specific kind of falsehood present in models are deliberately introduced misrepresentations of the represented system. In this section, we will take a closer look at how approximations and idealizations are characterized as misrepresentations (section 2.1). After an introduction to scientific understanding generally (section 2.2) and its connection to truth specifically (section 2.3), I conclude with a proposal of how misrepresentations can be considered in a more positive light (section 2.4) and how they can play a productive role for scientific understanding (section 2.5).

2.1 Representation, misrepresentation, and accuracy

The use of models in ubiquitous is science. Some system A is used as a surrogate to obtain insights into another system B. In other words, in such a surrogative use, A represents B. A simple criterion for why some A could represent some B is to ask that they be similar to one another. However, in order to restrict representation to intended cases, excluding accidents or chancy resemblance, one usually poses some further requirements on systems A and B beyond mere 'similarity'. One prominent strand in the philosophy of scientific modeling bases representation on denotation, i.e., representation requires that elements of A denote, or indicate, elements of B. A fairly minimalistic account of representation, which will suffice for the purposes of this paper, is presented by Suárez, who argues that a model A represents a target B if its 'representational force' points toward the target, essentially requiring its denotation, and if it additionally allows competent and informed agents to draw inferences about B (Suárez 2004, p. 773; see also Hughes 1997; Frigg and Nguyen 2022; Suárez 2024).

A successful model use should not only allow inferences to be made about the target; these inferences should arguably also not be possible within the target system itself. Otherwise, the construction of a model would be unnecessary. On the contrary, differences between the model and target can be exploited under specific constraints. In order to do so 'safely', we need to investigate how exactly a model misrepresents the target.

An important form of deliberate misrepresentation is idealization. Contrary to an abstraction, which refers to a simple omission of some target feature considered irrelevant and which remains neutral with regard to truth, idealizations are explicitly false about their targets. Generally, an idealization is considered to be a simplifying assumption that is made within the model system or its use, and which deliberately distorts the original target (see e.g. Frigg 2023; Shech 2023). If idealizations were asserted to be literally present in the target, this would be false. Usually, idealizations are introduced with specific goals or values in mind, meaning that misrepresentation is in some sense relative to an intention, allowing an agent to view the target system through a specific lens.

Idealizations are distinguished from approximations. These are simplifications that are made purely within a model, based on goals specific to the modeling, such as computational ease, and which are not distortions of the model-target relationship. For example, assuming that $\sin\theta \approx \theta$ for small θ is an approximation that can be made in any model independently of the target represented. Hence, while it can sometimes be interesting to de-idealize some models to arrive back at the complete target, we are not interested in de-approximating a model (Frigg 2022, ch. 11). In more general terms, idealizations are about the distortion of the relationship of two systems and then replacing one with the other, whereas an approximation is a modification within a system itself, without anything to be replaced. Therefore, approximations can be considered propositional, while idealizations are referential (Norton 2012).

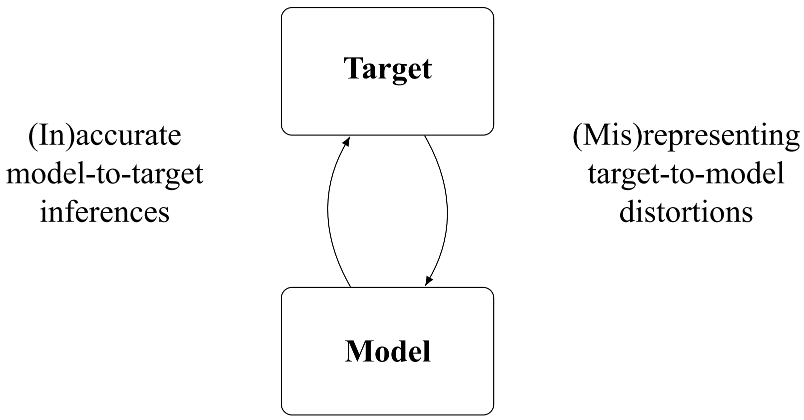


Figure 1: Two-way target and model relations.

If idealizations are distortions of the target present in the model, then how far can a target be distorted before we have to admit that there is no representative relation between model and target anymore? In other words, what makes a representation accurate? It is important to note that a loss of a representative relation between model and target is not due to the former misrepresenting the latter, in the sense of distorting it, if we define representation as being dependent on the ability to make inferences. Instead, inaccuracy will be considered the failure of being able to impute model conclusions onto the target system (Frigg and Nguyen 2022). As long as we do not conclude target falsehoods with the help of the model, we can consider it to be accurate. An idealized model therefore does not necessarily have to be inaccurate—accuracy and misrepresentation are independent of one another, as can be seen in figure 1.

Nonetheless, this does not mean that the conclusion of target truths from idealized models is entirely straightforward. Most accounts consider idealizations as derivatives. They are distortions of completely known targets. Elgin, for example, argues that idealizations are essentially derived from the target, being "telling instances" of a target property of special importance to the current modeling process. They "facilitate recognition of those aspects and appreciation of their significance" (Elgin 2010, p. 9). This dependency on the target ensures sufficient connection, even when the idealization is highly selective with what it exemplifies.

An advantage of this derivative view of idealizations is that a division of models into idealized and nonidealized parts is straightforwardly achieved via a simple comparison of the model to the target. Accounts of scientific modeling often assume this so-called decomposition strategy, which, according to Rice, can be broken down into are three decomposition assumptions. Firstly, we assume that the target can be decomposed into relevant and irrelevant features. Secondly, we assume that the model can be decomposed into distorting and non-distorting parts. Thirdly, we assume that the distorting model parts can be mapped onto the irrelevant parts of the target, and the non-distorting ones onto the relevant ones (Rice 2019, pp. 181–82). This decomposition strategy then implies that all 'work' in a model is done by its non-distorting features.

However, in recent years, it has been argued that the role of idealizations can go beyond the passive highlighting of specific, non-idealized target features. Rice himself argues that models can rarely be decomposed so that they meet all three decomposition criteria. If decomposition fails, Rice concludes, we cannot decompose a model, and thus cannot single out idealizations or their contributions. Instead, we should look at models holistically.

If we cannot single out idealizations, we cannot ensure the appropriateness of a model by it being derived from a (not relevant) target feature, opening up again the question of accuracy. According to Rice, accuracy is ensured by the model and target being in the same universality class, implying that they have the same high-level behavior (see also Batterman and Rice 2014).[1] If this is given, we can make accurate conclusions about the target with the model: "All that is required is […] that those macroscale patterns of counterfactual dependence will be preserved" (Rice 2019, p. 202).

Separating the target-to-model distortions from the model-to-target inferences as done in figure 1, we can recognize that it is sufficient that the inferences they facilitate are true, independently of a derivative relation to the target. If accuracy is dependent on inferences only, we can be more liberal about our model and its contents. We can especially admit idealizations that cannot be straightforwardly derived from the target. Potochnik, for example, takes this to heart and defines idealizations merely as falsehoods, without any reference to the target (Potochnik 2020a, p. 935; Potochnik 2020b). How exactly such a justification of accuracy can work will be discussed in more detail in section 2.3.

2.2 What is scientific understanding?

Science, it is generally assumed, strives to extend our knowledge of the world. Knowledge, famously, can be taken to be some form of justified true belief. In recent years, however, philosophers have highlighted that the possession of knowledge or facts is not enough, but that scientists also need to understand these facts. Understanding, as Elgin proposes, is "a grasp of a comprehensive body of information that is grounded in fact, is duly responsive to evidence, and enables nontrivial inference, argument, and perhaps action regarding that subject the information pertains too" (Elgin 2007, p. 39). It focuses on an individual's cognitive achievement of understanding why some phenomenon occurs. It is subjective, non-transferable, and comes in degrees: my understanding of some subject matter differs from that of the next person.

