Beyond Numerical and Causal Accuracy: Expanding the Set of Justificational Criteria[1]

Jeffry L. Ramsey

University of Chicago

1. Introduction

Until recently, realists and anti-realists alike have assumed that any approximations which appear in explanations and confirmations in the mathematically oriented physical and biological sciences are "mere distractions" (Laymon 1989, p. 353). When approximation techniques must be used, they are typically justified by appeals to their numerical accuracy. However, recent interest in computational complexity in the sciences has revealed that numerical accuracy is not always the only criterion which should be invoked to justify the use of approximations. Cartwright (1983), Franklin (1988) and others have suggested that causal accuracy should be added as a criterion.

Numerical and causal accuracy are important. However, they form a sufficient set of justificational criteria only when certain—often unstated—conditions obtain. These conditions are actually assumptions which are often not satisfied, especially in the computationally complex situations which characterize discovery and exploratory contexts. In these contexts, we need to concentrate on "delineating a justificatory set which is relevant to [the] given epistemic situation" (Duran 1988, p. 273). In short, we need to understand when and why numerical and causal accuracy must be supplemented. In this paper, I begin developing this rationale by suggesting that the evaluation of theories developed with approximations requires at least two (if not more) additional criteria: range of validity and intelligibility.

My argument is anchored in an examination of the effects of approximative procedures used by chemical kineticists in the 1920s and 30s. The phrase "approximative procedure" includes any methodological strategy which is used to generate or interpolate a result due to underresolved data or deficits of analytic or calculational power. Chemical kineticists determine rates of chemical reactions in order to help them deduce the reaction mechanisms which lead from reactants to products. (More on this in Section 4.) Although chemical kinetics happens to be an ideal science for the study of approximative procedures and the effect they have on the justificational activities of scientists, I think similar considerations arise in other sciences including physics (cf. Dresden 1974, Cartwright 1983), some of the geological sciences, and evolutionary biology and ecology.

PSA 1990, Volume 1, pp. 485-499

2. Numerical Accuracy as the Criterion of Evaluation

Philosophers of science commonly explicate justification, whether of approximate or non-approximate theories, in terms of numerical accuracy or empirical adequacy. (In this paper, I will bypass the question whether numerical accuracy is the same as empirical adequacy and concentrate only on the former.) Most theories of confirmation "rest on the assumption that theories can be made to agree completely with the known facts, and . . . use the amount of agreement reached as a principle of evaluation" (Feyerabend 1975, p. 65). Explanatory realists contend that the consilience of theory and observation is a strong argument for their position. Although empiricists have long challenged realists on the conclusion which can or should be drawn from numerical accuracy, they have accepted the assumption that such accuracy is attainable and is the operative justificational criterion. Finally, the semantic view of theory structure, which can be given either a realist or anti-realist interpretation (van Fraassen 1972), epitomizes the desire to understand justification in terms of numerical accuracy. In their theory of approximations, the structuralists construct a numerical relation, ϵ, which assesses the amount of approximation between any two statements in a given language (cf. Balzer, Ulises-Moulines and Sneed 1987, Chap. 6).

Despite this emphasis on numerical accuracy, philosophers rarely discuss the conditions which allow it to function as a justificational criterion. To illustrate the conditions which I think are commonly assumed, I will examine briefly an argument by Kline and Matheson (1986). In response to Cartwright's (1983) attack on the truth of fundamental laws, Kline and Matheson advocate a model of explanation which "does not require that the description deduced from the explanans be an exact description of the event in need of explanation" (p. 39). Nonetheless, they emphasize numerical accuracy in their model developed from examples of gravitational phenomena (of planetary motion and balls dropped from towers). For instance, they require that "there are no other forces present that would produce an acceleration of magnitude f, *where f is very small.*" They conclude that "the trajectory of the body will be d(t) plus or minus e(t) *where e is the largest possible error* due to perturbing forces" (ibid., p. 35, my emphasis).

In these and other passages (cf. pp. 34-35), Kline and Matheson talk as if exactitude always implies numerical accuracy. However, they have conflated the precision of an explanatory prediction with its accuracy. Generally speaking, the relative precision, i.e. exactitude, of a prediction entails nothing about its relative accuracy—and thus its truth. Kline and Matheson's conflation of precision and accuracy is not harmful in the case they have chosen only because the following conditions obtain:

1) we have good reasons to believe that our theories (in this case, of gravitational phenomena) are true,

2) computational difficulties in the theory can be ignored except in special circumstances, and

3) the data (on the motions of dropped balls and the planets) is stable, determinate, accurate and precise.

I believe that most—if not all—accounts of justification in the mathematically oriented physical sciences rely on (1) through (3). We should be clear, however, that these conditions are actually assumptions which are not satisfied in all contexts. Sometimes they can be false, and sometimes we simply do not know whether they are true or false. When any one or more of them is not satisfied, the conflation of preci-

sion with accuracy is damaging in both predictive and explanatory contexts. In the next two sections, I examine instances where the conditions are not satisfied.

