THOMAS KUHN'S THEORY OF RATIONALITY

PAULO PIROZELLI
https://orcid.org/0000-0002-4714-287X
Federal University of Santa Catarina
Department of Philosophy
Florianópolis, SC
Brazil
paulopirozelli@gmail.com

Article info
CDD: 501
Received: 04.09.2019; Revised: 11.09.2019; Accepted: 11.09.2019
DOI: http://dx.doi.org/10.1590/0100-6045.2019.V42N3.PP

Keywords
Thomas Kuhn
Philosophy of Science
Rationality
Values

Abstract: According to a widespread view, Thomas Kuhn's model of scientific development would relegate rationality to a second plane, openly flirting with irrationalist positions. The intent of this article is to clarify this aspect of his thinking and refute this common interpretation. I begin by analysing the nature of values in Kuhn's model and how they are connected to rationality. For Kuhn, a theory is chosen rationally when: i) the evaluation is based on values characteristic of science; ii) a theory is considered better the more it manifests these values; and iii) the scientist chooses the best-evaluated theory. The second part of this article deals with the thesis of the variability of values. According to Kuhn, the examples through which epistemic values are presented vary for each person, and for this reason individuals interpret these criteria differently. Consequently, two scientists, using the same values, can come to a rational disagreement over which theory to choose. Finally, I point out the limitations of this notion of rationality for the explanation

of consensus formation, and the corresponding demand for a sociological theory that reconnects individual rationality with convergence of opinions.

INTRODUCTION

Thomas Kuhn's *The Structure of Scientific Revolutions* (*SSR*) generated a heated debate over the importance of rationality to science, since its publication in 1962. For many of his critics (and including some of his admirers), Kuhn's model of scientific development would minimize the role of rationality, while instead giving prominence to non-epistemic forces. A more radical line of criticism accused him of denying the relevance of reasons in scientists' making a choice in favor of a particular theory. His explanation of the resolution of controversies, these critics say, would be restricted to social and psychological factors only, thus making the reference to rationality dispensable (Shapere 1964; Scheffler 1982; Lakatos 1970; Laudan 1978, 1985; and for some of his supporters, see Bloor 1991; Barnes & Bloor 1982).

Kuhn rebutted these accusations on many occasions, stressing the importance of rationality to his analysis (1970b, 1970c, 1977d). Nonetheless, there are reasons for the existence of so many misinterpretations on this matter: Kuhn does not present a detailed and systematic discussion of the concept of rationality in his writings. The purpose of this article is to fill this gap, developing a theory of rationality compatible with Kuhn's main ideas on topics such as the process of choosing theories, values, and scientific pedagogy.

A better grasp of Kuhn's notion of rationality will make it easier to understand why he rejected accusations of irrationalism as being profoundly misguided. Furthermore, Kuhn's notion of rationality is intertwined with his perspective on the nature of science and scientific behavior:

the considerations used by scientists in choosing theories, the sources of divergence in opinion and the possibility of rational disagreement. Lastly, this examination will make clear the limitations of rationality to produce widespread agreement within the community, and the need of social mechanisms in order to explain consensus formation.

As in every philosophy of science, it is in the process of theory-choice that the nature and importance of rationality arises more urgently.[1] For this reason, this article starts with a discussion about the role of epistemic values for the selection of theories, and how they allow for the definition of an implicit notion of rational choice. In the second part, I will discuss the structure of scientific pedagogy, and the impact it has to the application of values by individuals. Lastly, a few observations will be made concerning the

---

[1] I am here making two assumptions. The first, which I consider less problematic, is that periods of scientific controversy display more clearly the role and nature of rationality. That is why it is more illuminating to study rationality in these contexts. Second, I am implicitly assuming that the the same kind of rationality is operating in all stages of scientific development. It would be possible to think that, since Kuhn speaks of two types of theoretical change, cumulative and non-cumulative, there would be accordingly two different types of rationality, intra- and extra-paradigmatic (for this position, see De Langhe 2012). However, it is important to note that, in any stage of science, scientists are judging how well a certain accretion or modification of the accepted body of beliefs fits according to values shared by the community. Whether this change implies the substitution of a paradigm (or taxonomy) or not, it is a different, unrelated matter, irrelevant to the assessment of the rationality of the modification. This unified theory of rationality seems to me one of the consequences of the evolutionary perspective exposed by Kuhn in (1991a).

problems arising from these conceptions — particularly, the explanation of how consensus is formed.