There are different approaches to how understanding is obtained in practice. Some emphasize the relationship of understanding to explanations of the same subject matter, such as Strevens,

[1] In statistical and many-body physics, universality describes the fact that systems different at the micro-scale exhibit the same behavior at a much higher macro-scale. These systems then lie in the same universality class. Batterman and Rice draw on this notion to argue that a minimal (i.e. highly misrepresenting) model is explanatory of its target if they are in the same universality class, as it can then be shown that those details that differ between model and target are irrelevant, and that those that are shared between them are relevant (Batterman and Rice 2014, p. 365).

who argues that understanding is the grasping of an explanation (Strevens 2008). Others argue that these two should be considered independent (to varying degrees), as the success criteria for understanding differ from those for an explanation (Lipton 2009; Verreault-Julien 2019). In this paper, I will follow de Regt, who, emphasizing the skill aspect of understanding, advocates a view that features the intelligibility of a theory to its user as the main characteristic that ensures it can provide understanding. An intelligible theory, he argues, allows one to recognize a theory's qualitative consequences at a glance, and is distinguished from mere 'guesswork'. A theory is not intelligible intrinsically, but is so to a specific agent in a specific context. If that agent finds a theory intelligible, they can grasp dependencies without complicated (mathematical) analysis and apply the theory and its consequences elsewhere. A phenomenon, de Regt continues, is understood if an agent possesses an explanation of that phenomenon that is based on an intelligible theory which, in addition, is empirically adequate and internally consistent (de Regt 2017, p. 92).

2.3 Understanding and truth

If scientific understanding is best characterized as a skill, rather than a different form of knowledge, the question arises whether the same criteria that we pose for knowledge or explanation are similarly suitable.

According to de Regt, whether a theory can provide understanding depends first and foremost on whether it is intelligible. This implies that false theories are not necessarily excluded from providing understanding, as intelligibility is independent of truth (de Regt 2015). Specifically, de Regt argues that the intelligibility of a theory can even be enhanced by deliberate falsehoods, such as idealizations or simplifications.

In order to accommodate falsehoods as understanding-providing, one can give up the expectation that scientific understanding has to be factive, that is, that it is brought about exclusively by veridical statements (Elgin 2017; de Regt 2015). Then, we can admit, for example, idealizations as understanding-providing. We can also argue that theories such as the ether theory have provided genuine understanding, even though we now know that they are wrong. Hypothetical scenarios about the possible behavior of some system or related ones can provide understanding, too, independently of whether they can ever actually be instantiated. The same holds true for models that refer only to possible explanatory factors (Bihan 2016; Reutlinger, Hangleiter, and Hartmann 2018; Rice 2025).

The rejection of factivism as a criterion for understanding raises an important question: if falsehoods are admitted as understanding-providing, how do we ensure that we obtain genuine understanding of the actual target? de Regt, who prioritizes intelligibility over truth, argues that all we need to require for genuine scientific understanding are empirical adequacy and internal consistency. A direct relation to something true is not needed, as long as the falsehoods improve the intelligibility of the theory.

An alternative is given by rejecting the literalism rather than the factivism of understanding. Instead of claiming that falsehoods are 'true enough'—distortions exemplify those features they share with the facts and are always considered in relation to the true ideal (Elgin 2007, p. 41)—we simply accept falsehood as being false. Simultaneously, we reject the falsehood as being part of the content of understanding. Lawler, for example, argues that falsehoods can play crucial and irreplaceable roles in the acquisition of understanding, but as vehicles only. Falsehoods, she argues, are introduced deliberately to provide epistemic access to what she calls the content of understanding, and they often do so in a way that cannot be achieved by truths. The content of the understanding itself is then made up of truthful sentences about the target of the understanding, often as related to the falsehood. For example, the ideal gas law is known to be an idealization and to make literally false assumptions. However, we can use it to obtain truthful understanding about real gases. This understanding consists of statements on how the behavior of a real gas we experiment on differs from the predictions of the ideal gas law. Approximations can be related to the target in a truthful yet not literal reading in the same manner (Lawler 2021).

Relatedly, Potochnik, who understands idealizations primarily as falsehoods with no presumed target connection, justifies their use via their epistemic role. Idealizations, she claims, provide understanding by helping to make causal patterns visible in a way that truthful representations could not. Understanding of a phenomenon then comes from grasping such causal patterns. Idealizations are not "a first step toward scientific understanding, to be improved on later, but full participants in the epistemic success of achieving scientifc understanding" (Potochnik 2020a, p. 935). Thus, idealizations have a unique epistemic value despite not being part of the causal patterns that provide the content of understanding.

It appears that while workarounds can be created to allow idealizations to contribute to modeling, the fact that they are target falsehoods hinders them from contributing to the actual content of understanding.

2.4 Idealizations as carriers of true information

In the following section, I advocate a notion of idealization that is clearly distinct from the notions of misrepresentation and distortion. This approach stands in the tradition of Frigg and Nguyen 2021a, Potochnik 2020a, Lawler 2021 and other non-literal interpretations of idealizations. The resulting separation allows us to see the value of an idealization without blurring it with its falsehood stemming from a target distortion.

Statements such as 'the molecules in a gas are point particles' or 'the electrons in an atom orbit the nucleus on fixed, circular paths' are generally considered idealizations. We know that molecules in a gas are spatially extended, and that electrons cannot orbit the nucleus because this means they would sooner or later crash into the atom's core. In such cases, where we know the true behavior of a system studied, for example because we can observe it experimentally, it is natural to consider idealizations as distortions, misrepresentations of these known systems. As assumptions about their targets—real gases and real electrons—these statements are false. We can, however, also consider these statements independently. Take, for example, the point particle assumption. If we look at it without the real gas target in the back of our minds, we get a statement along the lines of: 'Considered at a macroscopic scale and under normal thermodynamic conditions, the influence of the physical extension of gas molecules on the gas volume is negligibly small compared to that of the pressure of the system'. When we recognize the autonomy of this statement when evaluating its truth, we can see that this is a true sentence.[2] Just because the statement is composed with reference to some scale (e.g. a macroscopic scale) and thus in some sense is restricted in its applicability, or because an idealization considered as a distortion of some target system is, per definition, false about that target system, it does not follow that this makes the statement, considered by itself, false. Even more, the insistence on the falsehood of an idealization or an approximation obscures the fact that idealizations tell us things that go beyond what is present in the target itself. In the case of the ideal gas law, the point particle idealization can tell us something about the forces present in a gas, the scales at which they act, and their relative magnitudes. This information is independent of the specific real gas we are modeling, but when put in relation to it, it provides us with relevant knowledge about this gas.[3]

Idealizations and approximations are thus not only derivatives; they are also information carriers. Idealizations carry information in two ways: intrinsic information about themselves, and extrinsic information about the model's relationship to the target. Intrinsically, the idealization or approximation is a truthful statement of the behavior of a system at some given scale. The familiar distortion of a target, however, makes an idealization a falsehood extrinsically. Using an appropriate interpretation, we can exploit both the intrinsic and extrinsic content of the idealization with respect to the target. As Frigg and Nguyen put it: "when we interpret the 'there are no interactions' feature of the model as representing the feature of the target that 'the interactions aren't difference makers with respect to the relationship between volume and pressure' we arrive at an accurate representation" (Frigg and Nguyen 2021a, p. 2444).

The idea of an intrinsic truth has always tacitly been accepted in the context of approximations, which are classified by Frigg as model-internal and thus not derivative. The statement that $\sin\theta \approx \theta$ for small θ and up to some degree of accuracy is a true statement independently

[2]It lies beyond the scope of this paper to establish a precise notion of truth within the context of scientific modeling. Instead, in this paper, I consider true statements, or facts, as referring to statements that are empirically adequate in their given context. Accuracy is maintained when a model allows only inferences that preserve these truth, i.e. that do not allow the conclusion of falsehoods.