3. Challenges to Numerical Accuracy

Ronald Laymon (1989) discusses the consequences when assumptions (1) and (2) are not satisfied. He focuses on a case where computational difficulties manifest themselves as computational quirks in the theoretical laws. Three theories are ranked in increasing realism according to the slogan "the more (numerical) accuracy, the better." However, the structure of the laws is such that an increase in the sensitivity of initial conditions combined with an increase in computational precision produces predictions which behave non-monotonically. In Laymon's example, three scientists believe the following coupled equations express the correct laws:

$$ax_1 + x_2 = 1$$

$$0.5x_1 + 0.5x_2 = 1$$

The first scientist, who is restricted to three place precision, believes falsely that a = 0. Thus, $x_1 = 1.00$ and $x_2 = 1.00$. The second scientist, who is also restricted to three place precision, believes correctly that a = 0.001. If the equations are solved by Gaussian elimination beginning with variables of the lowest coefficient, then $x_1 = 0.00$ and $x_2 = 1.00$. The third scientist has access to five place precision and also correctly believes that a = 0.001. For this scientist, the same method of solution gives $x_1 = 1.0000$ and $x_2 = .99990$.

Thus, even though the second theory is more realistic than the first and less realistic than the third according to the slogan, the results of the calculations indicate that the second theory is worse off than the other two. The realist who relies on monotonic behavior as evidence for his or her position is in trouble since "the current example suggests . . . that a theory may be retained in the face of non-monotonic behavior on the ground that computational quirks are involved" (Laymon 1989, p. 372). While Laymon uses the example to defeat the realist, I think the example cuts just as strongly against the anti-realist. In Laymon's example, the instrumentalist would conclude the instrument is defective when, in fact, it may not be.

Laymon then drops assumption (1). He concludes that increased accuracy "may win out, but in the absence of known-true theories (and the requisite computational capabilities), we will not know when this will occur" (1989, p. 372). In other words, unless conditions (1) and (2) hold, increased accuracy does not guarantee anything *per se*.

While Laymon does not suggest in his paper what criterion should be substituted for or added to numerical accuracy if we are to make our predictions reliable, others—most notably Cartwright (1983)—have endorsed causal accuracy as an alternative criterion. In doing so, they rely heavily on condition (3). To a lesser extent, they rely on (1) and (2) as well.

Franklin (1988) relates a simple example which gets this point across. Suppose we have two theories which differ primarily in the fact that they accept different causal mechanisms. For instance, theory T_1 accepts the conservation of energy and theory T_2 accepts that energy is not conserved. For any given process of energy transformation, T_1 predicts that the energy difference for the before and after states is

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192726 Published online by Cambridge University Press

0 energy units. T_2 predicts that the energy difference is 10 energy units. The experimental value for the energy difference is found to be 2 energy units. "It would seem that on any reasonable view T_2 is closer to the truth even though T_1 is numerically more accurate" (Franklin 1988, p. 528). *Attending to causal accuracy alone*, we are more likely to accept the theory T_2 as closer to the truth since it allows the existence of causal processes which can explain our experimental result.

The example is counterintuitive since Franklin asks us to endorse a causal story which controverts our currently held beliefs. I think it gains whatever force it has because he relies on conditions (1), (2) and (3). Most importantly, I think he relies on (3), the existence of an experimental result which is certain. In his derivation of the conclusion using Bayes theorem (cf. his fn. 2 on p. 528), he does not consider the option that the numerical value may be mistaken. Likewise, in his historical example, the value for the advance of the perihelion of Mercury is assumed to be known determinately and accurately. (Of course, this is because the value *was* known determinately and accurately.) Franklin has chosen examples where the stability and resolution of the observational data are not in question and transferred this condition to his imaginary example. In contrast, I think most people would question the results of the experiment before they adopted the new causal story which endorses the non-conservation of energy.

On a side note, I think an example similar to Laymon's could be used to argue against Franklin's assumption that causal accuracy is a reliable indicator of good explanations. To do so, one would simply have to assume that causal hypotheses are embedded in the three laws of Laymon's example.

4. Augmenting the Challenge

I conclude that assumptions (1) through (3) may be nice idealizations, but they are inappropriate in many or most cutting-edge research areas in science. What other qualities of justificational arguments can be invoked to generate confidence in theoretical or experimental results when other combinations of assumptions (1) - (3) are not satisfied? In the cases I have studied in this kind of situation, range of validity and intelligibility arguments are often deployed to fill the gap.

Before examining these arguments, I should set the stage by saying a few words about the tasks of chemical kinetics. Chemical kinetics is the branch of physical chemistry which "deals with the rates of chemical processes and how the rates depend on factors such as concentrations, temperature, and pressure. The ultimate objective of kinetic investigations is to gain information about the mechanisms of chemical reactions" (Laidler 1987, p. 74). A reaction mechanism is a series of elementary reactions detailing the postulated individual steps of the reaction pathway. To a good approximation, To a good approximation, the rate of a reaction is given by the Arrhenius equation $k = Ae^{-E/RT}$ where k is the rate, A is the pre-exponential factor and E is the activation energy. Theories of kinetics attempt to provide explanations for the A and E factors.

