Empirical Techniques and the Accuracy of Scientific Representations

Dana Matthiessen

ABSTRACT: This paper proposes an account of accurate scientific representation in terms of techniques that produce data from a target phenomenon. I consider an approach to accurate representation that abstracts from such epistemic factors, justified by a thesis I call Ontic Priority. This holds that criteria for representational accuracy depend on a pre-established account of the nature of the relation between a model and its target phenomenon. I challenge Ontic Priority, drawing on the observation that many working scientists do not have access to information allowing them to describe this relation between model and target or compare them accordingly. I critique the ability of an ontic-first approach to provide accuracy criteria in such cases and present historical support for an alternative according to which a model is accurate if and only if integrating the model into a theory of the data acquisition process yields well-fitting predictions of patterns in the data.

Keywords: Scientific Representation; Accuracy; Models; Experimentation; Data-Phenomenon Inference

1 Explaining the accuracy of scientific representations

A central explanatory project for the philosophy of scientific representation concerns the accuracy of scientific models. Representational accuracy is tied to the fact that competent users draw informative inferences about real-world target phenomena from models, some of which are more successful than others. Enabling informative inferences is identified by Suárez (2004) as a necessary condition for model-based representation in science. Informative inferences typically involve positive or negative attributions of states, properties, or behaviors to a target system in a way that is specific to this target and goes beyond knowledge explicitly assumed in a model’s design. On this usage, that an inference is informative with respect to a target does not entail that it is accurate; such attributions may lead to stark failures. Rather, this condition is meant to capture the fact that theoretical models are useful tools for investigative reasoning, and to distinguish scientific representation from other forms of representation established by stipulative fiat or arbitrary association. To be judged accurate, then, a model must satisfy some further criterion that accounts for the *success* of the inferences it enables in contexts of empirical application. Paradigmatic cases of success can be found throughout the history of science, often involving the prediction and subsequent uncovering of novel phenomena. This makes accuracy a worthy philosophical explanandum of its own.[1]

One line of approach found in the extant literature seeks to explain representational accuracy by means of a definitional account of scientific representation, understood as a relation between model and target phenomena. This approach seeks to first answer the question: "in virtue of *what* do scientific models represent their targets?" A common strategy of this kind specifies how distinct features of a model and a target phenomenon relate to one another—e.g., as elements in a formal similarity or morphism relation—and explains the ability to reason with models on this basis. If this question were settled, one could argue that models are accurate if and only if the proper representation relation in fact obtains between them and their targets. The motivation for this approach is based on a presupposition that I call *Ontic Priority*: the view that a theory of scientific representation must first possess a particular account of the *what*—the nature of scientific representation qua model-target relation—in order to explain the *how*—that is, the ability to use models to successfully reason about target phenomena. On this view, the activities of epistemic agents do not affect whether the proper relation between a model and its target obtains; a model may be accurate whether or not its accuracy is ever checked. Facts about empirical techniques and instrumentation, about how one *comes to know* whether a model affords inferential success, are separate concerns. This entails that the accuracy criteria of a model with respect to some target can be analyzed independently from criteria involved in evaluating the model, say, with respect to data generated from that target. Call this an ontic-first approach. A commitment to Ontic Priority thus involves two conjoined claims: first, that a set of ontic conditions are responsible for, and so explain, the

---

[1] Frigg and Nguyen (2017) refer to this desideratum as the issue of formulating standards of accuracy, which answer the question "how do we identify what constitutes an accurate representation" (p. 55).

outcome of epistemic evaluations of modeling accuracy; second, that an investigation of these conditions is therefore methodologically prior in any effort to establish accuracy criteria.

After providing some evidence of this approach, I will present a challenge to it. I begin from the premise that an adequate theory of scientific representation ought to provide us with accuracy criteria that scientists themselves can apply. My argument is based on the observation that, in many cases of scientific inquiry, researchers do not have prior access to a rich, model-independent characterization of a target phenomenon. Thus, many scientists do not have access to information that would allow them to describe the relation between model and target or straightforwardly compare one to the other. Yet they still draw informative inferences from models and evaluate their success in these situations. It is desirable to explain how working scientists do this from within their horizon. I assess the ability of an ontic-first approach to provide criteria that reflects this usage and find it wanting. This is because the evaluation of a model's accuracy depends on the reliable characterization of its target, which in turn depends on their possessing an account of the data acquisition process. In circumstances where scientists lack a model-independent account of their target, the evaluation of a model and the characterization of that model's target from data are inferentially intertwined. To support these claims, I draw on an extended episode from the early history of X-ray research. Here, the fact that researchers could successfully reason from model to target depended on their incorporating this model into an account of the data collection process. I propose a corresponding criterion on this basis, according to which a model is accurate if and only if integrating the model into a theory of the data acquisition process yields well-fitting predictions of patterns in the data.[2]

---

[2] This usage of accuracy builds on Mitchell (2020), where it is claimed modeling accuracy is evaluated with respect to data. Note that my presentation here is not meant to encompass all kinds of uses for theoretical models, such as investigations of their internal properties, relations to other models, or compatibility with some range of hypotheses. For the purpose of this paper, I take representational uses of models to be those aimed at explaining, predicting, or otherwise understanding the properties and behaviors of something external to the models themselves or their parent theories. These I refer to as their target phenomena. Phenomena may be characterized differently depending on the goals of the modeler; some target an extremely specific event or aspect of a system, others a highly general pattern or class of phenomena. My concern is primarily restricted to what is sometimes called target-directed modeling, as opposed to modeling aimed at very high levels of generality and sometimes called targetless modeling, how-possibly modeling, or purely exploratory modeling. Here I do not directly engage with these and related forms of modeling and representation discussed in the history and philosophy of science literature (as in Weisberg 2013 or Coopmans et al. 2014; see also critical remarks by Elliot-Graves 2020).

2 Ontic-first approaches in the philosophy of science

In a thoroughgoing review article on scientific representation,[3] James Nguyen writes, “We still do not have an account of in virtue of what scientific models represent their targets. This is particular worrying given that it is plausible we should answer this question before we investigate representational accuracy” (2016, p. 188). Nguyen cites Suárez (2004), Contessa (2007), and Frigg (2010) as authors who, despite their differences, “all take the question of representation as conceptually prior to accurate representation,” (p. 188). There is a sense of conceptual priority that is trivially satisfied here, insofar as the notion of accurate representation is a species of representation simpliciter. Less trivial is the philosophical methodology this recommends. Such priority for Nguyen appears to call for a specific order of investigation: first establish what grounds scientific representation, then seek criteria for accuracy on this basis.

The main concern of this paper is those who seek to characterize representation in terms of a fundamental, ontic relation between model and target.[4] For these authors, Nguyen’s worry might be expressed as follows: how can we judge that a model represents its target accurately, and so explain inferential success, without an account of the relation that the model stands in with respect to its target? This is precisely how Chakravartty defends “informational” accounts of representation, which appeal to model-target similarities:

> How, one might wonder, could such practices [of model-based interpretation and inference] be facilitated successfully, were it not for some sort of similarity between the representation and the thing it represents—is it a miracle? […] Indeed, it is precisely *because* the informational view is satisfied in this way that cognitive activities such as interpretations and inferences regarding target systems are successful in the first place (2010, pp. 201-203).

Bueno and French draw directly on this argument to support their partial morphism account of representation: “It is the requirement that such facilitation be accommodated that effectively demands that there be some relation of similarity that holds between the representational device and the target

---

[3] Cf. Frigg and Nguyen (2017).

[4] In other words, I am largely putting aside deflationary views, such as Suárez’s.

system" (2018, p. 65). Here inferential success is taken as a fact that, by a form of no-miracles argument, necessitates the existence of a particular kind of model-target relation. "[W]ithout it, the success of the very functions that functional [e.g., inferentialist] accounts take to be central would appear to be inexplicable" (Chakravarrty 2010, p. 201). In other words, one must specify the nature of the model-target relation in order to explain successful model-based inference.

On the other hand, "scientists *discover* that their model is a misrepresentation (typically) through lack of empirical success," spurring the creation of improved models. A partial morphisms account will accordingly describe "the interrelationships between the increasingly successful models that, because of this success, can be regarded as increasingly better, more informative representations." (Bueno and French 2018, p. 69). Improvements in accuracy are accounted for in terms of progressive mappings from less to more successful models. Whatever modicum of accuracy is possessed by an earlier model in this sequence is explained by structural similarities to its target. These are imported into more successful offspring by structural similarities between the model and its successors, thereby preserving the kernel of model-target relation responsible for the accuracy of the prior model. Scientists' confirming a model's accuracy may spur the production of improved models, but this process supervenes on the structural relations that underwrite their successes and failures. For Bueno and French, the practice of confirming accuracy—say, by checking it against data—is a separate epistemic matter from the accuracy of the model itself.

Weisberg's *Simulation and Similarity* agrees on this last point: "I will not primarily be focusing on experimentation, data, or confirmation. Before one can develop a theory of confirmation for models, one needs to be clearer about the nature of the model-world relationship" (2013, p. 90). This relationship is given by Weisberg's weighted feature-matching account. On this view, similarities between a model and target are evaluated in terms of the degree of fit between model features and target features, with some points of overlap or divergence given more weight than others. As a whole, this view is meant to characterize "how scientists represent the world with models" and "how their representational goals and ideals shape the standards of fidelity that they apply" (p. 135).

In doing so, it "reflects judgments about the relationship of models to their targets that scientists can actually make, because it draws on resources that are cognitively available: feature sets and weighting functions" (p. 155).

Again, this approach seeks to derive criteria of accuracy ("standards of fidelity" in Weisberg's parlance) from an account of the model-world relationship in a way that is independent from—and deemed conceptually prior to—matters concerning experimentation, data, or confirmation. To understand the rationale for separating these topics, we might look to remarks by Frigg, who likewise argues against the notion that a model's relation to data plays a role in scientific representation. Frigg claims that data are restricted to "an evidential function […] for the fact that the model is a (more or less) faithful representation of what is happening in the world" (2010, p. 111). From this standpoint, representation is a matter of establishing a particular relation between model and target phenomenon, standards of accuracy are understood in terms of whether this relation in fact obtains, and the data generating process is simply a means of producing evidence for this fact.

This shows that the two claims associated with Ontic Priority can be found in the literature. Authors routinely seek to define a model-target relation that is independent from a particular class of epistemic activities, such as data collection, and derive an account of model accuracy therefrom. This relation is the principal factor that explains scientists' judgments of accuracy; the assessment of a model in light of data is thus relegated to a secondary function from the standpoint of philosophical methodology, serving only to confirm or disconfirm a presumptive model-world relation.[5] This is not to claim that authors completely disregard the role of epistemic and pragmatic factors in their accounts of representation. Agreement with practice is a widely shared goal in philosophy of science,

---

[5] It should be noted that Weisberg presents his view as one concerned with the epistemology of model, and that he sometimes wavers on this demotion of data collection. He notes that empirically collected data are "used to make inferences about the nature of targets and mathematical representations of targets" (p. 96) and so play a foundational role in characterizing a model's target. At another point he states that "Modeling practice often involves an interaction between the development of the model and the collection of empirical data" (p. 155). Still, there is a general neglect in this literature of both the process of target characterization (though see Elliot-Graves 2020) and the dynamics of the interaction that Weisberg references.

and many take care to note that aspects of the interpretation of models depend on “the research questions of interest, the context of research, and the community’s prior practice” (Weisberg 2013, p. 149) or on factors “having to do with the use to which we put the relevant models” (Bueno and French 2018, p. 66). However, these are often incorporated as addenda or specifications within a framework in which the order of philosophical inquiry begins with the nature of the model-world relation. That is, accuracy criteria are still derived from this analysis without reference to data-gathering practices.

I will argue that an adequate account of model accuracy—that is, the kind of account that can provide criteria for working scientists—demands that we place this concept in closer relation to data acquisition. Having provided evidence of the ontic-first approach as a distinct tendency within the representation literature, I now sketch the elements of an ontic-first approach that follows the general pattern discussed in this section: a basic relation is assumed between model and target, accuracy is defined in these terms, and epistemic factors are relegated to “checks” on accuracy in a process of model-target comparison. I then present an ordinary scenario in scientific practice that is not captured by this framework, one in which a detailed model-independent target characterization is lacking. This scenario calls attention to the role of the empirical data-gathering techniques in target characterization, in a way that provides an important constraint on scientists’ judgments of model accuracy.