[3]It stands to reason that the autonomy of idealizations that I advocate for here implies a realist stance towards them. While a detailed discussion goes beyond the scope of this article, as I argue that idealizations can be considered equivalent to other forms of contextual information, a stance that denies the realism of idealizations yet is realist about non-idealized- statements appears hard to defend.

of the system that it might be applied to or used within. The intrinsic truth value of idealizations can be thought of analogously: the content of an idealization, considered autonomously and not in relation to a target system, can be true while, at the same time, the target-related extrinsic truth value is false. Considering idealizations as independent statements also accommodates idealizations that cannot be mapped directly onto target elements, as the derivative target relationship is not the idealization's defining quality. We thus do not run into the decomposition problem as noted by Rice 2019.

I conclude that it is essential to clearly distinguish between a misrepresentation on the one hand, which refers to a target derivation, and an idealization, as an information carrier, on the other hand. Both have their own, independent content and truth value. Idealizations are not only pointers to target irrelevancies. They are not only telling instances of the target features that are non-distorted in the model, they are also telling of the idealization's actual content. This intrinsic content is true. When this content appears in a modeling situation, it is put into relation to a target and subsequently interpreted also as a distortion of this target. This is an independent act that does not necessarily define the content of the idealizations itself. Idealizations, when put into a specific representative context, can become subject to an analysis of its relation to the target, i.e. what elements can be derived from it. However, this relation does not have to, and often is not, the defining quality of an idealization. Rather, an idealization encompasses independent, scale and context-dependent knowledge that exists also independent of a specific target. They are more than just target derivations, they are tools deliberately put into relation with the target to achieve a specific goal.

2.5 Productive idealizations for scientific understanding

If we take target falsehoods like idealizations to have a true intrinsic content independent of the falsehoods that result from their target relation, we are able to make better sense of the way they contribute to understanding. First and foremost, they add additional content to the model with which we want to understand the phenomenon. They contain information about the relevant context and the scale at which the model operates. Recognizing that these elements were introduced deliberately improves our understanding by providing a better grasp of the phenomenon or, in de Regt's terms, improves the model's intelligibility to the agent. Indeed, the introduction of idealizations and approximations might even allow the construction of an intelligible model in the first place. This is especially true of approximations, whose primary use is to improve the intelligibility of a system internally.

The improvement of understanding that is obtained through the introduction of target falsehoods is not (only) due to them pointing at target irrelevancies, or at any target features for that matter. They not only contribute in this passive sense, but also actively, by introducing contextual information about the target. In many cases, the idealization is obtained not by abstracting away from the specific target system currently studied. Instead, it is information that has been established scientifically in an independent procedure, and which can be straightforwardly used in the study of many different target systems. The existence, construction, or analysis of an idealization or approximation is not dependent on the specific target phenomenon studied. This independence should also be emphasized in their role of providing understanding.[4]

Nonetheless, we do not want to infer false conclusions about the target or, in this case, gain false understanding. The prima facie falsity of idealizations and approximations with regard to the target means that, while intelligibility of the theory is increased, maintaining accuracy still requires a sophisticated reading of the model. After all, a theory is only truly intelligible if the understanding provided is empirically adequate and consistent, and hence inferences drawn should not oppose this (de Regt 2015). As discussed in the previous section, we need to distinguish between the internal and external readings of an idealization. In the context of understanding, we connect this with a distinction between the content and the vehicle of understanding, as in the framework proposed by Lawler, but we make the information that idealizations carry autonomously explicit (Lawler 2021). We can consider distorted models, in which idealizations are interpreted as target-falsehoods, to act as 'trailblazing vehicles' for target understanding, but the target-falsehoods themselves do not enter the content of the understanding. Rather, the autonomous and contextual interpretation of an idealization, being true internally, enters the content of understanding, helping the agent to make the theory intelligible and use it, for example, for further model building. Figure 2 shows an adaptation of the previous figure 1 that accommodates this distinction. To be more precise, statements about possible behavior of a system under (varied) idealized conditions, comparisons of an observed behavior of the phenomenon to predicted model behavior, and similar statements that relate model and target while keeping in mind the scale and context of the introduced target falsehoods, populate the content of understanding. Taking into account also the positive truth value of idealizations when considered autonomously, which is added to the features a model might take from its target, we can conclude that the understanding so obtained is true, and consequently maintains the accuracy of the representation. Additionally, it stands to reason that keeping the modeling goal and intentions of the model user fixed, it seems that the same (or at least, a comparable) understanding would not be achievable without idealizations or approximations.

For an illustration, consider again the ideal gas law and the statement that 'considered at a macroscopic scale and under normal thermodynamic conditions, the influence of the physical extension of gas molecules on the gas volume is negligibly small compared to that of the pressure of the system'.[5] The point particle assumption is not present in the target, waiting to be

[4]See also Fletcher 2019's notion of minimal approximations and idealizations, whose removal does not improve the model. I extend on this approach by allowing an explicitly positive contribution for approximations and idealizations, not only a non-negative or neutral one.

[5]For the historical development and the several phenomenological laws involved in the formation of the ideal gas law, see, for example, de Regt 2017, pp. 31–35.

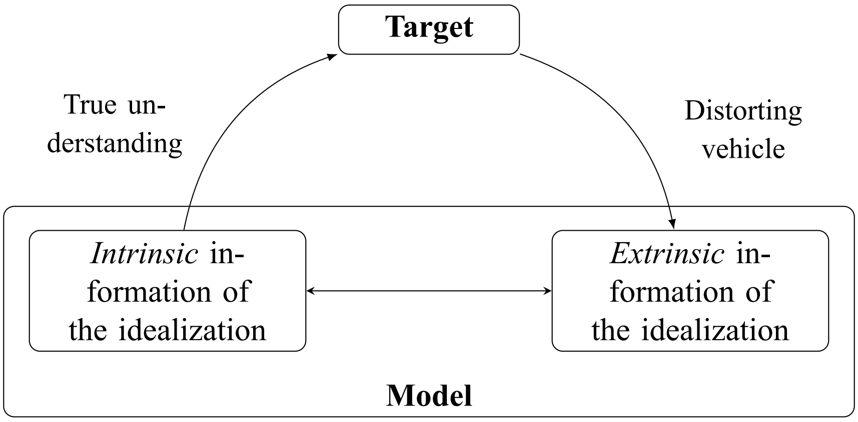


Figure 2: Adaptation of figure 1 that incorporates the two elements of idealizations that can figure in models.

found. It does not make the ideal gas law an intelligible theory by abstracting away from the behavior of real molecules in real gases. It is not this procedure of abstraction that helps us understand. Instead, the understanding stems from the true additional information that is being introduced, and which is placed in relation to the modeling aim and the modeling scale. The distorted ideal gas model is a vehicle for understanding, in which the point particle idealization specifically plays a productive role in understanding the behavior of the real gas target through the context, the comparisons, and especially the information encapsulated in it. While the existence of an abstracting relation should certainly not be denied, we do not analyze each individual real gas to see whether we can indeed obtain a point particle idealization; it is an established and true piece of information autonomously. Misrepresentation, the distorting relation between model and target, is not the most important feature of the idealization, and it is not what provides (the most) understanding. Understanding is dependent on the other direction of the model-target relation: it is provided by the non-literal model-to-target inferences, not the possible target-to-model abstractions. In the end, what matters for scientific understanding are the agent's ability to make accurate inferences, which increases when more contextual information is introduced and the theory's intelligibility is increased. More creative approaches to modeling can be admitted as understanding-providing on this account: Something can be a misrepresentation yet true.