In chemical kinetics during the 1920s and 30s, the available theories were not known to be true, calculational difficulties were rampant, and the experimental data was unstable and underresolved. Sir Cyril Hinshelwood, who won a Nobel prize for his work in chemical kinetics, noted that the theories available in the 1920s and 30s involved calculations which were possible in a "rough and ready manner only." "Rough guesses" were needed to assign "vague" and "arbitrary" values to some of the theoretical parameters (Hinshelwood 1951, p. 384). Louis Kassel, a prominent chemical kineticist, remarked, "On the experimental side much confusion has been pro-

duced by the complex interdependence of great masses of data. . . . It is distressing to have a dozen respected mechanisms collapse when a single reaction is more carefully studied" (Kassel 1932, p. 7).

In the 1920s and 30s, theoretically inclined chemical kineticists tended to focus on the following classes of reactions:

1) $H_2 + H$ and its various deuterium (D) analogues

2) $H_2 + X$, $H_2 + X_2$ where X is a halogen, i.e. fluorine (F), chlorine (Cl), bromine (Br) or iodine (I)

3) $M + X$, $M + X_2$ where M is an alkali metal such as sodium (Na), potassium (K), cesium (Cs), etc. and X is halogen

4) $Na + RX$ where RX is an organic halide such as CH_3X, C_2H_5X, C_6H_5X, etc.

Of all these reactions (and thousands more which were not of primary interest to theoreticians), only the mechanism for the $H_2 + I_2$ reaction, a member of class (2), was thought to be known with any degree of certainty. The only activation energy which could be calculated theoretically with any rigor was the energy for the $H_2 + H$ reaction, and this value was only within an order of magnitude of the spread of values allowed by the experimental evidence. This combination of computational difficulties and lack of necessary data hampered justificational efforts in the discipline.

It might be natural to suggest these classes of reactions involve only a change in initial conditions under a covering law. However, even if this is the case, it was not apparent to chemical kineticists in the 1920s and 30s. More importantly, it does not seem to be the case. The change in the atoms and molecules is a material variation which affects the processes and mechanisms involved in chemical change. As such, the theories, concepts and equations which apply to one class are not *a priori* guaranteed to apply to any others. In general, talk about the mechanisms was not easily translatable into talk about laws (cf. Wimsatt 1976 and Cartwright 1983 for much the same point). To supplement the rough numerical correspondence between predictions and data, kineticists appealed to range of validity and intelligibility considerations to argue for the validity of their concepts and equations.

Range of Validity: In my first example, assumption (2) is defeated due to the presence of equations which cannot be solved (rather than, as in Laymon's example, the presence of computational quirks). Further, assumption (1) is not satisfied *because* (2) is not. (Laymon does not link them in this causal fashion in his example.) In other words, drastic approximative procedures are applied to one theory to overcome calculational difficulties; this process produces a new theory which is not known to be true. When this happens, it may be that errors introduced by successive approximations partially or wholly cancel out each other. Thus, instances of causal and numerical accuracy do not, by themselves, rule out the possibility that the instances are isolated since the cancellation may not occur when the approximations are applied to other cases. They must be supplemented with an argument that the causal and numerical results have a range of validity.

By range of validity, I mean—roughly—the applicability of a theory, law or concept across a variety of classes of phenomena. This property has been called scope by other authors (cf. Giere 1983). In the absence of any one or more of the conditions

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192726 Published online by Cambridge University Press

which allows accuracy to be the operative justificational criterion, a range of validity argument is important because it is a claim that the relevant law, theory or concept has uncovered a pattern which may have a possible causal basis in the data. In the exploratory phases of an investigation where it is not taken for granted that the available laws, theories and concepts are the appropriate ones to employ, uncovering a pattern is an important explanatory achievement (Campbell 1966, Simon 1977). For instance, it is an indication that the principle "similar causes, similar effects" may be applicable. In the example I discuss, the pattern turned out to be spurious, but this was not apparent for a number of years.

Sir Cyril Hinshelwood (1897-1967) developed his 'simple' collision theory to calculate the rates of reactions prior to their use in determining mechanisms. The theory was not known to be true because some forbiddingly difficult calculations could not be performed. "The complete theoretical investigation of reactions . . . would be quite hopeless. . . . Even if the mechanics of a particular collision could be treated rigorously, a fivefold integration, by quadratures, would be needed to obtain the actual reaction rate" (Kassel 1932. p. 48). Hinshelwood thought he could produce an argument that there were no computational quirks in the theoretical apparatus he set up in place of the rigorous solution. (More on this in a moment.) Further, for the purposes of the 'simple' theory, the experimental data was adequately stable and resolved.

To calculate the rates of reactions, Hinshelwood borrowed the concepts and some of the equations of the statistical mechanical theory of gases. He postulated that $k = Ze^{-E/RT}$ where k is the rate, Z the number of collisions in the gas, and E the combined energy of the reacting molecules. Thus, the pre-exponential factor A of the Arrhenius equation was identified with the number of collisions in the gas. The theory was originally developed to interpret the experimental evidence for a small set of bimolecular gaseous reactions, most of which belonged to class (2) above (McLewis 1918, Hinshelwood 1926). The experimental rates for these reactions, which could be calculated precisely using a phenomenological equation, compared favorably to the theoretical values. Hinshelwood used this consilience to infer the correctness of the causal mechanism postulated in the theory, viz. the exchange of translational kinetic energy *via* a collision of the reactant molecules.