3 Ontic-first approaches and scientific practice

To start, it is useful to distinguish between consideration of the products and the activity of science, or what van Fraassen (2008) calls science viewed from “from above” and “from within.”[6] The view from above is retrospective, considering science in terms of its achieved theories and models in abstraction from the historical practices from which they resulted. The view from within seeks to characterize science

---

[6] Earlier instances of this distinction are found, e.g., in Cartwright (1983), Longino (1990) and Pickering (1992).

from the standpoint of these practices themselves, restricting its arguments to epistemic resources available to working scientists. These views are logically compatible, though approaches to an explanatory project for scientific representation will differ depending on which view is emphasized. Over-emphasis in either direction may carry risks: singular focus on practice can leave one swimming in particularities, while an excessively elevated view may lead to obscurities or formulations of questionable relevance to science in practice. This presents a problem, as a basic desideratum for a theory of scientific representation is that it be adequate to the practical realities of science in the making.

For instance, at a high level of abstraction the relation between a model and its target may be described "from above" as a relation between model components and some aspects of a target phenomenon. One can then decompose a model into constituents (be they elements, states, properties, relations, or transition rules), and define representation as a specific relation between these model constituents and the focal aspects of a target. Candidates may include forms of similarity, structural, or denotative mapping.[7] One could argue that a model user is able to reason about and draw inferences with respect to focal aspects of a target phenomenon in virtue of this relation. Further, the relation could be

---

[7] Among those viewing models as fictions or abstracta, it is common to see appeal to similarity relations: "Representation of a real-world system involves two distinct relations, the *specification* of a model system and some relevant *similarity* between model system and the world itself […] Many of the special features of model-based science come from the role played by resemblance relations between model system and target" (Godfrey-Smith, 2006, p. 733); "One way, perhaps the most important way, but probably not the only way [that scientists use models to represent] is by exploiting *similarities* between a model and that aspect of the world it is being used to represent […] It is the existence of the specified similarities that makes possible the use of the model to represent the real system in this way" (Giere, 2004, pp. 747-748). Structuralists tend toward mathematical relations: "a mathematical scientific representation can be schematically summarized as claiming that the concrete system S stands in the structural relation M to the mathematical system S*" (Pincock 2011a, p. 28); "the inter-relationships between physical structures are most appropriately captured by partial isomorphisms, those between mathematical structures and both physical structures and other mathematical structures are best captured via partial homomorphisms" (Bueno, French, & Ladyman, 2012, p. 45). In yet another account, the representational function of models is accounted for by "the fact that the user interprets the model in terms of the system" (Contessa, 2007, p. 51) via a denotative mapping between objects and relations in the model and those in the target.

investigated through a *comparison* of the model and target so characterized, in which the truth of statements linking model and target features is checked.[8]

For such an approach, a model is accurate if and only if the specified representation relation obtains between model and target. Model users apply this criterion through comparison, checking whether the proposed relation between model and target features does in fact obtain. There are reasons to doubt, however, that this is a sensible criterion for scientific practice. Importantly, it assumes there is a well-described target phenomenon, decomposable into a set of relevant features, that model features are mapped onto or interpreted in terms of for the sake of comparison. The use of certain toy examples is an aid to understanding representation in these terms. One regularly encounters appeals in the literature to models of swinging pendulums, city maps, or fictional characters like Sherlock Holmes as exemplary vehicles for representation. In these cases, the targets of each model are macroscopic objects with properties recognizable through ordinary acts of perception and concepts common to everyday discourse. The task of comparing these properties to a model appears straightforward. But these examples are disanalogous to the multitude of scientific models of systems that elude ordinary perceptual capacities and common concepts: models of nuclei, ecosystems, proteins, galaxies, atomic lattices, neural networks, black holes, cognitive modules, and their kin. In such cases, everyday concepts do not provide an adequate tool for describing the target phenomenon, nor should we expect this from them. On the contrary, scientists' best prima facie means for characterizing phenomena are often based on the very models used to represent them.

---

[8] Some instances of the language of comparison: "In each case, what we seem to be doing is comparing properties associated with one system with properties associated with another […] Whichever way the details go, these comparisons are guided by mappings between properties of the kind exhibited by ordinary objects" (Godfrey-Smith, 2009, p. 105); "[T]he apparent comparison [of a real system] with a nonexistent object eventually comes down [to] the unproblematic comparison of properties, and the statement making this comparison is true iff the statement comparing the properties with each other is true" (Frigg, 2010a, pp. 263-264); "Modelers intend talk of similarity between a concrete system and a model as an abstract object to be understood as a comparison between the model and the properties—perfectly respectable abstract objects—instantiated by the concrete object being compared" (Teller, 2001, p. 399); "Taken together with the mathematical entity, these propositions [describing how model parts correspond to physical features of a target system] impose a vast array of conditions on the system. These conditions are the content of the model. […] we will say that a model is accurate with respect to aspect A of the system when its content concerning A is correct" (Pincock, 2011b, p. 22).

If we grant that this is the case in some areas of model-based inquiry, it raises a threat of circularity for ontic-first accounts of model-to-target comparison. If models are the primary means of characterizing a target, then to what are these models being compared? On what grounds can it be asserted that models are similar or homomorphic to their real-world targets when these models are themselves the primary means of characterizing their targets?[9] Lacking such grounds, practicing scientists have no clear way to apply the accuracy criteria proposed by an ontic-first approach. In light of this, two strategies for specifying the notion of comparison suggest themselves. Either,

(i) **Posit:** the target description is a hypothetical posit and the model-target relation is justified abductively (or not) through predictive success or failure, or

(ii) **Data-mediation:** the representation relation is checked by comparison to an empirical result, which either provides the best independent characterization of its target or is the basis for any further model-target comparison.[10]

I'll consider each in turn.

Bueno and Colyvan (2011) address a case where there is no prior theoretical description of a target's structure. "This is clearly a problem for the mapping account […] When there is no such structure, we might impose some suitable structure or other and let the resulting mathematical model help us fine tune or revise the starting structure" (p. 347). This is an appeal to **Posit**. The target description is simply assumed and the fact that a model then successfully predicts an empirical result (however it might do so) licenses the abductive inference that a specific relation exists between the model and its target. **Posit** works well with views like that of Chakravartty (2010), for whom successful model-based inference would be inexplicable if such a relation were not in place. But what kind of explanation is being offered

---

[9] Van Fraassen (2008, Part III) details how this conceptual problem has arisen repeatedly in 20th century philosophy. With respect to a Putnam-Lewis debate, he writes, "As long as we are not given an independent description of both the domain and range of an interpretation, we do not *have* any such interpretation, nor any way to identify one […] we can grasp an interpretation—i.e. function linking words to parts of THE WORLD—*only if we can identify and describe that function*. But we cannot do *that* unless we can independently describe THE WORLD" (pp. 234-235).

[10] By an empirical result, I refer to the output of a processing sequence by which the data initially recorded from some target-apparatus interaction is given an interpretation or structure. I use the term interchangeably with data structure and data pattern in this paper.

here? By assumption, no model-independent account of the target—the basis for comparison—is available. For a practicing scientist in this scenario, the principal work in evaluating the model is done by determining whether the model can or cannot be used to predict the relevant empirical results. The notion that a real-world target must be similar or homomorphic to a model on this basis is a post hoc reconstruction, rather than an explanation from within the horizon of the model user. Moreover, the claim that one or another form of model-target relation obtains is underdetermined by the facts on hand. In this case, a successful prediction alone is insufficient grounds for a model user to distinguish between one explanatory relation or another, or to determine which features are primarily responsible for the model's success.[11] Accounting for accuracy in these terms would appear to offer little guidance for scientific practice.

What, then, of **Data-mediation**? In a footnote, Chakravartty takes it "as given in discussing relations between scientific representations and 'the world'" that authors are referring to the relation between representational models and models of data (2010, p. 203). This provides a criterion for accuracy that is amenable to scientific practice: a model user can typically check whether the appropriate relation holds between their model and data. Most authors defending semantic views attest to the mediating role of data (Cf. Bueno 1997; French and Ladyman 1999; van Fraassen 2008).[12] Weisberg also notes that many mathematical models "are compared to mathematical representations of targets […] very similar to what Suppes called a *model of data*" (2013, p. 95). Yet at the same time, most authors take models to represent something *more* than mere data.[13] From this initial comparison between model and data, then, one might think scientists can infer the accuracy of the model qua model-target relation. But here too there are difficulties, as there exist a diversity of techniques for comparing models to empirical results. It

---

[11] A version of this point is found in Rice (2019). For a historical example, see Kaveh (forthcoming), who argues that the successes and failures of Bohr's atomic model cannot be explained in terms of a semantic/referential relation between a model component and target.

[12] This goes back at least to Suppes (1962).

[13] Frigg (2010) critiques van Fraassen's view on this point. Weisberg claims that "Targets and their mathematical representations are not the same as empirically collected data" (2013, p. 96). Nguyen writes, "If it is phenomena that are ultimately represented by our scientific models, then it is those that are accurately or inaccurately represented." (2016, p. 175).

is not clear these are reducible to a single form of relation, such as similarity or homomorphism. Nor is it clear that the existence of such a relation between model and data entails further claims about the relation between a model and its target phenomenon. I'll consider an example of each relation holding between model results and a data pattern. In each case, this relation between model and data does not warrant inference to a specific relation between model and target.

Goodness of fit and similarity

Consider the likelihood ratio, a common measure of model-data fit used in fields ranging from particle physics to social psychology. This is a statistical measure designed to compare results derived from experimental data and expectations derived from models. It relates observed measurement values to a distribution of expected values conditional on a given modeling hypothesis H_0. We can write this as (L_{H0}/L_{best}), where H_0 is the hypothesis under test.[14] This ratio compares the goodness of fit of recorded data to two statistical models. One, L_{H0}, fits the data with a model based on hypothesis H_0; the other uses the same base model but parameterized to maximize fit. So the full ratio measures how well H_0 can account for data, returning values around 1 when the fit is optimal and 0 when it fails completely.

This provides a simple comparison between a model and data, and there is indeed a sense in which this measure informs us about similarities, but these are only similarities between two distributions, which are not necessarily informative about the relation between model and target. One could, for example, use a simple quantum mechanical model to simulate the frequency, over a large number of trials, with which a spin-½ particle would be found in an "up" state when acted on by a rotation operator restricted to a single axis. Using this model for H_0, the likelihood test would show the expected distribution of states works well to predict the number of "heads" observed in an equivalent number of fair coin tosses. In a very precise and limited sense, the model system and the coin tosses behave

---

[14] More precisely, it is given by $-2\ln(L_{H0}/L_{\text{best}})$. The nature of the statistical models represented by the Ls may vary widely across disciplines. In particle physics, for example, each model is often given by a product of Poisson distributions of the form $P(x|y) = (y^x e^{-y}/x!)$, where the variables are counts of events inferred from recorded data or generated using a theoretical model.

similarly; in the long run both generate similar distributions. But this statistical similarity should not warrant an inference to a deeper similarity between the components of a quantum mechanical model and a fair coin. Given what we know about the model and target in this example, we should resist the claim that anything beyond the recorded frequency of ‘H’s and ‘T’s has been accounted for by this comparison; no deeper aspect of a fair coin’s make-up has been revealed that is *really like* an observable acting on a mixed spin state, and there is reason to object to this inference: one problem with claiming that a standard coin toss has a similar structure to a quantum spin model is that the model must be arbitrarily restricted to generate the right distribution. Excess structure in the model, corresponding to other dimensions in which its states can be rotated, has empirical content that was left out with no principled justification other than to force a “match” with the statistics of a coin toss. Thus, a similarity between a model derivation and an empirical result, as determined by goodness of fit, is insufficient grounds for claiming that the model is accurate with respect to its target, or for explaining success on this basis.