As a consequence of the decomposition problem introduced earlier, Rice concludes that we should not believe that "the accurate parts of the model are what 'do the real work' while the inaccurate parts of the model are justified by distorting only what is known (or assumed) to be irrelevant" (Rice 2019, p. 196). While I agree with this analysis, I disagree with his conclusion that we cannot single out the contributions by target-distorting and target-faithful model elements, and that models are best understood as providing understanding holistically. Rice emphasizes the role of approximations and idealizations as necessary steps for further holistic model use, and especially for the consequent application of (mathematical) modelling techniques (Rice 2019,

p. 197). I agree that model-internal modifications stemming from approximations are potentially impossible to single out within a final model product, not to mention a possible mapping to target irrelevancies. Idealizations, on the other hand, given the explicit positive contribution they make by providing additional input about the target, can, however, indeed be singled out as contributing productively to understanding, and autonomously so.[6] The utility of a model should be dependent not on its (mis)representation, but on its ability to make inferences about the target. We should therefore be less interested in whether model and target features can be decomposed and mapped against one another, and more interested in which model elements contribute positively to model conclusions. The latter does not require the former.

My claim that idealizations and approximations provide understanding autonomously is similar to that of Potochnik, who argues that idealizations provide understanding directly, without further steps required (Potochnik 2020a, p. 937). My understanding of idealizations is, however, more comprehensive. Not only are there no further steps, there is also no reference to something 'more true' that needs to be referenced for a misrepresentation to provide understanding. Distortions contribute to understanding positively, in that the same understanding cannot be achieved without them. Importantly, they contribute to understanding independently, in that they do so without being derived from the target. I thus agree with Weingarten that such idealizations can be called productive (Weingarten forthcoming).

In the following sections, we will use the framework introduced here to show that effective field theories provide scientific understanding not despite, but because they contain idealizing intrinsic cutoffs that limit the scope of their applicability.

3 What are Effective Theories?

Effective field theories are non-fundamental theories that provide a tailored description of a physical process of interest. Physicists use them to study phenomena at only those energy scales they deem relevant by explicitly including contributions up to a cutoff Λ but ignoring contributions at higher energies (or, equivalently, short distances[7]). These effective theories are empirically and predictively successful below Λ but not above; by construction, they do not strive to be valid there.

The fundamental idea behind EFTs is the reduction of the degrees of freedom that need to be considered by filtering out those that are irrelevant at a given scale. In order for this to work, a so-called separation of scales, the independence of high and low energy contributions, is necessary.

[6] A special case are idealizations that have a speculative nature (think of systems with irregular limits, AI models, or similar). While I am positive that these kinds of idealizations can also contribute to scientific understanding, it is hard to define statements which are true internally, and they thus do not contribute to understanding in the same way.

[7] In natural units, where $c = 1$, mass and energy have the same dimensions, and length is anti-proportional to energy.

Then, renormalization methods can be used to construct a predictive EFT.

As we will see, effective theories, and effective field theories in particular, have distinct properties that distinguish them from 'fundamental' or 'complete' theories. They are, by design, incomplete; the distortions are introduced deliberately. There appears to be a trade-off between different theoretical virtues, such as the completeness of the representation, the tractability of the theoretical model, or comprehensibility of the description (Weisberg 2007; Norton 2012; Shech 2023).

In this section, I will provide a short introduction to effective field theories and how to obtain them using renormalization methods (section 3.1). I will then explore the special 'effective' nature of EFTs in more detail (section 3.2).[8]

3.1 Different types of EFTs—how to derive them, and how to use them

The construction of an effective field theory to describe some phenomenon consists of two important steps. First, a suitable Lagrangian needs to be identified, which in a second step needs to be regularized and renormalized to make it useable. This process introduces a cutoff beyond which the theory is inapplicable, accounting for an EFT's 'effectiveness'.

The first construction step can proceed in two different ways, either top-down or a bottom-up, depending on the phenomenon studied (Hartmann 2001; Franklin 2018a; Koberinski and Fraser 2023). In the case of a top-down construction, a fundamental theory valid at high energies, is available. Its Lagrangian, adapted to include only those elements needed to describe the phenomenon to be studied, serves as a basis for the construction of an EFT applicable at lower energies. Bottom-up EFTs, on the other hand, are constructed without reference to a fundamental framework theory. Instead, physicists start with heuristic assumptions, such as empirical data from observed phenomena (including particle properties, symmetries, etc.) or theoretical principles (such as naturalness, symmetries, etc.) one assumes to hold for the phenomenon investigated, to construct a first Lagrangian.

These initial Lagrangians are not yet predictive. To be able to calculate things like scattering amplitudes or correlation functions, integrals need to be computed which often diverge. To take care of these infinities, we need to both regularize and renormalize the theory. In the following, I will provide a schematic overview over these techniques, with emphasis on explaining the concept involved, rather than on mathematical completeness. For a more technical introduction, see, for example, Peskin and Schroeder 1995; McComb 2007; Duncan 2017.

A simple way to regularize any integral is to cut off the theory above some energy Λ, and to only consider the non-diverging terms that appear below. This cutoff is introduced 'by hand',

[8]For more work on the effective nature of EFTs, see, for example, Franklin 2018b; Williams 2019; Rivat and Grinbaum 2020; Fraser 2020; Palacios 2022; Koberinski and Fraser 2023; Koberinski 2024.

based on the assumption that the physics above Λ is not relevant for the interactions considered and that the scales above and below the cutoff are separable, i.e., that high energy contributions can be described by effective variables at low energies. Indeed, we are justified in making this assumption, as was proved in 1975 (Appelquist and Carazzone 1975). There exist also other regularization methods, such as dimensional regularization, which make use of the same principle.

As a next step, the theory has to be renormalized. In the perturbative sense, historically the first renormalization scheme to have been used, this means that we take the couplings in the initial integral (now dependent on the regulator) and shift them in such a way that diverging elements of the integral are absorbed into counterterms, rendering the integral finite. If a less than infinite number of such terms needs to be introduced, the theory is perturbatively renormalizable.[9]

However, the theory is still dependent on the cutoff, and as this cutoff is not given by the theory itself but introduced by hand, the whole scheme seems somewhat arbitrary. One way to remove the dependence on the cutoff Λ is by fixing the values of the parameters after renormalization with so-called renormalization conditions, which could be, for example, obtained experimentally. If the theory is renormalizable, the limit $\Lambda \to \infty$ can be safely taken in the end.

Still, renormalization underlies the idea that physics at lower energy scales is independent of that at higher scales. It thus seems natural to investigate the behavior of a theory when the value of the cutoff is changed. The resulting differential equations that show the behavior of the theory dependent on the energy scale μ, $\mu < \Lambda$ are called the renormalization group (RG) equations. They describe a 'flow' with the change of μ, which gives us an alternative, nonperturbative sense of renormalizability. I will call this approach the Wilsonian renormalization, after the physicist credited with its invention (Wilson 1975). A theory is renormalizable in the Wilsonian sense if the RG flow hits a finite fixed point for $\mu \to \infty$.

In addition to these two notions of renormalization, a pragmatic mixture of features of both approaches is used in physics practice to obtain predictions also with theories that are not renormalizable in either the perturbative or the Wilsonian way. I call this pragmatic strategy the Wilsonian perturbative renormalization. Here, an energy scale E_C characteristic of the phenomenon of interest is identified, the Lagrangian expanded as a power series of terms proportional to $(\frac{E_C}{\Lambda})^n$, and the appearing terms renormalized order-by-order by utilizing dimensional analysis and introducing counterterms (Georgi 1993; Manohar 1996). This yields a predictive approximation up to some desired order n that, using renormalization methods, renders also nonrenormalizable theories useful.