Unfortunately, the collision theory failed to be either numerically or causally accurate when applied to reactions other than bimolecular gaseous reactions. For example, when applied to the reactions in class (3)—in particular to the reaction of sodium (Na) vapor with the halides—the theory overestimated the rate of reaction by a factor of between 7 and 30 (Meer and Polanyi 1932; Beutler and Rabinowitsch 1930). When applied to the reactions in class (4) above, the collision theory underestimated the rate by a factor which ranged from 10^1 to 10^6 (Meer and Polanyi 1932). In short, Hinshelwood's inference about the correctness of the collision theory was premature because the theory had no numerical range of validity.

In addition, the theory possessed no causal range of validity. Since the theory postulated that a successful exchange of energy was a function of the kinetic energy of translation only, energy exchanges could occur only when the molecules were within the collision theoretic diameter. However, for simpler reactions such as that between ions and molecules, the exchange of energy occurred at distances far in excess of the kinetic theory diameters (Kallman and London 1929). In more complex reactions, the efficiency of the energy exchange was much less than that predicted by the collision theory. In these reactions a successful collision turned out to be a function of factors which were not included in the collision theory (Eyring 1935; Evans and Polanyi 1937).

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192726 Published online by Cambridge University Press

In order to account for the numerical and causal inaccuracies, Hinshelwood used an approximative procedure which he called the steric factor, P. He transformed the original collision theoretic equation, $k = Ze^{-E/RT}$, into $k = PZe^{-E/RT}$. The introduction of the steric factor was an approximative procedure which allowed him to adjust theoretical and experimental results so they were in rough numerical agreement. It was used only to bring the results to within an order of magnitude of each other; a higher degree of precision was deemed unnecessary. In effect, the steric factor was a curve fitting parameter with a putative causal interpretation.

At one point, Hinshelwood included no fewer than five physical considerations in P: the approximate fractions of the two molecular surfaces which could come into contact with each other, the collision number, the rate at which energy could find its way into the correct part of the molecule, the rate at which the preactivation of the molecules was destroyed, and an unknown function which linked the preactivation energy and the number of degrees of freedom active in the molecule (Hinshelwood 1935). These mechanisms were postulated to explain any deviations from what he continued to believe was the primary mechanism, the transfer of translational kinetic energy via a collisional encounter. Hinshelwood eventually admitted that these causal mechanisms could not be tested and accepted that the collision theory was only a useful interpolation formula (Hinshelwood and Winkler 1937).

Hinshelwood used the fact that the value of the steric factor changed predictably and gradually over a wide series of reactions to argue that there were no computational quirks embedded in the theory. P ranged from 1 to 10^{-8} over a series of increasingly bulkier and more complex molecules; the PZ and E values were nicely correlated for these reactions, i.e. E was correspondingly greater when the PZ values were larger (Moelwyn-Hughes 1933, Fairclough and Hinshelwood 1937). Although the value of P could not be predicted *a priori*, it could be rationalized once the experimental data had been determined. This rationalization indicated that the collision theory had a wide range of validity (Hinshelwood and Winkler 1937).

For my purposes the important points are: 1) the accuracy of the 'simple' theory in one set of instances did not guarantee that it would be numerically or causally accurate in other instances, and 2) the range of validity argument was used to argue against the presence of any computational quirks. To generate a range of validity for his 'simple' collision theory, Hinshelwood had to employ the steric factor. To be sure, he intended the steric factor to be interpreted causally. However, without an account of the actual operation of the causes, he could only keep the collision theory testable by conjoining numerical correspondence with an argument that P varied in a qualitatively predictive fashion. In other words, the steric factor looked like it was providing an intelligible—but as yet causally undetermined—explanation of the data. This leads me to my next example.

Intelligibility Arguments: I move now to a discussion of an example where an intelligibility argument provided a strong justificational warrant for a theory since causal and numerical accuracy were unattainable. In the 1920s and 30s, kineticists argued for some of their theories and data by employing something similar to what Einstein later called an intelligibility argument. This kind of argument claims that some concepts and theories make "intelligible what is sensorially given" without being derived from the senses (Einstein 1949, p. 669). To speak very roughly and metaphorically, an intelligibility argument imposes a map on an unknown terrain rather than constructing the map from the known terrain.

My example is the determination of the activation energies and reaction mechanisms of class (1), the reactions between molecular and atomic hydrogen and deuterium. These reactions were mechanistically simple and, in addition, were the only ones for which the theoretical calculations could be performed with any rigor. There are two somewhat separate parts to the story: the calculation of activation energies from potential energy surface diagrams and the experimental determination of the activation energy. Chemical kineticists attempted to get the two parts in some relation of agreement with each other; this proved more difficult than many had anticipated.