Test responses and homomorphism

The same issues arise when considering a relation like homomorphism. Taking an example from Borsboom (2005), we could construct a four-item test for agreement with the following statements:

1) I have biological parents (yes: 1; no: 0)
2) I am over the age of seven (yes: 1; no: 0)
3) I menstruate (yes: 1; no: 0)
4) I have recently given birth (yes: 1; no: 0)

The test responses from a sufficiently large group of people would result in a triangular data structure, where all subject responses appear as one of the following lines:

Item 1	Item 2	Item 3	Item 4	Sum score
1	0	0	0	1
1	1	0	0	2
1	1	1	0	3
1	1	1	1	4

This is the same form of data predicted by a Guttman model, by which the strength of a single latent psychological variable (say, distrust of government) is measurable in terms of a series of threshold prompts. The observed data is homomorphic to the measurement scale resulting from a Guttman model. Still, there is no compelling reason to think this data measures a latent variable underlying test subjects' answers. This nesting is not sufficient to think the progression from people with sum score 1 (humans under seven) to those with score 2 (males and non-menstruating females over seven), and so on to new female parents corresponds to a gradually increasing measurand within these populations. The nature of the questions simply forces the triangular nesting structure into the data. A structural relation such as homomorphism between a model-based prediction and an empirical result is thus insufficient by itself to ground judgments about the accuracy of this model with respect to the target generating this data. As Borsboom summarizes, "What matters is not just the structure of the data, but also the question of how these data originated. The relations in the data must not only exist, they must exist because of the right reasons" (2005, p. 105).

To summarize: ontic-first approaches produce criteria for the accuracy of models that are not directly applicable to many cases of scientific practice. They claim that a model is accurate if and only if a specific relation obtains between model and target. It is suggested that this criterion is applied through a comparison of model and target features, where this is understood to verify whether the appropriate relation between model and target obtains. Prima facie, this fails to account for typical cases where a characterization of a target's features is only given by the model under consideration. In such cases, a model's ability to successfully predict an empirical result does not by itself warrant the conclusion that a specific relation obtains between the model and its target. This remains true even if the appropriate relation exists between a model and a data structure. One may still posit that the model successfully predicts results in virtue of the proper relation existing between model and target, though this is best viewed as a metaphysical reconstruction that does not reflect the standpoint of scientific practice. If a theory of scientific representation is meant to provide accuracy criteria that scientists themselves can apply, then we must look elsewhere.

4 A role for epistemic factors

Something must be added to **Data-mediation** if this view is to provide adequate criteria for judging a model to be an accurate representation. As Borsboom puts it, the patterns in the data must be there for the "right reasons." Representational accuracy thus depends on the nature of the empirical result to which a model is related. More precisely, the result must be able to serve as a reliable basis for characterizing the target phenomenon. This requires the data be generated and processed through a set of procedures designed for this purpose.[15] If judgment of the accuracy of a model with respect to a target depends on both the relationship of the model to an empirical result *and* facts about the provenance of that data, then an ontic-first approach to accuracy criteria will not suffice. In the following sections I will lay out criteria that are adequate to scientific practice.

The relation between a model and an empirical result, such as a data pattern, is only grounds for judging the accuracy of a model when data is reliable.[16] The activities that underlie reliability consist in the technical procedures of preparing, executing, and interpreting results produced by experimental techniques. Reliability is ultimately a matter of the consistency and fidelity of the outputs of experimentation and measurement practice. It requires that data only vary within acceptable limits and cohere with available understanding of the techniques employed in their production. Securing the consistency and coherence of data is typically a matter of addressing statistical and systematic uncertainties. The former result from variations due to innate imprecisions of the instruments in use and,

---

[15] See also Van Fraassen (2008): "There is nothing in an abstract structure itself that can determine that it is the relevant data model, to be matched by the theory. A particular data model is relevant because it was constructed on the basis of results gathered in a certain way, selected by specific criteria of relevance, on certain occasions, in a practical experiment or observational setting, designed for that purpose" (p. 253).

[16] Though this topic is not routinely considered in the scientific representation literature, it is a common subject in philosophy of science elsewhere. Bogen and Woodward (1988) provide a much-cited analysis of the reliability of data (see also Woodward 2000). Related discussion is found in the epistemology of measurement (Cf. Chang 1995; Tal 2013) and the work of Deborah Mayo (1996; 2010). Among authors writing on modeling and representation, van Fraassen (2008; 2012) and Peschard (2012) stand out in considering this matter in detail (see also the edited volume by van Fraassen and Peschard 2018). Colaço (2018) presents an extended study on characterization of phenomena, a notion which also arises in the extensive literature on biological mechanisms, particularly Darden and Craver (2013).

in some cases, natural fluctuations in the system of interest. These are reduced by collecting sufficient data. Systematic uncertainties refer to factors in the experimental arrangement that systematically bias results in a particular direction. Reducing these requires a theory of the empirical techniques employed—a conception of the way the experimental apparatus interacts with the target phenomenon and an understanding of how this interaction is best carried out in light of local contingencies.

A typical theory of technique lays out the conditions for which a data generating practice is suitable and provides researchers with a guide to the myriad factors contributing to data. This guidance typically comes in the form of causal generalizations pertaining to the context at hand, as in statements that specify general dependency relations between different factors present in the data generating set-up. A theory for modern microscopy will note, for example, that thermal activity within a CCD detector may kick out free electrons then captured by the instrument, producing a layer of background noise that systematically increases its registered signal. For an experimentalist, the import of such a generalization is ultimately practical; it directs the researcher to act in ways that alter the initially recorded data or the experimental set-up.

Explicit modeling of components may be used where precision is required. A model relating the average amount, *x*, of ambient signal build-up per CCD cell to the duration of a detection process readily leads to an imperative like, “apply a uniform signal reduction of *x* to CCD recordings!” This is an action taken after the initial recording event. Other causal claims may direct researchers to achieve a similar end by altering the experimental set-up prior to recording, say, by shielding the apparatus from ambient sources of artifactual data. A theory of technique thus bottoms out in actions meant to ensure data gathering procedures are properly calibrated, and control for sources of noise and artifacts. In this way, the experimental procedure can generate results used to reliably characterize the target, by being executed in a way that coheres with researchers’ best understanding of the techniques in use. These claims are consistent with Woodward’s (2000) account of reliability, described in terms of a systematic pattern of counterfactual dependence between data outcomes and claims about a phenomenon of interest. For Woodward, this pattern is assessed in light of scientist’s understanding of “a large number of highly specific local empirical facts about the causal characteristics of the detection or measurement process" (p. S170). This understanding, which I have called a theory of technique, allows scientists to work "backward" from data outcomes to claims about the phenomenon generating this data, and so characterize the phenomenon on this basis.

Scientists' judgments of the accuracy of their models depend on their ability to characterize the models' target phenomena. Their ability to characterize targets depends on the reliability of the data that can be produced from these phenomena. Reliability is secured through the execution of technical procedures in accord with a theory of technique. Therefore, we should expect judgments of representational accuracy to be constrained by the empirical techniques available to scientists, and to develop in tandem with them. The early history of X-ray research demonstrates this dynamic. Further, it illustrates how scientists judge the accuracy of a model even when a reliable, model-independent characterization of their target phenomenon is not available. Here there was no reliable characterization of a target phenomenon, X-ray diffraction, prior to the process that established the accuracy of a particular model of X-rays. This model was widely disregarded until it was considered alongside means for producing a novel empirical result. The accuracy of the model was thus affirmed—and a competitor model rejected—on the basis of its ability to predict this result *when integrated into a theory of technique*.

5 Empirical techniques and model accuracy in early X-ray research

It is now commonplace to model X-rays as comprised of a continuum of periodic transverse waves. For example, this model remains a principal tool for protein structure determination, despite fifty years of attempts at ab initio protein folding (Mitchell and Gronenborn 2017). Yet it had very few advocates for the better part of two decades following the discovery of X-rays. In this section, I discuss the relationship between judgments of the accuracy of this model and the body of empirical knowledge accumulated over this time. This period culminates in Lawrence Bragg's explanation of data produced Friedrich and Knipping as an instance of X-ray diffraction, modeled as the effect of a continuum of transverse waves interacting with a crystal lattice.

5.1 The technical groundwork

In a well-known episode from 1895, Wilhelm Röntgen found a high voltage cathode ray tube, completely shielded with black cardboard, still produced a fluorescent glow in a film of photoreactive material. Cathode rays, produced when an electrical field is run through a vacuum tube, had been a subject of physical research since the 1860s, and several variants of specially outfitted vacuum tubes had become recognized laboratory instrumentation by this time (Fig. 1).

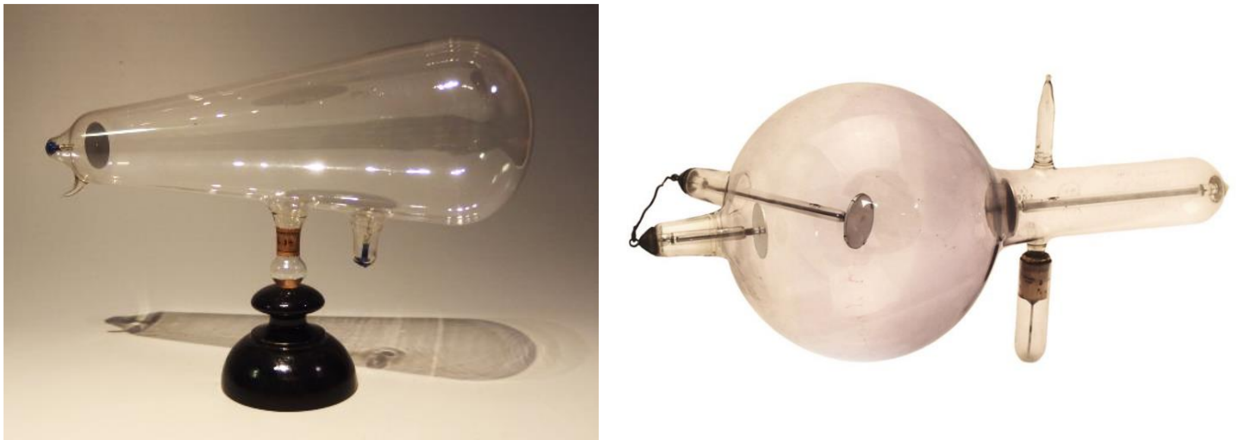


Figure 1: Left: a Crookes-Hittorf tube from 1896, with cathode is on the left side of the image. Right: a dual anode tube (c. 1900-1905), with the central anode plane angled 45 degrees to the cathode. These were sometimes referred to as "focus tubes" with a concave cathode invented by Elihu Thomson to purportedly focus the rays onto a small point of the central anode.

After a series of experiments, Röntgen announced in his paper of the same year that he had discovered "a new kind of rays." The term *strahlen*, or "rays," was already applied to visible light, ultraviolet, and the electrically charged canal or cathode rays. Characteristic among these is the tendency to travel in straight lines unless interrupted. Hence Röntgen wrote, "I find justification for using the name 'rays' […] in the very regular formation of shadows that are produced if one brings more or less transparent materials between the apparatus and the fluorescent screen (or the photographic plate)" (1995, p. 374). But these rays were different. Foremost among their novel behavior was the ability to differentially penetrate a variety of materials of the same thickness according to their density. Unlike transverse light waves, they could not be reflected or refracted by the same materials, nor was Röntgen able to produce any polarization or interference phenomena (as with diffraction patterns). Unlike cathode rays, they were not deflected when a powerful magnet was placed alongside their trajectory and their absorption was strictly proportional to the density of matter. These effects thus warranted a primitive model of X-rays as highly penetrating linear propagators with no apparent charge or characteristically transverse-wavelike behavior.

In a subsequent paper, Röntgen detailed what would become a standard technique to isolate the effects of X-rays on an experimental target. He built a box out of lead shielding and thick plates of zinc—a known absorbent of ambient electricity—save for an airtight aluminum window through which the rays could pass. Further X-ray technology was adopted rapidly, spawning a diversity of tubes (Fig. 2). For many practical purposes, there was little need to understand the theoretical nature of the rays beyond Röntgen's earliest findings. A 1910 textbook by Mihran Kassabian, a pioneer on the rays' medical applications, quotes heavily and almost exclusively from Röntgen's first papers in its summary of basic X-ray properties (Kassabian 1910). On the other hand, the book attests to a developed body of technical skills and knowledge of the experimental apparatus. One early contribution in this vein came from the engineer E. Wilbur Rice, who found his image focus noticeably improved after introducing a lead plate with a small aperture in the immediate path of the rays (Thompson 1896, Sec. 106). This addition—essentially a beam collimator—was soon adopted in virtually all X-ray experiments. In the same year, Wilbur Morris Stine tested one of Röntgen's claims about the source of the rays by irradiating photographic plates positioned at different points surrounding a tube and comparing their results. This confirmed that the source of the rays was the point where the cathode rays struck the material opposite them, not necessarily the anode (Thompson 1896, Sec. 108).