We can thus distinguish three types of renormalizability in total, all of which are independent of whether the initial Lagrangian was derived in a top-down or bottom-up manner. All effective

[9] We do not actually add any new terms, rather, we split the terms present in the Lagrangian in an advantageous way (Peskin and Schroeder 1995, p. 325).

theories are at least renormalizable up to some E_C with Wilsonian perturbative renormalization. Renormalizable theories are renormalizable also at all energies, either in the perturbative or Wilsonian sense, and sometimes in both. What unites the group of effective field theories is not whether they are explicitly dependent on a cutoff, or whether the cutoff could be taken to infinity or not. Rather, they are all founded on the assumption of scale separability, that is, the assumption that there is some value for Λ that splits the present physics in a relevant low-energy and an irrelevant high-energy part. Even if the theory is perfectly renormalizable, it can only be constructed with this assumption and the cutoff thus remains significant.

3.2 The "effective" nature of effective field theories

All EFTs have in common that they are constructed using a renormalization scheme and cutoff. They all operate on the assumption of scale separability and hence that physics at scales above the cutoff energy can be ignored or expressed effectively. The introduction of the cutoff Λ and the subsequently reduced scale is a misrepresentation of the actual physics. More specifically, the cutoff is an idealization as described in section 2.1, a simplifying assumption that is, technically, wrong of the target. We omit things in the description of our phenomenon that we know or expect to be present in the real-world phenomenon, and we do so because we know that while these high-energy contributions are not absent, they have a negligibly small influence on the phenomenon that we are interested in. The introduction of a cutoff reflects our knowledge of the phenomenon: contributions above the cutoff are irrelevant, uninteresting, inexpressible or experimentally inaccessible.

In addition to this idealization, we can see that, following the distinctions introduced previously, any effective field theory also contains an approximation. This is due to the renormalization procedure, with which we 'rewrite' the theory obtained after the cutoff is set in order to remove infinities and make the theory conceptually as well as computationally useable. While the introduction of a cutoff into a theory is a distortion of the phenomenon that it represents, the renormalization scheme is entirely internal to the EFT. It is employed as a mathematical technique to obtain finite predictions independently of the specific phenomenon the EFT descries. While we might be interested in what happens when we remove the idealization in an EFT—letting Λ go to infinity to obtain a UV-complete theory—it does not make any sense to try and undo the renormalization methods that were employed.

The fact that we cannot undo the distortions for most EFTs, and that they are applicable only up to and dependent on a cutoff introduced by hand, has prompted physicists and philosophers alike to make light the importance of effective theories (see, e.g., Redhead 1988; Williams 2019). Driven by a unificatory paradigm, physics has favored universally applicable UV-complete theories as the ultimate goal of its scientific endeavors. Effective theories, because of their limited applicability, are, following this account, considered nothing more than temporary tools. While

EFTs have been widely accepted on pragmatic grounds—they are generally very successful at providing accurate values when compared to experimental data—it is nonetheless assumed that, in the end, an EFT will always be reducible to some (more) UV-complete higher-energy theory, be it one that is already established or one that has yet to be discovered. Only theories valid independent of scale are considered fundamental.

One possible way to justify the use of EFTs is to argue that they are constructed with specific extra-theoretical virtues in mind. In cases where there are fundamental theories available, we can easily justify the construction of incomplete EFTs for the achievement of specific goals, such as the derivation of values to be tested experimentally (similar to the construction of models, see e.g. Gelfert 2016). In the case of bottom-up EFTs, however, where no fundamental theories are available, we might have to judge differently. In practice, EFTs play an important role in physics—even the acclaimed Standard Model of Particle Physics is now thought of as a bottom-up effective theory (see e.g. Bechtle et al. 2022). In times where experimental data fail to indicate the direction towards a theory of everything and an 'infinite tower of effective theories' at the foundation of fundamental physics becomes more and more accepted, the question arises whether effective theories can meet the standards that physical theories have been held up to in the past, and whether we can rely on them in the same way when it comes to questions of, for example, understanding.

4 Understanding effectively with EFTs

EFTs, we have just seen, are inherently incomplete theories that provide an idealized and approximate description of phenomena. This raises doubts about their status and reliability within the (philosophy of) high energy physics. At the same time, we saw in section 2 that idealizations and approximations can contribute productively to scientific understanding. In this section, I will show that the framework developed there can be applied to EFTs. They are accurate representations of the target phenomena they describe (section 4.1). The cutoff Λ as an idealization and the renormalization as an approximation contribute productively to the understanding so obtained (section 4.2). This allows the conclusion that scientific understanding does not depend on the degree of fundamentality ascribed to a theory. On the contrary, I will argue that effective theories might even be better vehicles for understanding than (more) fundamental ones. At least where understanding is concerned, the absence of a theory of everything gives us no reason to worry (section 4.3). The section concludes with the discussion of an example effective theory, the Fermi theory (section 4.4).

4.1 EFTs as representative vehicles

Before we turn to the specific role of the 'effective' elements of EFTs, we first need to establish how an EFT represents a target phenomenon, and ask if and how this representation is accurate.

An effective theory can be considered a model that is representative of the target phenomenon that we want to investigate and for which we construct the initial Lagrangian. This EFT denotes the target, or equivalently, its representational force points towards the target. It also allows competent and informed agents to make inferences about it. This should not come by surprise; after all, an EFT is constructed with this specific purpose in mind. It contains the key aspects of the phenomenon of interest, such as the interaction studied, the symmetries and mechanisms one expects to be involved, and the energy scale at which the process transpires. From the considerations made in section 3.2, it follows not only that an EFT is a representation of the phenomenon but also that it is a distortion, a misrepresentation. High-energy contributions are ignored, and the initial Lagrangian is renormalized, distorting some features we know or expect to be present in the phenomenon.

Despite these distortions, EFT models can be considered accurate: the inferences (such as predictions) that are made are true. While conclusions based on faithful representations of target features should not lead any competent and informed agents to inaccurate conclusions, distortions might well prompt falsehoods if not interpreted correctly.

Let us thus begin with the interpretation of the cutoff Λ, which was previously identified as an idealization, a distortion of the model-target relationship. In section 2.4, I proposed to interpret idealizations autonomously, and consider also their intrinsic truth value independently of any relation to a target. When we know how to read an idealization truthfully, we can consider it to contribute to true target inferences, maintaining the accuracy of the model in the process. For the cutoff that sits at the heart of an EFT, this means the following. Λ contains important information about the phenomenon. It is a telling instance not only in the negative sense, of the thing it actively distorts and thus deems irrelevant—the contributions above the cutoff Λ—but also in the positive sense, of the forces or particles involved and the energy scales relevant to the interaction that is modelled. Considered independently and taken in their respective contexts, these are true facts that can be related to the phenomenon. The truth of the intrinsic content does not stand in conflict with the fact that at the same time they are extrinsically false distortions of elements of the original target system. Λ has been consciously introduced at a specific energy scale by the physicist constructing the EFT. We can go even further and consider Λ, despite being an idealization, to represent a difference maker in the phenomenon: it indicates the scale after which we know we need not consider further contributions. Employing a non-literal interpretation allows us to learn these truths from the idealization Λ.

After the distortion of the model-target relationship Λ is introduced, we continue by renormalizing the initial Lagrangian. While the consequences of the idealization Λ can be clearly identified and located within the model, this does not hold for the approximation that the application of a renormalization scheme introduces. The renormalization constitutes a model-internal transformation that is due to values or goals that are internal to the modeling endeavor. After we have obtained the final EFT, we cannot disentangle conclusions based on distorted or undistorted elements. It is thus not sensible for a transformation that is independent of the represented target to require a split into faithfully and unfaithfully represented target elements, and this is indeed what happens when a renormalization scheme is applied. The EFT model is, in the words of Rice, holistically distorted (see Rice 2019). However, in order to obtain this holistically distorted model, the cutoff idealization needs to be introduced, which can nonetheless be singled out. Λ thus has a primary positive and independent contribution in idealizing the theory, as well as a secondary role in enabling the approximation that follows. We therefore arrive at a model that is holistically distorted, yet where we can identify specific elements as explicitly positively contributing to the model use. It is not only the final product that is interesting, but the assumptions needed to get there, too.