Because of the formidable calculational difficulties of solving the many-body Schrödinger equation, chemical kineticists employed a number of drastic approximative procedures to generate potential energy surfaces (see Eyring and Polanyi 1931, Coolidge and James 1934, Hirschfelder and Wigner 1939). The approximations were so drastic that the resultant equations were not known to be true. The best calculations on the simplest system, the $H_2 + H$ reaction, were only within an order of magnitude of the experimental result. While the theoretical value was around 13 kcal (Polanyi and Eyring 1931), the best available experimental value (calculated with a phenomenological equation) was between 4 and 11 kcal in 1930 (Farkas 1930) and 5 and 7 kcal by 1935 (Geib and Harteck 1935). (The total experimental error in the value for the energy was, respectively, 1% and 3%.) Clearly, numerical accuracy could play a partial but not determinate role in judging whether the calculations were reliable. Since no theory existed which was known to be true, the order-of-magnitude theoretical calculation could function only as an indication that the theory was headed in the right direction.

While kineticists believed that potential energy considerations did determine the behavior of molecules, they acknowledged that the gap between the surfaces represented in the calculations and those in actual reaction situations was quite large (Farkas and Wigner 1937, Hirschfelder 1941). In addition, no-one had ever seen or touched a potential energy surface, so they certainly were not derived from the senses. Finally and importantly, the surfaces got some of the causal story incorrect because of the approximative procedures used in their construction. Speaking of the $H_2 + H$ reaction, Kassel noted, "It is unfortunate that the approximation . . . seems to break down rather badly at small distances, leading to the prediction of H_3 with a very compact structure" (Kassel 1932, p. 57). In sum, the surfaces provided an intelligible picture which provided behavioral clues but was not intended to be a strict causal interpretation of what was happening in the actual system.

In addition to the computational difficulties which made the truth of the theory unknown, an additional complicating factor in this instance was the lack of necessary experimental data which might help validate the theoretical calculations. In the $H_2 + H$ reaction, experimenters measured changes in pressure, a measure of concentration changes for gases, to determine the rate constant. Once the rate constant was determined, the activation energy could be calculated using a phenomenological equation. However, due to limitations in experimental technique, a number of macro-level causal stories were possible. (At the micro-level, viz. the utilization and interaction of the different molecular energy levels, the causal story was left open since it was simply not relevant to interpreting the causal interactions of the molecules.)

Farkas (1930) used pressure dependence measurements to determine the order of the $H_2 + H$ reaction. On the basis of these measurements, he postulated two possible mechanisms:

$$H + H + H \qquad \text{and} \qquad H_2 + H$$

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192726 Published online by Cambridge University Press

He ruled out the first mechanism because the probability of a three atom collision was exceedingly small and focused his attention on the second mechanism.

Unfortunately, he could not calculate the activation energy with much precision. The experimental rate constant was dependent on both temperature and pressure. These dependencies could only be separated by incorporating the pressure dependence into the equation phenomenologically. However, this forced Farkas to determine the hydrogen atom concentration, [H], in the system. This could not be measured experimentally, so Farkas resorted to three different causal hypotheses about the relation between the measured hydrogen molecule concentration, [H_2], and the unknown hydrogen atom concentration, [H]. The different hypotheses resulted in values between 4 and 11 kcal. The mechanisms were not derived from the experimental setup; they were causal stories which Farkas thought were plausible.

In 1935, Geib and Harteck recalculated the values for the activation energy using an experimental method which greatly improved the precision of the measurements (Geib and Harteck 1935). Even with the improved data, they—like Farkas—had to supply causal interpretations to generate values for the activation energy. Using different causal stories, they generated values of 5 to 7 kcal. Although Geib and Harteck's values were more precise and probably more accurate than Farkas', they were further away from Eyring and Polanyi's (1931) approximate theoretical value. However, the effect on the credibility of the theoretical calculation was slight. Eyring and Polanyi intended their value to be an order of magnitude calculation since the theoretical equations could be solved only by employing experimental values in theoretical equations.

In addition, the theoretical calculations made good predictions for the relative rates of change of the analogues of the hydrogen molecule-hydrogen atom reaction: H_2 + D, HD + D, HD + H, D_2 + H, and D_2 + D (Farkas and Wigner 1937). As with the H_2 + H reaction, the potential energy surfaces were not causally accurate since the values needed to tell a causal story were not known "with sufficient accuracy" (p. 718). Likewise, the numerical accuracy in each instance was no better than it was for the H_2 + H reaction. Even further, the predictions on the analogue systems had no range of validity: the "lack of more experimental and theoretical data . . . makes it impossible for us at present further to check [the theoretical] formula for a wide range of H_2 and D_2 concentrations, and also for different temperatures" (p. 720). However, the relative changes in the qualitative behavior were matched by the relative changes in the rough theoretical estimates. In short, the theoretical method made the behavior of the experimental data intelligible even if discrepancies "between experimental data and the theoretical rate must be ascribed . . . either to errors in the experiments or to some imperfection of the theoretical treatment" (p. 723).