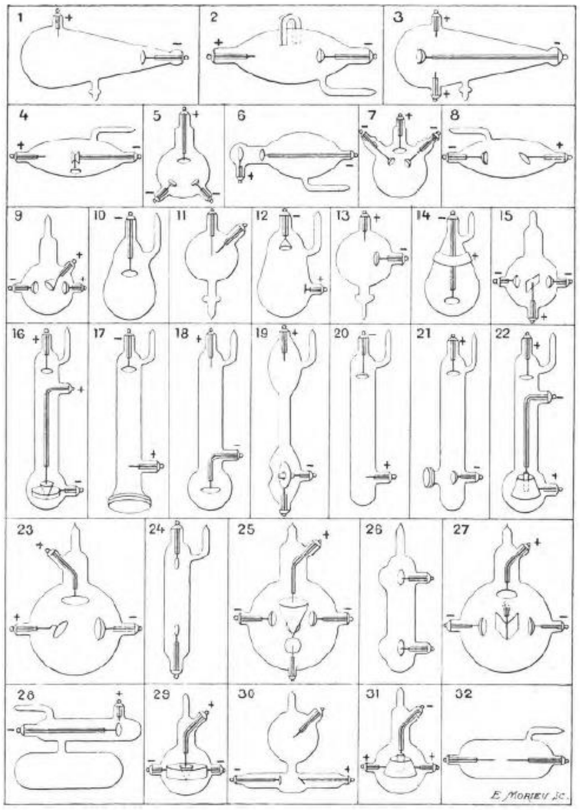
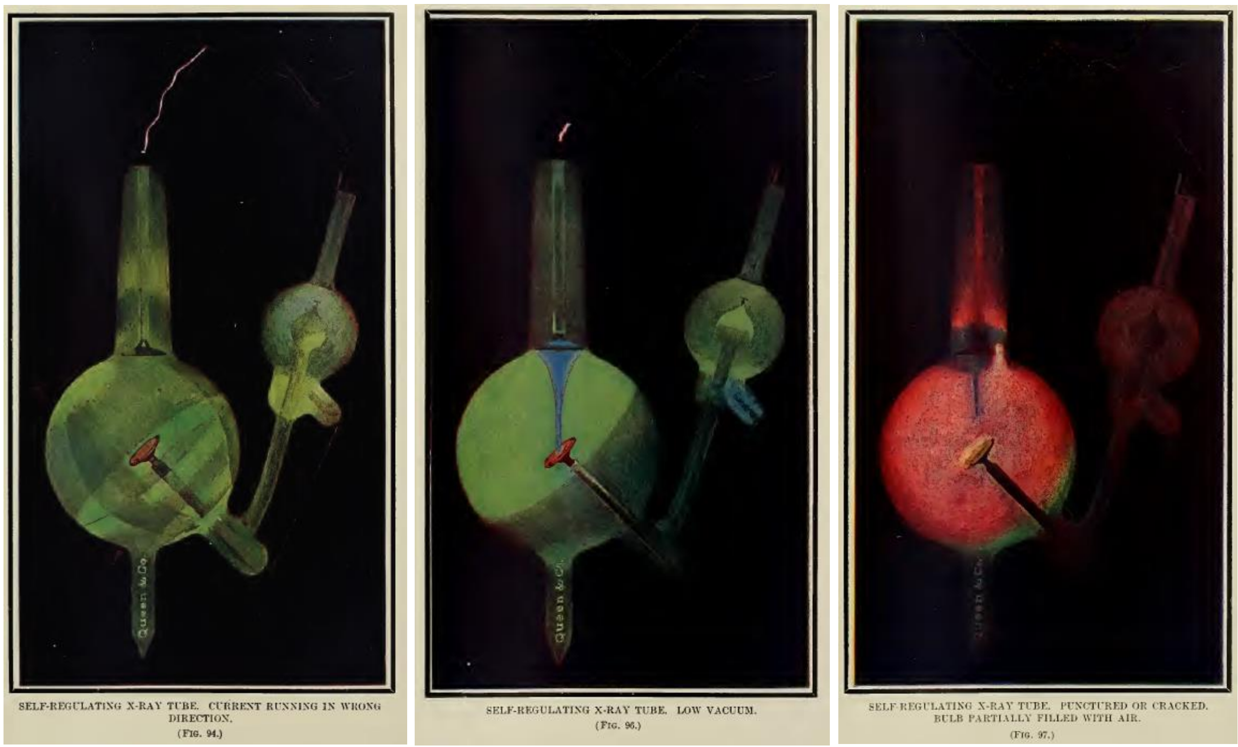


Figure 2: An assortment of tubes shown in a popular science article by the instrument maker and experimentalist Gaston Séguy (1897), including designs by Crookes (nos. 1, 2, 20) and Röntgen (nos. 24 and 32).

Despite widespread manufacture, X-ray tubes were often unreliable. The quality of the vacuum was inconsistent, not just between tubes but within the same tube over a period of use.[17] The gradual decrease in tube pressure "hardened" the X-rays emitted, increasing their penetrative power, which led to a host of additional problems. Harder rays provided less contrast in the resulting image, could puncture the glass of the tube, and led to a build-up of static electricity on its surface capable of damaging the tube (or its user) in discharge. Two ways of tackling these issues stand out in Kassabian's textbook. The first is the use of a regulator—a metallic wire "wick" extending into the tube from outside through an airtight plug that would release hydrogen into the tube when heated. This "softened the tube," in the parlance of experimentalists (Bragg 1975). The second involves a variety of naked eye tests, examining the coloration and distribution of gas within the charged tube to evaluate its quality (Fig. 3).

---

[17] Mould (1995) recounts an 1896 examination of the life span of a single tube by London hospital physicist Charles Phillips.



SELF-REGULATING X-RAY TUBE. CURRENT RUNNING IN WRONG DIRECTION.
(FIG. 94.)

SELF-REGULATING X-RAY TUBE. LOW VACUUM.
(FIG. 96.)

SELF-REGULATING X-RAY TUBE. PUNCTURED OR CRACKED. BULB PARTIALLY FILLED WITH AIR.
(FIG. 97.)

Figure 3: Plates from Kassabian (1910) showing, from left to right, a backward-running current producing illumination on the wrong side of the tube, low vacuum (i.e., high gas pressure) producing a blue light, and a cracked bulb.

In short, experimentalists developed a body of knowledge oriented around the operation of the apparatus and the manipulation of X-ray effects. In addition to the ray properties summarized by Röntgen, this included a refined understanding of the following factors:

- Conditions for preparation and production (source of emission, vacuum pressure required, induction coil strength),
- Factors modifying X-ray properties (vacuum pressure-ray hardness relation, use of regulators, anode material-beam intensity relation), and
- Factors affecting X-ray data (lead collimator, shielding, ray hardness, interactions with gas, interactions with photoreactive materials).

The roles of these factors could be summarized in a set of generalizations, such as those found in textbooks, that collectively provide a causal account of the experimental apparatus. Technical work guided by this theory of technique secured the reliability of empirical results. This is because the theory itself did not extend far beyond a summary of accumulated empirical results, with which the behavior of the apparatus was expected to cohere when producing further data. If the data was of poor quality, it provided a range of options to test and amend this: Was the tube properly positioned? The distance between the cathode and anode correct? The vacuum pressure too low? A Duhemian web, embodied in the working knowledge of the apparatus, was woven around the core characterization of X-rays established by Röntgen.

5.2 Competing characterizations

While this technical skillset developed, theoretical clarity was lagging. Kassabian's review of X-ray properties nears its conclusion by stating, "The nature and origin of the Röntgen rays is as little understood to-day as when first discovered" (Kassabian 1910, p. 151). This was not for a lack of theoretical models. Given their production by cathode ray tubes, it was natural for physicists to seek electromagnetic models of X-rays. This led Röntgen to suggest his non-diffracting rays might be the sought-after longitudinal ether waves. Schuster soon undermined the evidence favoring this model, arguing that high frequency transverse waves may also fail to diffract, refract, or reflect under the same conditions (Wheaton 1991, Ch. 2). For a moment, the transverse wave model drew interest, only to lose out to a new contender, the ether impulse. British physicists at the time suspected cathode rays consisted of charged material particles, a kind of intermediary between the purely wavelike elastic ether, and the mechanics of neutral matter. Electromagnetic theory predicted that a charged particle, when rapidly decelerated to a stop, would produce a radiative disturbance in the ether. In 1896, Stokes modeled this effect as a single spherically expanding sinusoidal "pulse" in the ethereal medium, like the shape traveling down a cracked whip. The X-ray beam, he proposed, "consists of a vast succession of independent pulses" of this nature (1898, p. 54). The paper introducing this model concluded with a theoretical explanation for the lack of observed diffraction fringes from X-rays. Being single pulses, they lacked the periodic oscillatory structure responsible for producing optical diffraction effects.

The following year, J. J. Thomson published a generally accepted proof of the material nature of cathode rays. He promptly connected this work to the production of X-rays, following Stokes's impulse theory (Thomson 1898). The German physicist Arnold Sommerfeld built on this work in 1901, providing a rich mathematical explanation for the appearance of a diffraction effect observed by Haga and Wind (1899). The Dutch team had sent X-rays through an initial slit followed by a trapezoidal slit tapering from 14 µm at the top to 2 µm at its base (Fig. 4, left). Using a square-shaped pulse model (Fig. 4, right), Sommerfeld (1901) showed these pulses would produce no diffraction fringes when passing through a slit but subtly broaden the recorded line in a manner consistent with Haga and Wind's image. With advocates among elite physicists in Britain and Germany, the ether impulse model prevailed over others.

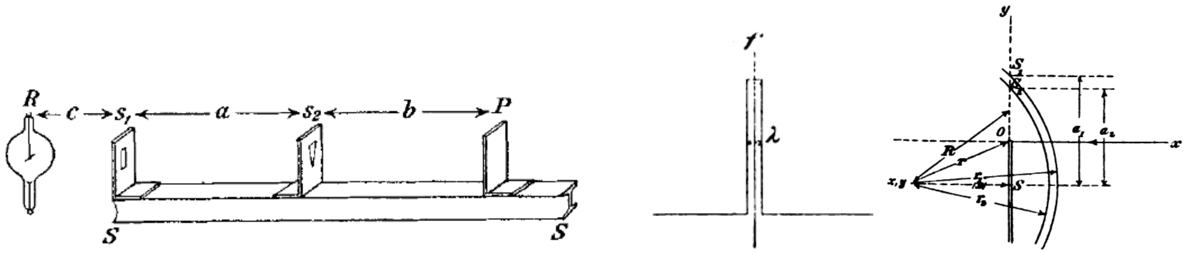


Figure 4: Left: the Haga and Wind diffraction experiment, as reproduced by Walter and Pohl (1909). Right: the radiating pulse shape posited by Sommerfeld to explain the Haga and Wind result.

A curious interplay of support for model and experiment was at work here. Haga and Wind's photograph, like many of its time, was of poor quality and difficult to interpret (Fig. 5). It was produced from over 100 hours of exposure, resulting in significant noise and, most likely, inconsistent ray quality. It was published against a background of failed attempts to demonstrate X-ray diffraction (e.g., Thompson 1896, Sec. 110 recounts attempts by Stine and Perrin) and its interpretation was disputed. A repetition of the experiment ten years later by Walter and Pohl (1909) was said to show no effect—prompting replies from Wind and Sommerfeld. As historian Bruce Wheaton recounts, "by itself the photograph submitted to support this claim would not have been convincing" (1991, p. 31). The main argument in favor of Haga and Wind's result was that they controlled for optical effects blamed for prior failures, though in Wheaton's assessment the real force behind its positive reception was coherence with Sommerfeld's treatise. At the same time, empirical results influenced what was included in Sommerfeld's model and what was not. For instance, when calculating behavior at the edge of a slit, he assumed there was no contribution from pulses reflected off the material, as no results had shown this to take place.