A GENERAL THEORY OF RATIONALITY

Values

Kuhn's concept of rationality is directly tied to that of value. For him, a rational choice is simply a choice based on the values that characterize a certain activity. Therefore, in order to understand his notion of rationality, we must first analyze his ideas on the nature of values.[2]

By definition, values are what someone uses to evaluate something: "a certain thing *T* is evaluated by the subject *S* according to a set of values *V*". One way to make this relationship clearer is to treat this definition as a mathematical function. *T* is the set of objects under consideration (in our case, scientific theories), $T = \{t_1, t_2, \dots, t_m\}$, or the domain of the function; $X = \{x_1, x_2, \dots, x_m\}$ is the outcome of the evaluations, or the image of T; and *v* is the function of evaluation, relating each theory t_i to a particular appreciation x_i. Formally, we have that

$$v(t_i) = x_i$$

[2] The references to values in Kuhn's writings are generally brief (e.g. 1962a, 1970a, 1970c; the exception is 1977d). Consequently, the absence of discussion on the subject could put some doubt on my claim that scientific values are indispensable for a proper understanding of his ideas on theory-choice and rationality. But as Kuhn himself assures, although his "work has been little concerned with the specification of scientific values, [...] it has from the start presupposed their existence and role" (1977b, pp. xxl-xxli).

in which x_i is the evaluation of the theory t_i according to value *v*. We further arbitrarily define the codomain as

$$x \in [0,1]$$

with 0 representing the minimum evaluation, and 1, the maximum.

The next step is trying to define the relation between values and rational choices. According to Lacey (2004), three aspects define *v* as being a value of an activity φ for individual *X*. First, *v* is something that *X* wants to achieve in φ. Second, *v* is partially constitutive of a good φ, helping to stipulate its standard of excellence; in other words, *v* indicates when realizations in φ are good or bad. Finally, everything else remaining the same, it is always better that a value be manifested to a greater degree (the *ceteris paribus* clause).

There is a considerable difference, though, between Kuhn's and Lacey's views on values. According to the latter, the manifestation of *v* is not the main goal of activity φ and does not suffice to define it. A simple example may help to illustrate this point. The employee of a car company may embrace a set of values, such as honesty and dedication, that dictate to a large extent many of his attitudes while he is working. His goal, nonetheless, is not that of being honest or dedicated, but producing cars. For him, honesty is at most a secondary goal. That is why Lacey considers values as being only partly constitutive of a good φ: they define good and bad achievements for an activity, but do not define a practice *per se*. They allow for distinguishing a good car worker from a bad one, but not a car worker from a teacher.

Kuhn believes that such a conception is based on a profound misunderstanding regarding the nature of values. Instead, Kuhn defends a view of the role of scientific criteria, similar to Hempel's "quasi-trivial" approach (Hempel 1981, 1983), that intends to solve both problems at once.

According to Kuhn, science is a social practice that is distinct from other activities, such as Art and Philosophy, because it possesses certain characteristic traits, such as accuracy and precision (1983d, pp. 212-13; McMullin 1993, pp. 65-66). If these features are what characterize science, then it is redundant to say that these features are pursued because they lead to the fulfillment of some specific goals: producing theories in which they are manifest is simply, by definition, to make science (1993a, pp. 251-52). Put differently, values are not the *means* to achieve independent goals in science, but the *goals* of science themselves. Contrary to what we find in Lacey, values are for Kuhn *totally constitutive* of a good φ: practicing an activity is synonymous with pursuing the achievement of its socially determined values.

With this, Kuhn avoids two traditional epistemological problems at once. First, that of finding the goals of science. Those, for Kuhn, are simply the traits that differentiate science from other activities in societies. The epistemological problem of determining the ends of science thus becomes a matter for sociology. A second difficulty, the problem of finding the values that would promote the aims of science more than any others, is naturally dissolved by this quasi-trivial approach. There is no sense in looking for values capable of promoting objectives of science, because the values and objectives of science are one and the same thing.

Values and Rationality

Science is, for Kuhn, an activity guided by multiple values (1977d; see note 3 below for more references). This means that there is not just a single function of evaluation v but instead a set of functions v_i, one for each element of the set $V = \{v_1, v_2, \dots, v_n\}$. Therefore, evaluation is more properly represented as a system of equations, formally expressed as

$$\begin{Bmatrix} v_1(t) = x_1 \\ v_2(t) = x_2 \\ \dots \\ v_n(t) = x_n \end{Bmatrix}$$

in which x_i represents the result of applying the value v_i to theory t. Also, we can define arbitrarily that

$$v_i \in [0,1]$$

As stated before, values are what define an activity as such, and therefore, serve as parameters for evaluating achievements in a certain activity (in our case, scientific theories). Generalizing this idea to the set *V*, we can define the evaluation of a theory as a function of the various constitutive values of science. In a formal way, we have that

(1) $$f\big(v_1(t), v_2(t), \dots, v_n(t)\big) = y$$

in which *y* represents the outcome of the evaluation of theory *t* from the various values v_i. Arbitrarily, we can stipulate that

$$y \in [0,1]$$

We also know from the definition presented by Lacey (2004) that it is always better to have more of a value than less. That is,

If $v_k(t_1) > v_k(t_2)$,

then
$f\big(v_1(t_1), v_2(t_1), \dots, v_n(t_1)\big) > f\big(v_1(t_2), v_2(t_2), \dots, v_n(t_2)\big)$,