The lesson I wish to draw from this example is two-fold. First, faced with underdetermined and underresolved experimental data, experimentalists had to postulate plausible causal stories which were not derived from the data. They used the coincidence of the values to argue that the actual mechanism or mechanisms was close to the ones they postulated, but it is important to note that none of the values was derived directly from the experimental data. Thus, placing the restriction that our explanatory stories must arrive at *the* correct causal story seems too stringent at the exploratory or discovery phase of investigation since causal stories often have to be supplied by the researchers. Often, a plethora of possibilities helps get the investigation going. Second, the intelligibility argument for the concept of a potential energy surface carried the day even though numerical accuracy was lacking and the causal story implied by the surface was either underdetermined or wrong. The experimental data was too uncertain to warrant throwing out the theoretical calculation, but the theoreti-

cal calculations were too compelling to support the judgement that the experimental values were inaccurate. This was especially true given the qualitative prediction of the rates of the analogue reactions.

This brings me to an additional, more general point. In kinetics, one does not attempt to prove a causal story as the story. "An investigation, kinetic or otherwise, can disprove a proposed mechanism but cannot establish a mechanism with absolute certainty" (Laidler 1987, p. 75). This is not just Humean, Cartesian or inductivist "in principle" skepticism. The mapping of possible mechanisms to the available data is always many to one. Today, when we can get a fair amount of detailed information about the micro-causal state of affairs, we can eliminate a large number of causal pathways. However, the mapping from the micro mechanisms to the (relatively) macroscopic reaction mechanism is still a difficult task. In the 1920s and 30s, the information on the micro-causal stories was not available, so the mapping was even more indeterminate. The intelligibility argument runs in tandem with the causal accuracy arguments to produce a causally consistent story. Population genetics and ecology often face a similar plethora of possible alternative mechanisms (cf. Gould and Lewontin 1978).

More generally, then, theoretical interpretations can be intelligible without necessarily being an exact causal representation or mirroring of the phenomena. There is a difference between "making sense of" and "representing." After all, a representation can be wrong but still make partial sense of experience. Further, although my endorsement of an argument from intelligibility may sound too idealistic, I would argue that the idealism inherent in the criterion is not noxious. The intelligibility criterion involves an idealism which takes seriously the claim that experience is difficult to interpret and that it takes a lot of work to make our ideas and experience consistent.

5. Towards an Expanded Descriptive Set of Justificational Criteria

I wish to reiterate that I believe numerical and causal accuracy are important justificational criteria. However, the presumption that they are the *only* criteria is a presumption about certain ends and conditions for scientific justification. That end is "complete agreement between theory and data," and the conditions are: (1) the presence of known-true theories, (2) the absence of computational difficulties, and (3) the presence of stable and resolved data.

(3) seems to be more commonly assumed than the others; among the authors I have discussed, only Laymon explicitly excludes situations in which the necessary data is not available (cf. Laymon 1985, p. 147). However, data can be absent, too imprecise to be useful, give conflicting messages, etc. Problems with the data are not limited to the examples I have cited; Hedges (1987) has shown that the amount of disagreement about the character of the data is roughly the same in both the "hard" and the "soft" sciences.

I think assumptions (1) through (3) are not generally satisfied, especially in exploratory and discovery contexts. At the very least, they are not warranted in the determination of rates and mechanisms of chemical reactions in the 1920s and 30s. When conditions (1) - (3) are not satisfied, other descriptive qualities are invoked as justifications for theories and concepts. In the 1920s and 30s, intelligibility and range of validity arguments were used together with numerical and causal accuracy arguments to motivate the acceptance of theories and concepts developed with approximative procedures. Kuhn (1977) has suggested that scientists use a number of such qualities as criteria to guide theory choice. In contrast, I have been concerned with how the criteria are used to generate opinion about the acceptability of a theory, concept or

law independently of the choice between theories. [This is not an illegitimate abstraction. Alternative versions of the collision theory did not part with Hinshelwood on any important issues until the late 1930s (cf. Fowler 1929 and 1937) and no alternatives existed for the potential energy surface calculations for many years.]

For the kineticists, all four justificational qualities worked in tandem since assumptions (1) through (3) were not satisfied. It is certainly possible that range of validity and intelligibility arguments could be superfluous when the conditions in the assumptions are met. However, note that even though Hinshelwood thought the collision theory satisfied all three conditions, he still relied on a range of validity argument to motivate acceptance for the theory. Thus, a range of validity argument seems to be an important and somewhat independent justificational criterion. More importantly however, I would argue that the conditions which might make the arguments superfluous are often not met. In addition, kineticists were trying to tell a causally consistent story; this fact shaped their use of numerical and causal accuracy arguments in significant ways.

If we accept that other qualities are used by scientists in certain cases, I think it becomes apparent that both realists and anti-realists place too many restrictions on scientific justifications. To pick only one of many possible examples, consider van Fraassen's dictum that "acceptance of a [scientific] theory involves as belief only that it is empirically adequate" (van Fraassen 1980, p.12). Apart from the fact that he buys into a corollary of condition (3), viz. that empirical adequacy can be demonstrated, the dictum is too rigid if "empirically adequate" is interpreted as "numerically accurate." There are at least two justificational qualities which contribute to belief by going beyond empirical or numerical adequacy. van Fraassen's theory of truth has led him to process scientific justifications through a particular epistemological grinder.