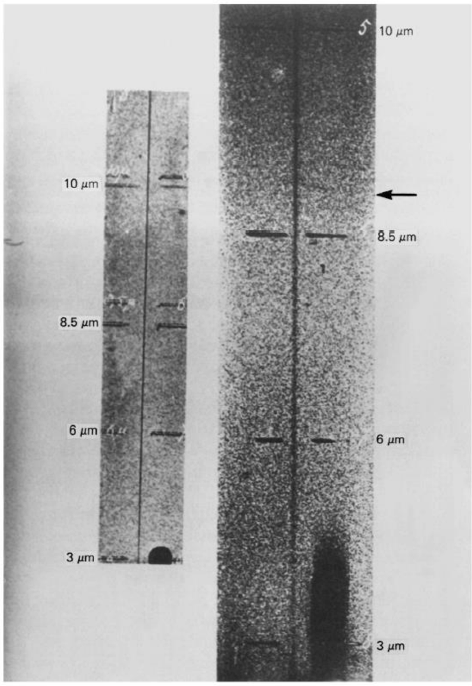


Figure 5: Photographic image from Haga and Wind's (1899) experiment, with width of the slit marked on the side. They emphasized the apparent blurring and widening of the line when the photographic plate was placed at a distance behind the second, wedge-shaped slit (right image), compared that from a single slit (left). Note also the fading of the right line image toward the bottom.

By the mid-1900s, fundamental research had slowed on the continent due to the poor quality of X-ray photographs. In Britain, where other experimental techniques were on hand, a war raged between dominant theoretical models for X-rays: Charles Barkla defended the ether impulse, while William Henry Bragg (WHB)[18] proposed a corpuscular theory according to which X-rays and the similar acting γ-rays consisted of "neutral pairs" of positive α and negative β particles.[19]. Both models could account for some X-ray phenomena—such as their penetrating power—and struggled over others. Neither, however, was a periodic wave model. Material particles were taken to behave mechanically, whereas ether pulses were independent and discontinuous disturbances in the ether with none of the extended structure required for periodicity. In fact, since the physical pulses lacked a periodic structure, the related notions of spatial frequency and wavelength could not apply. This led to confusion among Sommerfeld's readers, some of

---

[18] From here on, I refer to William Henry Bragg by these initials, WHB. I will refer to his son, William Lawrence Bragg, as Bragg or Lawrence Bragg.

[19] For an account of WHB's corpuscular theory, see Steuwer (1971).

whom took his work to provide an estimate of the wavelength of X-rays based on diffraction results.[20] Though he used the symbol for wavelength, λ, this referred only to the *width* of single pulses (as seen in Fig. 4, right). By taking their Fourier transform, a pulse could be viewed as a spectral sum of wavelengths, but for Sommerfeld this representation was non-physical; diffraction for X-rays could not be explained as a pattern of destructive and constructive interference.

This explains Sommerfeld's lack of enthusiasm in 1912 when Max Laue, a colleague in Munich, sought out his laboratory assistants, Friedrich and Knipping, for an X-ray interference experiment. Recent results had begun to suggest that X-rays and light, the latter typically modeled as transverse waves, were not as different as first thought.[21] A conversation with Paul Ewald had inspired Laue to see whether the atoms of a crystal could function as a diffraction grating for X-rays, on the assumption that these rays possessed a short wavelength. Munich had been home to Leonhard Sohncke, a theorist of crystal symmetries, and it was commonly accepted that crystals were comprised of a regular, repeating molecular structure that would provide such a grating (Forman 1969). But to Sommerfeld, the vanishing line image at the base of Haga and Wind's photograph signified that X-rays could not pass through a 2 μm slit, much less a gap between crystal atoms. Further, any X-ray pulses that might pass through would be of varying widths and scatter in every direction, producing only a uniform darkening of the photograph (Wheaton 1991). He strongly discouraged Laue's experiment on these grounds.

A model of X-rays as comprised of many periodic waves was clearly not deemed accurate by theorists. Yet Laue proposed that the so-called secondary rays emitted by an irradiated crystal (as opposed to the "primary" rays produced by the X-ray tube) might have a more uniform structure capable of producing a diffraction pattern.[22] Laue, Friedrich, and Knipping selected a crystal of copper sulfate, believing this would yield the most uniform secondary rays, and set to work (Fig. 6).

---

[20] This includes Lawrence Bragg's (1975) recounting of Sommerfeld's derivations.

[21] For instance, two years prior researchers found that the photoelectric effect had the same forward scattering asymmetry that WHB found in X-rays (Wheaton 1991).

[22] This was based on Ewald's work (following Lorentz), which studied the optical effects of crystals modeled as a lattice of resonators that would collectively vibrate when stimulated by incident light (Ewald 1962).

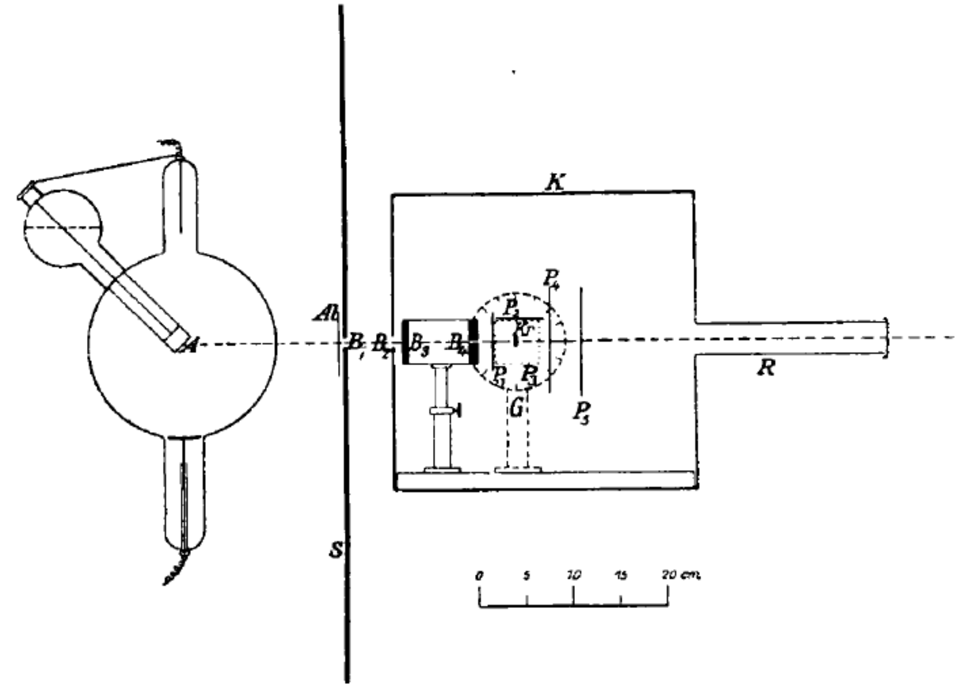


Figure 6: Friedrich and Knipping's experimental set-up (Friedrich et al. 1913). The crystal Kr *is located within a lead box* K. *After passing through an aluminum window and several collimators (to shield off secondary rays), the X-ray beam strikes the crystal and is scattered onto the plate* P_4 *placed behind it.*

They initially placed photographic plates placed on each side of the crystal, parallel to the beam line, under the assumption that its molecules would emit spherical waves.[23] When they failed to produce a clear image, other arrangements were tried until they found that a plate positioned behind the crystal yielded a crude, but roughly symmetrical two-dimensional pattern. When the plate was moved back, the image spread out. When the crystal was ground up, the data was uninterpretable. The effect thus appeared to spread outward from the internal structure of the crystal. A repetition with zincblende crystal irradiated for many hours produced more clearly resolved data (Fig. 7, left).

Laue developed a mathematical model of diffraction in three dimensions, seeking to interpret the pattern of spots as the maxima of interfering secondary waves from the crystal lattice. Using this model, he showed the proportions between (i) the coordinates (x, y) of each spot on a plate at a distance z from the crystal should match those of (ii) the scattering angles (α, β, γ) of the contributing wave in each

---

[23] This assumption was also due to Ewald's resonator model.

dimension. Laue used these to derive fixed ratios for integer parameters (h_1, h_2, h_3) that could be used to determine a wavelength λ responsible for this spot via the "Laue conditions":

$$h_1\lambda = a\alpha \;\; ; \;\; h_2\lambda = \; a\beta \;\; ; \;\; h_3\lambda = a(1-\gamma), \;\; where \;\; \alpha{:}\,\beta{:}\,\gamma :: x{:}\,y{:}\,z$$

One merely had to take the values x, y, z for recorded spots, find the corresponding ratio $\alpha{:}\beta{:}\gamma$, and then derive the ratio of $\alpha{:}\beta{:}(1 - \gamma)$ to see how the h's were related. However, Laue's method did not uniquely determine his results. For any ratio $\alpha{:}\beta{:}(1 - \gamma)$, infinite triplets of integers h_1, h_2, h_3 might satisfy it. To account for this, Laue decided to stick to small integer values of h and assumed that the secondary radiation from the crystal consisted of only a small number of wavelengths. He provided five that yielded a pattern similar to that found in the photograph from zincblende (Fig. 7, right).

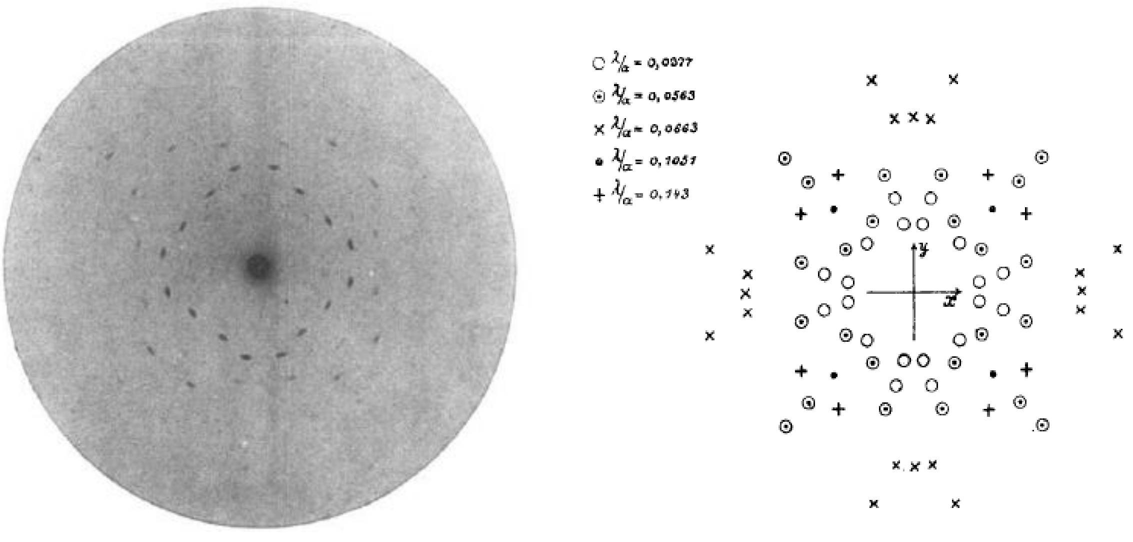


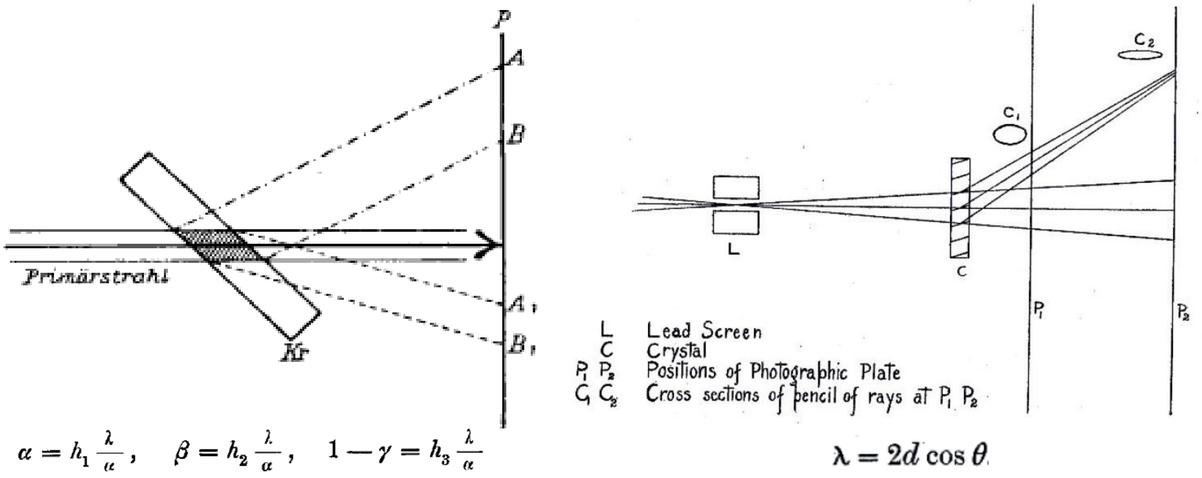
Figure 7: Laue's derivation (Friedrich et al. 1913). Left: the diffraction pattern produced by zincblende crystal. Right: Laue's reconstruction of the data using five selected wavelengths.