Conclusions made from the EFT that result not only from an idealization but also from an approximation need to be checked for their accuracy, too. Not only do we want to learn from Λ itself, but also from the EFT as a whole. As the cutoff also figures prominently in all renormalization schemes, we can employ a similar strategy for interpreting renormalization: non-literalism. We take into account that the renormalization, as an approximation, is a distortion of the model system, and that conclusions made need to be interpreted accordingly. Because approximations are model-internal and do not have a target-relative truth value in the sense that idealizations do, we do not need to justify their use relative to the target. Instead, we rely on their appropriateness being settled independently. Competent and informed agents, it can be assumed, should generally possess this background knowledge. Employing an appropriate interpretation, the intrinsic and true content of both idealizations and approximations can be interpreted in such a way that it can be truthfully related the target. We thus conclude that an effective theory can accurately represent its target phenomenon.

4.2 The productiveness of Λ and renormalization

It has already been shown that we can interpret both the cutoff and the renormalization scheme, the two distinct features of an effective theory, in an intrinsically truthful way. Building on the conclusions from section 2.5, in this section I will explore their contributions to scientific understanding specifically.

In general terms, the whole process of constructing an effective theory for some phenomenon, starting with an initial Lagrangian followed by the introduction of a cutoff idealization and renormalization approximation, provides us with understanding of that phenomenon. More specifically, Λ provides us with information relevant to the target phenomenon that is both true and not evident in a completely faithful representation. This information about the target phenomenon improves our understanding, adding to its content with information about the relevant energy scales, symmetries or particles involved. It also carries information about expected or predicted behavior both below and above the cutoff. With this, it makes the theory that describes the phenomenon more intelligible. Λ also makes the application of a renormalization scheme possible, which provides us with a theory that is comprehensible, mathematically useful, and generates numbers that can be compared to experiment. This makes the theory intelligible in de Regt's sense: We can easily[10] see the EFT's consequences and make comparisons.

We use the renormalized EFT and the cutoff idealizations as a vehicle for target understanding, where we do not admit these idealizations as literal statements into the content of our understanding, but in a non-literal, context-dependent way. We do not interpret the cutoff as 'there are no contributions to our target interaction at energies above Λ', rather, we admit to the content of our understanding statements such as if 'on the energy scale of the target phenomenon, interactions stemming from particles whose masses correspond to energies above Λ contribute negligibly at the desired accuracy'. Similarly, statements about the relation between the behavior of real and hypothetical systems where there were no contributions above Λ, can contribute to scientific understanding. In the end, we obtain factual statements that take into account this context of the modeling. They populate the content of understanding.

Both the cutoff Λ and the renormalization scheme can therefore be called productive elements of the model. They contribute crucially and positively to the understanding provided by an EFT. They do so autonomously, independent of the target and of each other. Even though the renormalization scheme needs a cutoff in order to be applied, it contributes to understanding by allowing (successful) mathematical treatment, and does so even in the case of strong renormalizability where all dependencies on Λ can be removed. It also does so entirely model-internally, without reference to the target system. Hence, its contribution to understanding does not hinge on Λ specifically, which, on the other hand, contributes to understanding by its relation to the target phenomenon. Its contributions can be singled out and made explicit. We thus have an EFT that, at the same time, is a holistic distortion and contains features that are explicitly, autonomously, and directly productive for scientific understanding.

4.3 The (ir)relevance of fundamentality and (more) fundamental theories

It can be concluded that effective theories can provide understanding, and that they do so in virtue of, and not despite, their 'effective', i.e. idealized and approximative, nature. EFTs are intelligible theories that appropriately describe the phenomenon.

The renormalization scheme allows a successful mathematical treatment of the phenomenon, providing values that can be compared to experiment, and significantly increasing the theory's intelligibility by considering a smaller amount of degrees of freedom than would be needed for a theory applicable also at larger energy scales. The cutoff indicates to us those features that are, or are not, especially important to the phenomenon studied. With this information it

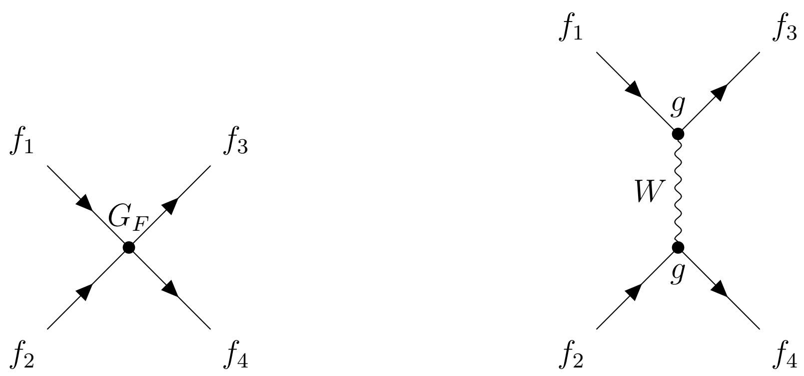
[10] Easy should be understood here relative to the discipline of HEP, which, after all, his highly technical and requires advanced knowledge.

increases the understanding that any theory of that phenomenon can provide. The incomplete nature of EFTs makes the whole description more intelligible, or allows an intelligible description in a first place. The presence of the cutoff thus increases our understanding. Fundamentality or UV-completeness, on the other hand, do not. All theories are judged on their intelligibility, independently of whether or not they contain a cutoff or a renormalization scheme was used in their construction. Not only does the understanding provided not depend on—or even make reference to—the UV-completeness or fundamentality of the theory that is used, an effective theory most likely does not provide 'worse' understanding. On the contrary, because EFTs can be easier to compute and more intelligible, they might have an edge on theories that are (more) UV complete but simultaneously unwieldy and impractical. Not only do the cutoff and renormalization scheme contribute productively to understanding, they provide to the understanding information that is hard to attain otherwise. Furthermore, there is no difference in the procedure or the quality of understanding obtained with EFTs that start with an initial Lagrangian constructed in the top-down or bottom-up way. No EFT, when used as an understanding-providing vehicle, refers to a (more) UV-complete theory. They provide understanding autonomously.

As we are interested in understanding the phenomenon, and not the EFT itself (or any other theory for that matter), which type of theory we would prefer based on other values is not relevant. Instead, all theories are judged on the same basis. We therefore conclude: scientific understanding is independent of fundamentality. Consequently, we do not need to be worried about the lack of evidence for proposals that lead towards a unified picture of physics or a theory of everything. Regarding scientific understanding, EFTs can be taken as seriously as any other type of theory.

4.4 Understanding beta decay with Fermi theory

To see how EFTs provide understanding in practice, we will look at Fermi theory as an example. This theory was proposed in the early 1930s by Enrico Fermi to explain an observed continuous energy spectrum in beta decay (Fermi 1934). The theoretical framework now available to describe weak interactions that partake in beta decay had not been developed yet, so Fermi set out to make sense of the observed phenomenon based on a few experimental constraints and theoretical assumptions, including the model of quantum electrodynamics, the neutrino hypothesis that had recently been proposed by Wolfgang Pauli, and assumptions regarding the nature of the interaction's Hamiltonian. He also took into account the energy regime at which the interaction was observed experimentally, and what was known about interactions in this energy regime. This allowed Fermi to construct a bottom-up effective theory that describes the observed energy spectrum via a direct interaction of four fermions with its strength given by the Fermi constant G_F, as visualized in figure 3a (Hartmann 2001, pp. 8, 22). This theory was able to match the experimentally obtained values with high accuracy, and, up to scales of about 100 GeV describes the weak interaction quite well, making the theory an important success of



(a) Four-fermion interaction as described by Fermi theory.