To my mind, the question becomes why both realists and anti-realists have touted predictive success and causal accuracy as the criteria of justificational arguments. I think assumptions (1) through (3) underlie that motivation. When they are not satisfied, attempts to make the theories agree with the known (causal or numerical) facts and to use the amount of agreement as a principle of evaluation seem misguided. My intent is not to advocate a third ontological position but simply to point out that the traditional formulations of realism and anti-realism are incomplete when confronted with actual justificatory practice.

If we take seriously the fact that not all science satisfies the assumptions, then the criterion of agreement with accurate (numerical or causal) representation becomes only one possible end of justification. This has been pointed out by authors concerned with models. We can use models in an ad hoc fashion, as a data-ordering device, as a computational device, or as a claim about reality (Wartofsky 1979). Or, we can use specific knowledge about the falsity of a model to improve it in later versions (Wimsatt 1987). In short, to presuppose that the only important part of the scientific game is to get the theory and experiment to agree to a specified degree of numerical or causal accuracy is to place too many restrictions on the game. At times, we simply want to explore, and in these instances the justificational game is altered accordingly.

How often is the justificational game altered to include criteria like range of validity and intelligibility? This is an interesting empirical question which deserves further investigation. What causes the shift out into the other explanatory criteria? While I cannot provide a complete answer to this question, I can sketch some considerations.

When assumptions (1) - (3) are satisfied, the focus of explanation seems to lie properly at the level of the theoretical laws. However, when they are not satisfied, the focus properly broadens to include theories, concepts, the stability of the data, and the laws. For instance, worry about the accuracy of fundamental laws only makes sense when the empirical laws governing the behavior of the phenomena are well known. Until that time, the systematization of data into classes is an important explanatory achievement in itself. Thus, range of validity and intelligibility arguments seem important in pattern induction. Scientists use them to see what concepts are needed in the empirical and theoretical laws and to explore how and if the concepts are related.

In addition, the move to an expanded justificatory set is conditioned by changing material conditions. In these cases, we are trying to do more than discover the appropriate initial conditions. We are also trying to discover if the equations which govern one set of phenomena govern another. We proceed by analogy to see if the processes and mechanisms in the new instances are similar or different to the old instances.

Finally, I would like to suggest that the shift occurs due to the nature of some sciences. In sciences that generate causal consistency rather than causal accuracy arguments, intelligibility and range of validity arguments are essential since the mapping between the experimental data and the possible causal stories is one to many. This is true in evolutionary biology just as much as it is chemical kinetics.

5. Conclusion

In this paper, I have argued that numerical and causal accuracy form a sufficient set of justificatory criteria only when: the theories in use are known to be true, computational difficulties do not exist, and the experimental data is stable and resolved. When any one or more of these assumptions is not satisfied, additional justificational considerations in the form of range of validity and intelligibility arguments must be invoked. I illustrated these claims by examining some cases in chemical kinetics. My arguments suggest that the presupposition common to realists and anti-realists alike that numerical or causal accuracy constitutes the whole of justification is unsound. Finally, I have sketched some suggestions why range of validity and intelligibility arguments are invoked as justificatory criteria in such situations.

Notes

[1]Many thanks to Bill Wimsatt and Dan Garber for helping me get clearer on the argument of this paper. I would also like to thank R. Stephen Berry and Jack Halpern for helping me understand the chemistry.

References

Balzer, W., Ulises-Moulines, C. and Sneed, J. (1987), *An Architectonic for Science: The Structuralist Program.* Boston: D. Reidel.

Beutler, H. and Rabinowitsch, E. (1930), "Über die Beziehungen zwischen Rotation, Warmetönung und Wirkungsquerschnitt der Reaktion mit chemischen Elementarprozessen", Zeitschrift für physikalische Chemie B8: 231-254.

Campbell, D. (1966), "Pattern Matching as An Essential in Distal Knowing", in *The Psychology of Egon Brunswick*, K. R. Hammond (ed.). New York: Holt, Rinehart and Winston, pp. 81-106..

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

Coolidge, A. S. and James, H. (1934), "The Approximations Involved in Calculations of Atomic Interaction and Activation Energies", *Journal of Chemical Physics* 2: 811-817.

Dresden, M. (1974), "Reflections on 'Fundamentality and Complexity'", in *Physical Reality and Mathematical Description*, C. Enz and J. Mehra (eds.). Dordrecht: D. Reidel, pp. 133-166.

Duran, J. (1988), "Discussion: Causal Reference and Epistemic Justification", *Philosophy of Science* 55: 272-279.

Einstein, A. (1949), "Remarks to the Essays Appearing in this Collective Volumes", in *Albert Einstein: Philosopher-Scientist*, P. A. Schilpp (ed.). LaSalle, IL: Open Court, pp. 663-688.

Evans, M. G. and Polanyi, M. (1937), "Inertia and Driving Force of Chemical Reactions", *Transactions of the Faraday Society* 34: 11-23.

Eyring, H. (1935), "The Activated Complex in Chemical Reactions", *J. Chem. Physics* 3: 107-115.