Even with this similarity on hand, there was resistance to a model that allowed for X-ray interference. Barkla expressed strong skepticism in a letter to Rutherford that the data was the effect of interference (Forman 1969). Laue followed Sommerfeld even after his experiment, claiming in 1913 that an incident beam consisting of a continuum of wavelengths would only darken a photograph (Ewald 1962). William Lawrence Bragg, the elder WHB's son, was unsatisfied with this result for another reason. By assuming only five wavelengths contributed to the recorded result, Laue had arbitrarily restricted the freedom provided by his model. Further, these same values chosen by Laue also predicted some spots that were absent from the photograph. These left the adequacy of Laue's model in question. As Bragg put it, "there are a great many combinations of h_1 h_2 h_3 which might give spots on the photographic plate which are in fact, not there, and there is no obvious difference between the numbers h_1 h_2 h_3 which correspond to actual spots, and those which are not represented" (1913a, p. 44-45).

While studying physics at Cambridge, Bragg encountered models developed by William Barlow and William Pope, two of England's foremost theorists on symmetry and atom packing in crystals. These showed that within any crystal one could locate evenly spaced parallel planes of repeating atomic arrangement. Bragg raised the possibility that X-rays might be reflected off these planes and was encouraged to attempt it with mica. A report of his success was published in *Nature* that year, noting that "variation of the angle of incidence and of the distance of plate from mica left no doubt that the laws of reflection were obeyed" (Bragg 1912, p. 410). These findings led to a model for diffraction that remains in use today.

Bragg's solution exploited a mathematical equivalence, well-known in British optics, between the diffraction of light modeled as either discontinuous pulses or as a continuous range of wavelengths—a routine model for white light. The condition for incoming X-rays forming a constructive maximum (i.e., a diffraction spot) could then be derived as a relation—later dubbed "Bragg's law"—between their constituent wavelengths and the angle at which they reflect off parallel planes of the crystal. In other words, a portion of each incoming ray would be reflected off an atom in a plane while another part may pass through (Fig 8 right). If the distance between planes and the angle of the incident ray was correctly balanced, a ray reflected off a subsequent plane would have traveled just the right distance to be in phase with those reflected from a higher plane, resulting in constructive interference. Different atomic planes within the crystal will be tilted at different angles with respect to the beam path, depending on the crystal's symmetries, and these will select for reflections of different wavelengths within the ray, forming a pattern of distinct spots corresponding to the internal structure of the crystal. Assuming a particular symmetry group for the zincblende and using this model, Bragg was able to reconstruct Friedrich and

Knipping's data, spot for spot. In summary, he claimed "the interference patterns can be ascribed to diffraction of a 'white' radiation by a set of points on a space lattice" (1913b, p. 276).[24]



$$\alpha = h_1 \frac{\lambda}{a}, \quad \beta = h_2 \frac{\lambda}{a}, \quad 1 - \gamma = h_3 \frac{\lambda}{a} \qquad \lambda = 2d \cos\theta$$

Figure 8: Two models of the diffraction experiment. Left: Laue's diagram showing the primary rays producing secondary radiation from the crystal and forming spots off the beam axis at angles corresponding to their wavelength. The core mathematical contribution of this model are the Laue conditions, giving three parameters (the h *terms) relating scattering angle (greek letters) and inter-atomic distance,* a*, to ray wavelength (Friedrich et al. 1913). Right: Bragg's diagram showing X-rays of a preferred wavelength reflected off crystal planes to the same point. Bragg's Law, below, relates the wavelength to the angle reflection and the distance,* d*, between crystal planes (Bragg 1913).*

Bragg's model initiated new research programs in crystallography and spectrometry—the first canonical applications that required knowledge of X-rays beyond Röntgen's initial results—which he and his father avidly pursued. Compelled by these results, physicists across Europe rapidly adopted this view of X-rays as comprised of a continuum of wavelengths. In a review at the end of 1913, Laue and Sommerfeld's research assistant, Friedrich, claimed this work showed "the presence of a continuous spectrum" emitted from the X-ray tube (1913, p. 1081). Following Bragg's announcement, Barkla and a collaborator sought to replicate crystal diffraction effects (Allen 1947). In an early statement of their findings, the former skeptic wrote, "we cannot at present suggest any probable explanation except the very obvious one of interference" (Barkla and Martyn 1913, p. 647). His Nobel speech several years later contains no references to ether pulses, instead characterizing primary rays as consisting of a range of wavelengths (Barkla 1967).

---

[24] Bragg's references to white light during this period, as well as Röntgen's adoption of the "ray" concept, could be fruitfully analyzed as instances of the scientific use of analogy (Hesse 1963; see Rowbottom (2019) for applications of Hesse's account that bypass her latent realism).

5.3 Model accuracy and techniques for X-ray diffraction

This historical process, which established the accuracy of the wave continuum model for specific X-ray phenomena, shows a mutual interplay between judgments of accuracy and empirical techniques. A skilled experimentalist would recognize the procedural steps taken by Friedrich and Knipping as part of an accepted theory of technique. The alignment of dual anode tube, aluminum window, multiple collimating apertures, lead shielding, and photographic plates (Fig. 6) were the standardized repertoire for producing reliable X-ray results, developed through an accumulation of causal generalizations over laboratory arrangements. These were the basis for characterizing X-ray phenomena in greater depth during the period leading up to 1912. They established an initial target for model-based inferences, which were in turn licensed by their consistency with electromagnetic or mechanical principles. As representational tools, these models allowed for predictive inferences beyond familiar results. With motivation from models, experimentalists drew on their theory of technique to construct novel experimental set-ups, as seen in the experiments undertaken by competitors WHB and Barkla.[25] New characteristic behaviors, once established in the lab, could then be integrated into the developing theory of technique. These, in turn, placed new demands on models designed to account for an expanding range of empirical results.

Bragg's model would not have been accepted by his scientific community if it were not for Friedrich and Knipping's technical work. There was scant enthusiasm for this model within the community of X-ray researchers prior to 1912. Even after Laue et al.'s paper was publicized, other models were first suggested. Pulse theorists like Barkla questioned whether the data was due to wave interference at all. Yet, given the reliability of their techniques, there was little doubt that their data were the effect of *some* interaction between X-rays and crystals. Both Laue and Bragg depended on this fact in applying their models. They were confident in abstracting from everything but the relations between the

---

[25] There, X-ray scattering and ionizing behaviors established by Röntgen were causally linked in set-ups designed to examine the characteristics of forward, backward, and right-angle scattering—behaviors over which different models disagreed.

rays, crystal, and photographic plate because Friedrich and Knipping had reliably isolated the interaction between these components in their data. But the model presented by Laue could only reproduce their data by admitting a degree of arbitrariness in its method of application. By contrast, Bragg's model uniquely determined the structure of the data on hand without predicting additional points that were not found in the recorded pattern. These advantages were easily recognized. In fact, it was largely because of Bragg's work that Friedrich and Knipping's data were fully accepted as evidence *of diffraction*. Broader reception of the wave continuum model of X-rays soon followed. In this way, Bragg's model played a constitutive role in the scientific community's characterization of its target phenomenon. Its success depended on confidence that the data were reliably connected to a particular interaction and its ability to simulate the process isolated by these techniques.

6 The model-integration criterion of representational accuracy

Scientists' judgments of the accuracy of the wave continuum model evolved alongside the empirical techniques used to characterize X-ray phenomena. A similar strategy for judging this accuracy is found in both Sommerfeld's application of the ether impulse model and Laue and Bragg's use of wave models. These authors proceeded by integrating their models into an inferential framework that reflected the specificities of the data generating process. Additional structure was added to the model of interest that reflected the interaction that experimentalists had reliably isolated: in Sommerfeld's case, the interaction between a traveling ether pulse and a geometrically modeled slit; in Laue and Bragg's case, that between traveling X-ray waves, both among themselves and with a crystal structure. By this means, they derived model results that could be compared with recorded data.

Let's see how this worked in the cases of Laue and Bragg. Guided by a theory of technique, Friedrich and Knipping were able to record an interaction between a crystal and X-ray apparatus. This allowed for an initial characterization of their data as reliably produced by this interaction (Figure 9).

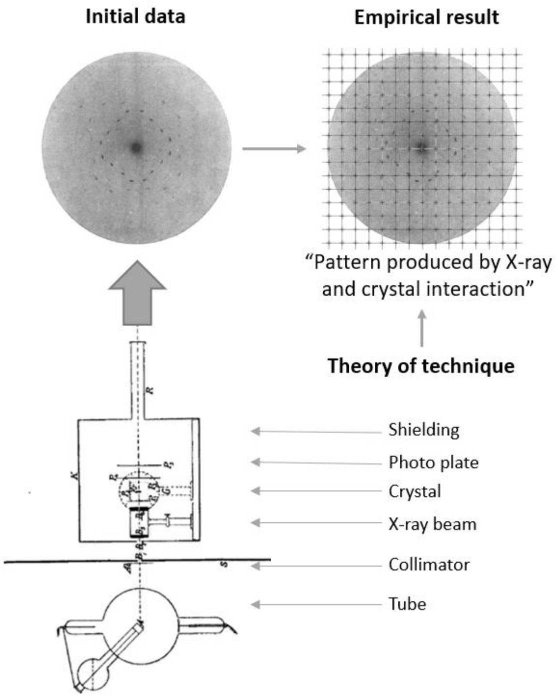


Figure 9: Producing an initial empirical result using a theory of technique.

Laue and Bragg developed two different models to account for this pattern as a diffraction effect. Each took a particular model of the X-rays—for Laue, a finite set of wavelengths emitted from an irradiated crystal; for Bragg, a continuum of waves reflected by the crystal layers—and incorporated this into the experimental setting as described by the theory of technique. Both Laue and Bragg developed a model of the process isolated within the experiment (Figure 10).