(b) Four-fermion interaction as described by Weinberg–Salam theory.

Figure 3: Fermi theory is an EFT that describes the interaction of four fermions for energies lower than those corresponding to the mass of the W-boson, and which can be considered a low-energy effective theory deducible from Weinberg–Salam theory.

early elementary particle physics.

Today, we can derive the same EFT also in a top-down manner, starting from Weinberg–Salam theory for electroweak interactions. Introducing a cutoff just below the mass of the W-boson turns the known Standard Model description of the weak interaction as being mediated by a W-boson (as seen in figure 3b) into the effective description given by Fermi theory. For energies $E \leq m_W$, it provides a framework that is both conceptually and mathematically simpler yet accurate enough for a range of purposes.

Renormalization theory needs to be applied to the initial Lagrangian on which Fermi theory is based, in the manner discussed in section 3.1. This is independent of whether the Lagrangian was constructed top-down or bottom-up. A cutoff Λ is introduced at an energy consistent with the assumptions about the interaction; today, we know that this energy lies just below the mass of the W-boson $\Lambda < m_W$. This is an idealization, as it is a distortion of the full energy spectrum that, in principle, could be considered for any particle interaction. In the Fermi theory case, the application of the perturbative or Wilsonian renormalization schemes fails, but Fermi theory is pragmatically renormalizable using Wilsonian perturbative renormalization up to some characteristic energy E_C below the cutoff Λ.

Fermi theory provides us with straightforward examples of the way the cutoff, as an idealization, and the renormalization, as an approximation, can contribute to understanding a phenomenon with an EFT. Λ tells us about relevant features of the interaction and at which energy the theory falls apart. Internally, these are true pieces of information that can be used to understand the phenomenon that the theory describes, i.e. those appearing at energy scales below the cutoff. Λ is also needed for renormalization, which provides us with a computationally useful theory and values that can be compared to experiment. The immediate four-particle interaction is intelligible and we can see its consequences immediately, using, for example, the insights provided by diagram 3a. All this makes Fermi theory a suitable vehicle to understand weakly interacting phenomena at low energies. Both the cutoff and the renormalization scheme contribute productively to this understanding.

While the original bottom-up construction of Fermi theory and the confirmation of its 'correctness' by top-down derivation from Weinberg–Salam theory are of historical interest, the motivation for and the construction of an initial Lagrangian play only a limited role in our discussion of Fermi theory in the context of scientific understanding. Fermi theory provided understanding of beta decay and other weak interactions before the introduction of a more fundamental and more UV-complete theory, and it still provides that same understanding today. It does so without reference to another theory but based on its own intelligibility, assisted by the idealizing and approximating assumptions. Still, sceptics might ask whether understanding provided by Fermi theory does not conflict with understanding provided by Weinberg-Salam theory. After all, they describe the same process differently. Here, it is important to emphasize the context-dependence of understanding. Both theories provide understanding relative to their own context, including, for example, the relevant energy scale. They are both true independently of one another. Understanding can come in depths, and it is reasonable to assume that these 'pieces' of understanding complement each other.

Fermi theory serves as an example of any effective theory which, due to its inherent cutoff, is considered incomplete. Other effective theories, such as QED or even the Standard Model of particle physics itself, can be subjected to an analogous analysis. In all cases, the same conclusion holds true: EFTs are intelligible theories, the understanding they provide is increased by the cutoff Λ and the renormalization scheme, and independent of the theory's fundamentality.

5 Conclusion

In this article, I argued that not only approximations, which are introduced into a model to reach goals independent of target reference, but idealizations, too, have an intrinsic truth value. The content of an idealization, understood at the scale and in the context in which it is set, can be true while, at the same time, that idealization is a distortion of the target and thus considered a falsehood. Idealizations do not have to be understood only as pointers towards irrelevant target features, they also contribute to scientific understanding through the true information that they carry. They do so entirely on their own, and can thus be considered productive elements of the model.

Idealizations, just as approximations, can contribute to scientific understanding in particular because, contrary to scientific knowledge or explanation, understanding is a subjective and context-relative cognitive achievement of an agent. It also is relative to the modeling process and the question the modeler sets out to answer. With this context in mind, the appropriate, truthful interpretation of a target distortion can be found and contribute to understanding of that target. Accurate representation, many philosophers agree, is about the ability to make true inferences with one system about another, rather than (structural) similarities between the two. True understanding, consequentially, should be based also on the model-to-target inferences rather than exclusively on the target-to-model relationship and its distortions. Emphasizing the truth of the internal information an idealization carries, and thus allowing it to productively contribute to truthful understanding, does exactly that.

Effective field theories, being inherently incomplete, are a textbook example of an idealized model. Prompted by the increasing pressure on unification as a guiding principle in high energy physics and the continuing presence of effective theories, the question arose whether EFTs—or only fundamental theories—can provide scientific understanding. And if such inherently incomplete theories can in fact provide understanding, what is the role of their incomplete nature: does it contribute to, or hinder understanding? I conclude that the distortions present in EFTs—the cutoff Λ as an idealization and the renormalization scheme as an approximation—do not hinder the understanding they can provide. On the contrary, the cutoff carries information about the scales, forces, or symmetries relevant to the interaction studied, and renormalization allows the calculation of values that can be compared to experiment. Both features increase the understanding of the target phenomenon, and they do so without reference to an undistorted UV-complete theory. Theories can be understanding-providing vehicles independently of the degree of idealization or, in an EFT context, the degree of fundamentality we grant them, or whether they can be reduced to some other, more fundamental theory. Hence, the lack of evidence for a theory of everything does not need to worry us: it does not diminish our understanding of the physical world. On the contrary, the scale-dependent information that EFTs provide makes them rather favorable tools for understanding. Understanding can be actively improved by idealizations that explicitly facilitate true inferences about the target.

References

Appelquist, Thomas and J. Carazzone (1975). "Infrared Singularities and Massive Fields". In: *Physical Review D* 11.10, pp. 2856–2861. DOI: 10.1103/PhysRevD.11.2856.

Batterman, Robert W. and Collin C. Rice (2014). "Minimal Model Explanations". In: *Philosophy of Science* 81.3, pp. 349–376. DOI: 10.1086/676677.

Bechtle, Philip et al. (2022). "Bottoms up: The Standard Model Effective Field Theory from a Model Perspective". In: *Studies in History and Philosophy of Science* 92, pp. 129–143. DOI: 10.1016/j.shpsa.2022.01.014.

Bihan, Soazig Le (2016). "Enlightening Falsehoods: A Modal View of Scientific Understanding". In: *Explaining Understanding*. Routledge.

de Regt, Henk W. (2015). “Scientific Understanding: Truth or Dare?” In: *Synthese* 192.12, pp. 3781–3797. DOI: 10.1007/s11229-014-0538-7.

de Regt, Henk W. (2017). *Understanding Scientific Understanding*. Oxford University Press. DOI: 10.1093/oso/9780190652913.001.0001.

Duncan, Anthony (2017). *The Conceptual Framework of Quantum Field Theory*. Oxford, New York: Oxford University Press.

Elgin, Catherine Z. (2007). “Understanding and the Facts”. In: *Philosophical Studies* 132, pp. 33–42. DOI: 10.1007/s11098-006-9054-z.

Elgin, Catherine Z. (2010). “Telling Instances”. In: *Beyond Mimesis and Convention: Representation in Art and Science*. Ed. by Roman Frigg and Matthew Hunter. Boston Studies in the Philosophy of Science. Dordrecht: Springer Netherlands, pp. 1–17. DOI: 10.1007/978-90-481-3851-7_1.