Eyring, H. and Polanyi, M. (1931), "Über einfache Gasreaktionen", *Z. f. physik. Chemie* B12: 279-311.

Fairclough, R. A. and C. Hinshelwood (1937), "The Functional Relation between the Constants of the Arrhenius Equation", *Journal of the Chemical Society* 1937: 538-546.

Farkas, A. (1930), "Über die Thermische Parawasserstoffumwandlung", *Z. f. physik. Chemie* B10: 419-433.

Farkas, L. and E. Wigner (1936), "Calculation of the Rate of Elementary Reactions of Light and Heavy Hydrogen", *Transactions of the Faraday Society* 32: 708-723.

Feyerabend, P. (1975), *Against Method*. Norfolk: Thetford Press.

Franklin, A. (1988), "How Nancy Cartwright Tells the Truth", *British Journal for the Philosophy of Science* 39: 527-529.

Fowler, R. (1929), *Statistical Mechanics*. Cambridge: Cambridge University Press.

_____. (1937), *Statistical Mechanics*, 2nd edition. Cambridge: Cambridge University Press.

Geib, K. and Harteck, P. (1931), "Die Einwirkung von atomaren auf molekularen Wasserstoff", *Z. f. physik. Chemie, Bodenstein-Festband*: 849-862.

Giere, R. (1983), "Testing Theoretical Hypotheses", in *Testing Scientific Theories*, Earman, J. (ed.). Minneapolis: University of Minnesota Press, pp. 269-298.

Gould, S. and Lewontin, R. (1978), "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptionist Programme", *Proc. R. Soc. (London)* 205: 581-598.

Hedges, L. V. (1987), "How Hard is Hard Science, How Soft is Soft Science?", *American Psychologist* 42: 443-455.

Hinshelwood, C. (1926), *Kinetics of Chemical Change in Gaseous Systems*. Oxford: Clarendon Press.

_____. (1935), "Considerations Concerning the Mechanism of Chemical Reactions", *Journal of the Chemical Society* 1935: 1111-1115.

_____. (1951), *The Structure of Physical Chemistry*. Oxford: Clarendon Press.

Hinshelwood, C. and Winkler, C. A. (1936), "On Bimolecular Reactions in Solution", *J. Chem. Soc.* 1936: 371-377.

Hirschfelder, J. and Wigner, E. (1939), "Some Quantum Mechanical Considerations in the Theory of Reactions Involving an Activation Energy", *J. Chem. Phys.* 7: 616-628.

Kallman, H. and F. London (1929), "Über quantenmechanische Energieübertragung zwischen atomaren Systemen", *Z. für physik. Chemie* 2B: 207.

Kassel, L. (1932), *The Kinetics of Homogeneous Gas Reactions*. New York: Chemical Catalog Co.

Kline, A. D. and Matheson, C. (1986), "How the Laws of Physics Don't Even Fib", in *PSA 1986*, vol. 1, A. Fine and P. Machamer (eds.). East Lansing: Philosophy of Science Association, pp. 33-41.

Laidler, K.J. (1987), "Kinetics (Chemistry)", in *Encyclopedia of Physical Science and Technology*, Vol. 7. New York: Academic Press, pp. 74-97.

Laymon, R. (1985), "Idealizations and the Testing of Theories by Experimentation", in *Observation, Experiment and Hypothesis in Modern Physical Science*, Achinstein, P. and O. Hannaway (eds.). Boston: MIT Press, pp. 147-173.

_____. (1989), "Cartwright and the Lying Laws of Physics", *Journal of Philosophy* 86, 353-372.

Lewis, W. C. McC. (1918), "Studies in Catalysis. Part IX. The Calculation in Absolute Measure of Velocity Constants and Equlibrium constants in Gaseous Systems", *J. Chem. Soc.* 113: 471-492.

Meer, N. and Polany, M. (1932), "Vergleich der Natriumdampreaktion mit anderen organish-chemischen Prozessen", *Z. f. physik. Chem.* 19B: 164-189.

Moelwyn-Hughes, E. A. (1933), Kinetics of Reactions in Solution. Oxford: Clarendon Press.

Simon, H. (1977), "On Judging the Plausibility of Theories", in *Models of Discovery*. Boston: D. Reidel, pp. 25-45.

van Fraassen, B. (1972), "A Formal Approach to the Philosophy of Science", in *Paradigms and Paradoxes*, R. Colodny (ed.). Pittsburgh: University of Pittsburgh Press, pp. 303-366.

_____. (1980), *The Scientific Image*. Oxford: Clarendon Press.

Wartofsky, M. (1979), "Models, Metaphysics and the Vagaries of Empiricism", in *Models: Representation and the Scientific Understanding*. Boston: D. Reidel, pp. 24-39.

Wimsatt, W. (1976), "Reductive Explanation: A Functional Account", in *PSA 1974*, R. S. Cohen et. al. (eds.). Dordrecht: D. Reidel, pp. 653-670.

_____. (1987), "False Models as Means to Truer Theories", in *Neutral Models in Biology*, M. Nitecki and A. Hoffman (eds.). New York: Oxford University Press, pp. 23-55.