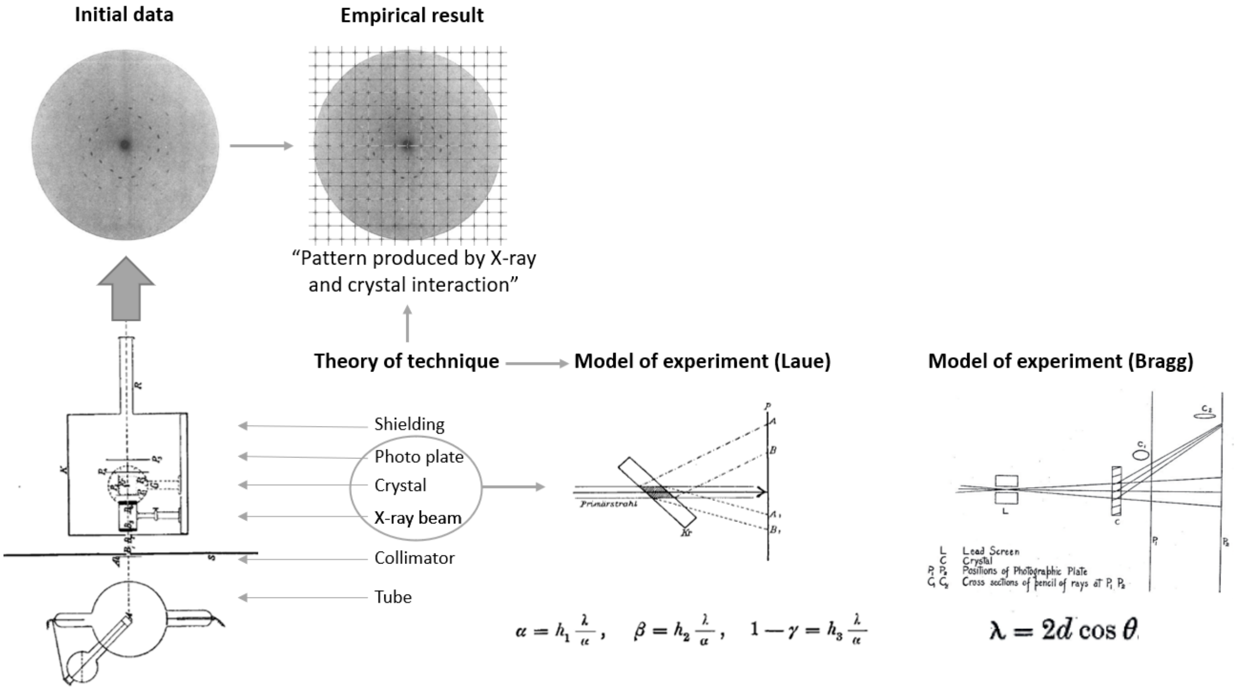


Figure 10: Incorporation of different wave models into the description of the experiment.

Once integrated with a theory of the technique, these models could be used to produce derivations that were comparable to the empirical result (Figure 11).

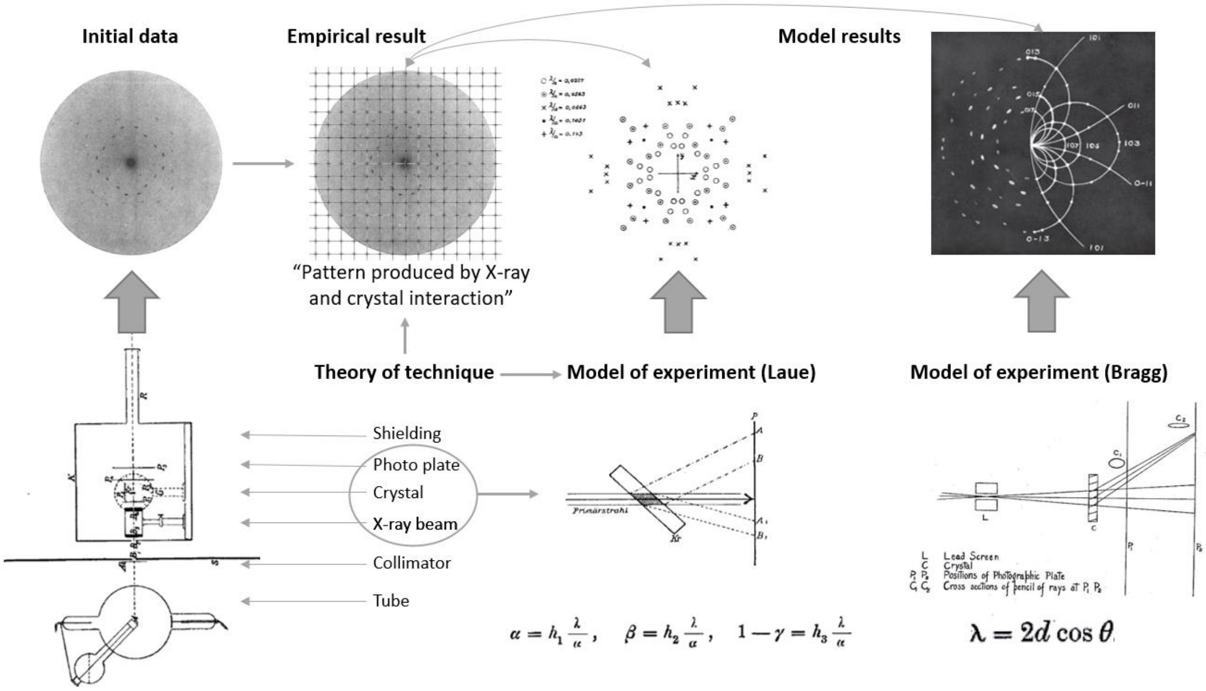


Figure 11: Derivation from model and comparison with empirical results.

This enabled the models to be compared to data *and* to further characterize their target phenomenon. The models provided a means to interpret Friedrich and Knipping's pattern as the product of X-ray wave interference. It wasn't the comparability to data alone that underwrote judgments of Laue and Bragg's models; it was the fact that these models could predict data patterns in a way that cohered with the theory of technique. Moreover, Bragg's model was accepted over Laue's not only because it more precisely fit the data pattern but also because no further, arbitrary decision was required on the part of the modeler to produce this fit once the model of the experiment was in place. In such a case, characterizing a target phenomenon from experiment is at the same time a model-based predictive process with respect to an empirical result.[26]

The likelihood ratio and test response examples from Section 3 showed that a mere fit between a model and an empirical result is insufficient for judging a model to be accurate. In addition, the empirical result must be reliably connected to the phenomenon that the model user targets. Cases from the history of X-ray research demonstrate a general strategy for judging a model's accuracy on this basis, even when a prior model-independent characterization of the target phenomenon is absent: include the theoretical model as a component of an account of the data-gathering process, further specifying or concretizing the model as necessary, and use *this* to derive results comparable to the empirical result. This is an instance of a general criterion according to which a model is accurate if and only if integrating the model into a theory of the data acquisition process yields well-fitting predictions of patterns in the data. The theory of technique functions as a linchpin in this account, serving two inferential roles. One is retrodictive, moving from recorded data to an interpretation of these in terms of causal factors within the data acquisition process. The other is predictive, working in conjunction with the model to simulate a data pattern. The first specifies those factors in the experimental set-up that are the primary contributors to the data. The second directs the inclusion of additional elements into a model and their interpretation in terms of these specified factors. Hence both Laue and Bragg incorporate the X-rays, the crystal, photographic plate, and

---

[26] There is a parallel here to Tal's (2017) account of measurement as the outcome of a predictive model of the measuring process.

their spatial relations into their accounts. By tying these terms to elements of their models, they embed it within the set of inferences that comprise an established theory of technique.

This provides scientists with a means for evaluating a model's accuracy: it must simulate a result that can be compared to a data pattern *and* it must cohere with the technique that generates that data. This is not a post hoc reconstruction of the model's success or failure in terms of its relation to the target. Rather, it is a way of incorporating the model into an account of the apparatus, and so into the practical inferences that guide actions in this context. For example, integrating a model into a theory of technique licenses the move from a claim like, "my spectroscopic analyzer isn't showing a strong reading at this angle" to "the reflected X-rays are not interfering constructively at this point." With the Bragg model on hand, one can sensibly respond to the above inference with the suggestion to rotate the *crystal* rather than the detector. A significant contributor to the model's accuracy is whether it the practical inferences it licenses bear out in such contexts. On the other hand, the only direct comparison taking place, of the kind that resembles a similarity judgment, does not concern the model and its target but its simulated result and a data pattern. When a model is properly integrated into this broader inferential network, and its simulated results satisfy standards of fit vis-à-vis recorded data, then the model is accurate with respect to its target phenomenon. This criterion of accuracy is broadly applicable by practicing scientists and is readily found in many areas of model-based research.[27]

7 Ontic Priority reconsidered

A few points of clarification are in order. First, this argument only pertains to how criteria of model *accuracy* should be addressed by a theory of scientific representation. It does not entail that, in all cases, scientists or philosophers ought to cease thinking of models in abstraction from data-gathering practices. I do not deny that scientific models are employed in abstraction from specific practices for legitimate

---

[27] To take one example, the methods of judging the accuracy of theoretical models of particle physics at the LHC accord with this criterion. There, models of collision events are first specified based on the operating energy of the accelerator using a suite of techniques referred to as particle phenomenology, their decay paths are simulated, and then integrated into a much larger model of the detector itself to generate a prediction of LHC data.

scientific purposes other than evaluations of accuracy.[28] By the same token, this argument does not entail that a model's accuracy can only be determined with respect to a single data acquisition process. Precisely because of their capacity to be abstracted from specific contexts, theoretical models can be integrated into multiple lines of inquiry. Though the success criterion for model accuracy always depends on the way it is integrated into a theory of technique, the specific character of this criterion—including the standard of fit applied to its predictions—may vary with the form of target-apparatus interaction theorized in each case. Accordingly, the same model can be judged by different criteria based on different forms of interaction with the same target phenomenon and may even yield different judgments of accuracy in each case. I have proposed a general criterion for model accuracy, but this is compatible with pluralism regarding the specific form the model-data comparison can take in different instances and the corresponding judgment that results.

Still, I anticipate this view will not satisfy proponents of Ontic Priority. I have said nothing about the nature of the relation between a model and its target phenomenon that, one might think, grounds the accuracy of a scientific model. On this view, it is a fact whether such a relation obtains, and thus it is a fact whether the model is accurate or not. I have only described how scientists can judge the accuracy of models, a form of judgment that might vary based on the theoretical background and technique in use. But, on an ontic view, whether a model is an accurate representation of a target is a fact that cannot vary due to these epistemic factors. Whatever the method that scientists use to judge a model's accuracy, its success is only explained by the appropriate relation obtaining between model and target.

I have responded to this line reasoning by raising the issue of scientists' epistemic access to such facts. In Section 3, I argued that the relation between model and target theorized in abstraction from epistemic factors is a kind of fact that is often inaccessible from the standpoint of scientific practice. There is no general way to infer *from* the kind of information that scientists can access *to* the precise nature of the relation between a model and a target, beyond how the model itself coheres with the data

---

[28] See footnote 2. Though models arguably retain implicit empirical content via the process of their construction and reception within a research community (Cf., Knuuttila 2011; Knuuttila and Loettgers 2017; Boesch 2017).

acquisition process. The argument I take to favor this model-integration account is based on the further desideratum that an adequate theory of scientific representation ought to propose success criteria for model accuracy that scientists themselves can, and do, apply. But a defender of Ontic Priority can push back here by distinguishing, on the one hand, the ontic relation in virtue of which a model is successful and, on the other, the facts about modeling success that are epistemically accessible.[29] They may concede that these latter facts might not include or even entail identifying information about the ontic relation itself, but still insist that it would be a mistake to draw conclusions about what accuracy *is* from facts about how one comes to judge a model to be accurate. One might modify my desideratum on this basis, proposing that a theory of representation ought to provide both a metaphysics of accuracy *and* an epistemology of those facts and procedures that determine accuracy judgments.

It's a fair proposal. The basic assumptions that comprise a metaphysics of accuracy—that representation is grounded in a strictly ontic model-target relation and that accuracy can be defined in terms of whether this relation obtains—are self-consistent and are not contradicted by my view. By these counts, ontically-minded philosophers are perfectly free to consider the possible relations between models and targets, while others fuss over practice. The only issue concerns the explanatory connection between these two realms. Defenders of Ontic Priority claim that the successes of model-based reasoning observed in scientific practice cannot be explained without an account of this ontic relation. But if the facts relevant to evaluations of modeling success that are accessible to scientists neither include nor entail facts about this ontic relation, and if philosophers do not have privileged access to certain facts over the methods available to science, then it seems we are left with two choices: Either we accept that this success can be explained in terms that reflect scientists' evaluative resources (i.e., without reference to an ontic relation) or we are left to explain this success by appeal to facts that are epistemically inaccessible.