Elgin, Catherine Z. (2017). *True Enough*. 1st edition. Cambridge, MA: The MIT Press.

Fermi, E. (1934). “Versuch einer Theorie der β-Strahlen. I”. In: *Zeitschrift für Physik* 88.3, pp. 161–177. DOI: 10.1007/BF01351864.

Fletcher, Samuel C. (2019). “Minimal Approximations and Norton’s Dome”. In: *Synthese* 196.5, pp. 1749–1760. DOI: 10.1007/s11229-018-1676-0.

Franklin, Alexander (2018a). “On the Renormalization Group Explanation of Universality”. In: *Philosophy of Science* 85.2, pp. 225–248. JSTOR: 26551920.

Franklin, Alexander (2018b). “Whence the Effectiveness of Effective Field Theories?” In: *British Journal for the Philosophy of Science* 71.4, pp. 1235–1259. DOI: 10.1093/bjps/axy050.

Fraser, James D. (2020). “The Real Problem with Perturbative Quantum Field Theory”. In: *The British Journal for the Philosophy of Science* 71.2, pp. 391–413. DOI: 10.1093/bjps/axx042.

Frigg, Roman (2022). *Models and Theories: A Philosophical Inquiry*. London: Routledge. DOI: 10.4324/9781003285106.

Frigg, Roman (2023). “Understanding Scientific Representation”. In: *Linkage: Studies in Applied Philosophy of Science* 3, pp. 1–6. DOI: 10.50824/linkage.3.0_1.

Frigg, Roman and James Nguyen (2021a). “Mirrors without Warnings”. In: *Synthese* 198.3, pp. 2427–2447. DOI: 10.1007/s11229-019-02222-9.

Frigg, Roman and James Nguyen (2021b). “Scientific Representation”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2021. Metaphysics Research Lab, Stanford University.

Frigg, Roman and James Nguyen (2022). *Scientific Representation*. Elements in the Philosophy of Science. Cambridge: Cambridge University Press. DOI: 10.1017/9781009003575.

Frigg, Roman and James Nguyen (2025). “Stabilising Understanding”. In: *Philosophical Studies*. DOI: 10.1007/s11098-025-02384-y.

Gelfert, Axel (2016). *How to Do Science with Models*. SpringerBriefs in Philosophy. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-27954-1.

Georgi, H. (1993). "Effective Field Theory". In: *Annual Review of Nuclear and Particle Science* 43.Volume 43, 1993, pp. 209–252. DOI: 10.1146/annurev.ns.43.120193.001233.

Hartmann, Stephan (2001). "Effective Field Theories, Reductionism and Scientific Explanation". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32.2, pp. 267–304. DOI: 10.1016/s1355-2198(01)00005-3.

Hughes, R. I. G. (1997). "Models and Representation". In: *Philosophy of Science* 64, S325–S336.

Koberinski, Adam (2024). "The Breakdown of Effective Field Theory in Particle Physics: Lessons for Understanding Intertheoretic Relations". In: *Philosophy of Science*, pp. 1–21. DOI: 10.1017/psa.2024.30.

Koberinski, Adam and Doreen Fraser (2023). "Renormalization Group Methods and the Epistemology of Effective Field Theories". In: *Studies in History and Philosophy of Science* 98, pp. 14–28. DOI: 10.1016/j.shpsa.2023.01.003.

Lawler, Insa (2021). "Scientific Understanding and Felicitous Legitimate Falsehoods". In: *Synthese* 198.7, pp. 6859–6887. DOI: 10.1007/s11229-019-02495-0.

Lipton, Peter (2009). "Understanding Without Explanation". In: *Scientific Understanding: Philosophical Perspectives*. Ed. by H. W. de Regt, S. Leonelli, and K. Eigner. University of Pittsburgh Press, pp. 43–63.

Manohar, Aneesh V. (1996). "Effective Field Theories". In: pp. 311–362. DOI: 10.1007/BFb0104294. arXiv: hep-ph/9606222.

McComb, William David (2007). *Renormalization Methods: A Guide For Beginners*. Oxford, New York: Oxford University Press.

Norton, John D. (2012). "Approximation and Idealization: Why the Difference Matters*". In: *Philosophy of Science* 79.2, pp. 207–232. DOI: 10.1086/664746. JSTOR: 10.1086/664746.

Palacios, Patricia (2022). *Emergence and Reduction in Physics*. Elements in the Philosophy of Physics. Cambridge: Cambridge University Press. DOI: 10.1017/9781108901017.

Peskin, Michel E. and Daniel V. Schroeder (1995). *An Introduction to Quantum Field Theory*. Westview Press.

Potochnik, Angela (2020a). "Idealization and Many Aims". In: *Philosophy of Science* 87.5, pp. 933–943. DOI: 10.1086/710622.

Potochnik, Angela (2020b). *Idealization and the Aims of Science*. Chicago, IL: University of Chicago Press.

Redhead, Michael (1988). "A Philosopher Looks at Quantum Field Theory". In: *Philosophical Foundations of Quantum Field Theory*. Ed. by Harvey R. Brown and Rom Harré. Clarendon Press, pp. 9–24.

Reutlinger, Alexander, Dominik Hangleiter, and Stephan Hartmann (2018). "Understanding (with) Toy Models". In: *British Journal for the Philosophy of Science* 69.4, pp. 1069–1099. DOI: 10.1093/bjps/axx005.

Rice, Collin (2025). “Which Possibilities Contribute to Scientific Understanding?” In: *Modeling the Possible*. Ed. by Tarja Knuuttila et al. Routledge.

Rice, Collin C. (2016). “Factive Scientific Understanding without Accurate Representation”. In: *Biology & Philosophy* 31.1, pp. 81–102. DOI: 10.1007/s10539-015-9510-2.

Rice, Collin C. (2019). “Models Don’t Decompose That Way: A Holistic View of Idealized Models”. In: *The British Journal for the Philosophy of Science* 70.1, pp. 179–208. DOI: 10.1093/bjps/axx045.

Rivat, Sébastien and Alexei Grinbaum (2020). “Philosophical Foundations of Effective Field Theories”. In: *The European Physical Journal A* 56.3, p. 90. DOI: 10.1140/epja/s10050-020-00089-w.

Shech, Elay (2023). *Idealizations in Physics*. Elements in the Philosophy of Physics. Cambridge: Cambridge University Press. DOI: 10.1017/9781108946742.

Strevens, Michael (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.

Suárez, Mauricio (2004). “An Inferential Conception of Scientific Representation”. In: *Philosophy of Science* 71.5, pp. 767–779. DOI: 10.1086/421415.

Suárez, Mauricio (2024). *Inference and Representation: A Study in Modeling Science*. Chicago, IL: University of Chicago Press.

Verreault-Julien, Philippe (2019). “Understanding Does Not Depend on (Causal) Explanation”. In: *European Journal for Philosophy of Science* 9.2, p. 18. DOI: 10.1007/s13194-018-0240-6.

Weingarten, Karla (forthcoming). “Understanding Particle Interactions: Feynman Diagrams as Representative Models”. In: *The British Journal for the Philosophy of Science*. DOI: 10.1086/737746.

Weisberg, Michael (2007). “Three Kinds of Idealization”. In: *The Journal of Philosophy* 104.12, pp. 639–659. JSTOR: 20620065.

Williams, Porter (2019). “Scientific Realism Made Effective”. In: *The British Journal for the Philosophy of Science* 70.1, pp. 209–237. DOI: 10.1093/bjps/axx043.

Wilson, Kenneth G. (1975). “The Renormalization Group: Critical Phenomena and the Kondo Problem”. In: *Reviews of Modern Physics* 47.4, pp. 773–840. DOI: 10.1103/RevModPhys.47.773.