I have proposed a notion of accuracy that follows the first disjunct. Contra Ontic Priority, I affirm that the model-integration criterion *does* explain how models are used to accurately reason about target

---

[29] I thank an anonymous reviewer for raising this point.

phenomena. It gives an account of models being integrated into a theory of technique, of their use to simulate data, and of a comparison being made between simulated and reliably recorded results. This is not the kind of explanation favored by an ontic-first approach, but the alternative appears deeply obscure from the standpoint of scientific practice. The second disjunct invokes a form of explanation that raises more questions than it appears able to answer, insofar as it depends on assertions of fact that cannot be confirmed through the methods of fact-finding available to us. What, after all, is the difference between an explanation based on the existence of an epistemically inaccessible fact and one based on an a priori postulate? Each explanans plays the same functional role, insofar as their truth cannot be evaluated a posteriori. This is a particularly strange role for factual claims, since these are just the things that we typically regard as evaluable a posteriori. Whatever these inaccessible facts and accompanying explanations are, they seem to belong to a realm that is unfamiliar to the naturalistic philosopher of science.

At issue here are the conjoined claims that the ontic explains the epistemic and that ontic questions have methodological priority over the epistemic. If a metaphysics of accuracy is a desirable complement to my epistemic account, and if it is to be compatible with naturalism, then one of these must be given up. First, a metaphysics of accuracy could be defended on grounds other than its explanatory merit. One could argue that there is conceptual value to considering hypothetical relations between models and targets in isolation from the question of how these are empirically instantiated and affect scientists' determinations of accuracy. In this case, a notion of accuracy might be investigated as a purely theoretical construct with no direct bearing on practices of model-target comparison. Distinct notions of ontic and epistemic accuracy could be developed for distinct philosophical ends. On the other hand, if this separation is unsatisfactory, then the connection between scientists' evaluations of modeling success and the ontic conditions in virtue of which their models are successful needs to be articulated in further detail. One way to bring an ontic account into harmony with the epistemic view defended here would be to incorporate the latter's pluralism and contextualism. On this view, different ontic relations will play the role of grounding representational accuracy in different cases, in a manner that is sensitive to the practical and theoretical context at hand. Here a practice-first approach may be more fruitful for analysis. Instead of starting from a singular relation, inquiry can proceed from scientists' evaluations of accuracy in contexts of empirical application, working out the metaphysical implications on this basis.

References

Allen, H. (1947): "Charles Glover Barkla, 1877-1948", *Obituary Notices of the Fellows of the Royal Society*, **5**, pp. 341-66.

Barkla, C. (1905): "Polarised Röntgen Radiation", *Philosophical Transactions of the Royal Society of London A*, **204**, pp. 467-79.

----- (1967): "Characteristic Röntgen Radiation", from *Nobel Lectures, Physics 1901-1921,* Amsterdam: Elsevier, pp. 392-99.

Barkla, G. and Martyn, G. (1913): "An X-Ray Fringe System", *Nature*, **90**, p. 647.

Bragg, L. (1912): "The Specular Reflection of X-Rays", *Nature*, **90**, p. 410.

----- (1913a): "The Diffraction of Short Electromagnetic Waves by a Crystal", *Proceedings of the Cambridge Philosophical Society*, **17**, pp. 43-57.

----- (1913b): "The Structure of Some Crystals as Indicated by Their Diffraction of X-rays", *Proceedings of the Royal Society of London, Series A*, **89**, pp. 248-277.

Bragg, W. H. and Glasson, J. (1907): "On a Want of Symmetry shown by Secondary X-Rays", *Proceedings of the Physical Society of London*, **21**, pp. 735-45.

Boesch, B. (2017): "There is a Special Problem of Scientific Representation", *Philosophy of Science*, **84**, pp. 970-81.

Bogen, J. and Woodward, J. (1988): "Saving the Phenomena", *The Philosophical Review* **97**, pp. 303-52.

Borsboom, D. (2005): *Measuring the Mind,* New York: Cambridge University Press.

Bueno, O. (1997): "Empirical Adequacy: A Partial Structures Approach", *Studies in the History and Philosophy of Science,* **28**, pp. 585-610.

Bueno, O. and French, S. (2018): *Applying Mathematics*. New York: Oxford University Press.

Cartwright, N. (1983): *How the Laws of Physics Lie*. New York: Oxford University Press.

Chakravartty, A. (2010): “Informational versus functional theories of scientific representation”, *Synthese*, **172**, pp. 197-213.

Chang, H. (1995): “Circularity and Reliability in Measurement”, *Perspectives on Science*, **3**, pp. 153-72.

Colaço, D. (2018): “An Investigation of Scientific Phenomena”, PhD diss., University of Pittsburgh.

Contessa (2007): “Scientific Representation, Interpretation, and Surrogative Reasoning”, *Philosophy of Science*, **74**, pp. 48-68.

Coopmans, C., J. Vertesi, M. Lynch and S. Woolgar (*eds*.), (2014): *Representation in Scientific Practice Revisited*. Cambridge, MA: MIT Press.

Darden, L. and Craver, C. (2013): *In Search of Mechanisms*, Chicago: University of Chicago Press.

Elliot-Graves, A. (2020): “What is a Target System?”, *Biology and Philosophy*, **35**. Accessed online: <https://link.springer.com/article/10.1007/s10539-020-09745-3>

Ewald, P. (1962): *Fifty Years of X-Ray Diffraction*, Netherlands: Utrecht.

Forman, P. (1969): “The Discovery of the Diffraction of X-rays by Crystals; A Critique of the Myths”, *Archive for History of Exact Sciences*, **6**, pp. 38-71.

French, S. and Ladyman, J. (1999): “Reinflating the semantic approach”, *International Studies in the Philosophy of Science*, **13**, pp. 103-121.

Friedrich, W. (1913): “Röntgenstrahlinterferenzen”, *Physikalische Zeitschrift*, **14**, pp. 1079-87.

Friedrich, W., Knipping, P., and Laue, M. (1913): “Interferenzerscheinungen bei Röntgenstrahlen”, *Annalen der Physik*, **346**, pp. 971-88.

Frigg, R. (2010): “Fiction and Scientific Representation”, in M. C. Hunter (*ed*.), *Beyond Mimesis and Convention*, New York: Springer, pp. 97-138.

Frigg, R. and Nguyen, J. (2017): “Models and Representation”, in L. Magnani and T. Bertoletti (*eds*.), *Springer Handbook of Model-Based Science*, Switzerland: Springer.

Giere, R. (2004): “How Models Are Used to Represent Reality”, *Philosophy of Science* **71**, pp. 742-52.

Glasser, O. (1936): “What Kind of Tube Did Röntgen Use When He Discovered the X-Ray?”, *Radiology*, **27**, pp. 138-40.

Godfrey-Smith, P. (2006): “The strategy of model-based science”, *Biology and Philosophy* **21**, pp. 725-40.

----- (2009): “Models and fictions in science”, *Philosophical Studies* **143**, pp. 101-16.

Kassabian, M. (1910): *Röntgen Rays and Electro-Therapeutics*, Philadelphia: Lippincott.

Kaveh, S. (Forthcoming): “What Helium Teaches Us About the Success and Failure of Dynamical Theories”, *Proceedings of the 8th Integrated History and Philosophy of Science Conference*.

Knuuttila, T. (2011): Modelling and representing: An artefactual approach to model-based representation”, *Studies in History and Philosophy of Science*, **42**, pp. 262-71.

Knuuttila, T. and Loettgers, A. (2017): “Modelling as Indirect Representation? The Lotka-Volterra Model Revisited”, *British Journal for the Philosophy of Science*, **68**, pp. 1007-1036.

Haga, H. and Wind, C. (1899): “Die Beugung der Röntgenstrahlen”, *Annalen der Physik und Chemie*, **68**, pp. 884-918.

Hesse, M. (1963): *Models and Analogies in Science*, London: Sheed and Ward.

Longino, H. (1990): *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton: Princeton University Press.

Mayo, D. (1996): *Error and the Growth of Experimental Knowledge*, Chicago: University of Chicago Press.

----- (2010): *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science*, New York: Cambridge University Press.

Mitchell, S. (2020): “Instrumental Perspectivism: Is AI Learning Technology Like NMR Spectroscopy?”, in M. Bertolaso and F. Sterpetti (*eds*.), *A Critical Reflection on Automated Science*, Switzerland: Springer.

Mitchell, S. and Gronenborn, A. (2017): “After Fifty Years, Why Are Protein X-ray Crystallographers Still in Business?”, *Philosophy of Science*, **68**, pp. 703-23.

Mould, R. (1995): “The early history of X-ray diagnosis with emphasis on the contributions of physics 1895-1915”, *Physics in Medicine & Biology*, **40**, pp. 1741-87.

Nguyen, J. (2016): “On the Pragmatic Equivalence between Representing Data and Phenomena”, *Philosophy of Science*, **83**, pp. 171-191.

Peschard, I. (2012): “Forging Model/World Relations: Relevance and Reliability”, *Philosophy of Science*, **79**, pp. 749-60.

Pickering, A. (1992): “From Science as Knowledge to Science as Practice”, in A. Pickering (*ed.*), *Science as Practice and Culture,* Chicago: University of Chicago Press, pp. 1-28.

Pincock, (2011a): *Mathematics and Scientific Representation,* New York: Oxford University Press.

----- (2011b): “Modeling reality”, *Synthese* **180**, pp. 19-32.

Rice, C. (2019): “Models Don’t Decompose That Way: A Holistic View of Idealized Models”, *British Journal for the Philosophy of Science*, **70**, pp. 179-208.

Röntgen, W. (1995): “On a New Kind of Rays”, *Veterinary Radiation and Ultrasound*, **36**, pp. 371-4.

Rowbottom, D. (2019): *The Instrument of Science,* New York: Routledge.

Sagnac (1899): “Transformations des rayons X par la matière”, *Journal de Physique*, **8**, pp. 65-89.

Séguy, G. (1897): “Tubes for the production of Roentgen rays”, *Nature,* **55**, pp. 296-7.

Seliger, H. (1995): “Wilhem Conrad Röntgen and the Glimmer of Light”, *Physics Today,* **48**, pp. 25-31.

Sommerfeld, A. (1901): “Theoretisches über die Beugung der Röntgenstrahlen”, *Zeitschrift für Mathematik und Physik,* **46**, pp. 11-97.

Steuwer, R. (1971): “William H. Bragg’s Corpuscular Theory of X-Rays and γ-Rays”, *The British Journal for the History of Science,* **5**, pp. 258-81.

Stokes, G. G. (1898): “On the Nature of the Röntgen Rays”, in G. Barker (*ed.*), *Röntgen Rays,* New York: American Book Company, pp. 43-65.

Suárez, M. (2004): “An Inferential Conception of Scientific Representation”, *Philosophy of Science* **71**, pp. 767-79.

Suppes, P. (1962): “Models of Data”, in E. Nagel, P. Suppes, and A. Tarski (eds.), *Logic Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*, Stanford: Stanford University Press, pp. 252-61.

Tal, E. (2013): “Old and New Problems in Philosophy of Measurement”, *Philosophy Compass*, **8**, pp. 1159-1173.

----- (2017): “Calibration: Modelling the measurement process”, *Studies in History and Philosophy of Science*, **65-66**, pp. 33-45.

Teller, P. (2001): “Twilight of the Perfect Model Model”, *Erkenntnis,* **55**, pp. 393-415.

Thompson, E. (1896): *Roentgen Rays and Phenomena of the Anode and Cathode*, New York: Van Norstrand.

Thomson, J. J. (1898): "A Theory of the Connection between Cathode and Röntgen Rays", in G. Barker (*ed.*), *Röntgen Rays*, New York: American Book Company, pp. 69-72.

van Fraassen, B. (2008): *Scientific Representation: Paradoxes of Perspective*, New York: Oxford University Press.

----- (2012): "Modeling and Measurement: The Criterion of Empirical Grounding", *Philosophy of Science*, **79**, pp. 773-84.

van Fraassen, B. and Peschard, I. (2018): *The Experimental Side of Modeling*, Minneapolis: University of Minnesota Press.

Walter, B. and Pohl, R. (1909): "Weitere Versuche über die Beugung der Röntgenstrahlen", *Annalen der Physik*, **25**, pp. 331-54.

Weisberg, M. (2013): *Simulation and Similarity: Using Models to Understand the World*, New York: Oxford University Press.

Wheaton, B. (1991): *The Tiger and the Shark: Empirical Roots of Wave-Particle Dualism*, New York: Cambridge University Press.

Woodward, J. (2000): "Data, Phenomena, and Reliability", *Philosophy of Science*, **67**, pp. S163-S179.