As long as

$$v_i(t_1) \geq v_i(t_2), \forall v_i \in V$$

Notice that $v(t)$ and $f(t)$ are supposed to be continuous functions, assuming any value in the specified range. Thus, given the restriction of the *ceteribus paribus* clause, the same idea can be rewritten as the following derivative:

(2) $$\frac{\partial f}{\partial v_i} > 0, \forall v_i \in V$$

Lastly, a choice can be considered rational when the scientist chooses the most well-evaluated theory. Suppose the elaboration of increasingly consistent theories is one of the defining features of science. Then, a scientist who opts for a less consistent theory, all other aspects remaining fixed, would be frontally violating the basic patterns of the activity she believes to be practicing (1983d, p. 209). This kind of "self-defeating action" is, according to Kuhn, the "surest index of irrationality" (1993a, p. 252). Formally, we have that the choice of a scientist is rational in the case where

(3) If $f\big(v_1(t_1), v_2(t_1), \dots, v_n(t_1)\big) > f\big(v_1(t_2), v_2(t_2), \dots, v_n(t_2)\big)$,

then

$t_1 \succ t_2$,

in which the symbol $\succ$ represents the choice relation.

Points (1)-(3) constitute what I call the *minimum criterion of rationality* (the first two describing the evaluation of theories, and the third, their choice). Although relatively vague, these axioms give a more precise outline to Kuhn's notion of rationality. A choice is rational when it meets the following requirements: theories are evaluated through values characteristic of scientific activity; the greater the manifestation of a value, the better the theory is considered;

and among two theories that differ in one or more values, the choice must fall on the one that exhibits them to a greater degree (as long as this theory is not worse in any aspect). This minimum criterion of rationality has also the advantage of treating the evaluation of theories as an intrinsically comparative process, something strongly defended by Kuhn (1962a, p. 146; 1991a; 1992).[3]

Nonetheless, these axioms have a very limited scope of application: the comparison of theories is restricted to the *ceteris paribus* clause. The minimum criterion of rationality gives no indication about the case in which a theory is better than another according to some values and worse according to others. But before addressing this issue, I will briefly speak about a particular type of value — scientific values, which, after all, are those that directly concern Kuhn's investigations.

Scientific Values

Kuhn does not present consistent observations regarding the precise values that constitute scientific activity. But we can obtain some suggestions from his writings — particularly from (1977d), where he examines the nature and role of values in a more detailed way. In this paper, Kuhn suggests five basic criteria that define a good theory: accuracy, consistency, scope, simplicity, and fruitfulness (1977d, p. 322).[4]

---

[3] The most obvious implication of the view that theory-choice is an essentially comparative activity is the elimination of any concerns about the truth of scientific theories (1991a, pp. 95-96).

[4] Kuhn's list of criteria changes from paper to paper. He mentions, for example: "simplicity, precision, and congruence with the

It is important to have in mind that this enumeration is more illustrative than exhaustive of scientific values. Kuhn has no intention of yielding a definitive description of the criteria actually employed by scientists. It is the general mechanism of the process of theory-choice, rather than its specificities, that concerns him (1977d, p. 321). For that, relatively accurate general observations about the typical behavior of the scientist are enough to sketch the problem. It is the task of sociologists and historians of science, not philosophers, to provide more precise descriptions of the values involved in the choice of theories.

Another reason that explains Kuhn's lack of concern for rigorously identifying the constituent values of a good theory is that these criteria are learned in practice, rather than through theoretical elaborations. Thus, it is unlikely that precise formulations can be given for elements whose transmission and application is largely tacit. Scientists themselves would probably have difficulty in verbally expressing the precise structure of their judgments.

Finally, there is a third aspect that further hampers any attempt to provide an exhaustive list of values (at least from the perspective of an author with a leaning for the history of science, such as Kuhn). It is the fact that these values change from time to time and from group to group. Although values such as precision and simplicity, when viewed in a very broad way, are permanent traits of science, their concrete

---

theories used in other specialties" (1970a, p. 21); "accuracy, scope, simplicity, fruitfulness, and the like" (1970b, p. 261); "accuracy, simplicity, fruitfulness, and the like" (1970c, p. 198); "accuracy, precision, scope, simplicity, fruitfulness, consistency, and so on" (1993a, p. 251). Broader presentations of values are given in: 1962a, p. 42; 1970a, p. 21; and 1970c, p. 184. See Hoyningen-Huene (1993, p. 149), for a complete list of references.

understandings vary greatly over time and through scientific communities (1977d, p. 335).

Whatever the difficulties in defining these values are, they all share one important characteristic: they are all *cognitive* or *epistemic* values. Broadly, I mean by this that these criteria serve to evaluate a theory regarding its capacity for producing knowledge (however this is understood — the capacity to solve puzzles, answer problems, broaden our understanding of nature, save the phenomena, or some other formulation). For Kuhn, cognitive values as precision, consistency, comprehensiveness, simplicity, and fecundity provide "*the* shared basis for theory choice" (1977d, p. 322). Moral or social considerations, on the contrary, have a small or non-existent influence on assessments. The scientist described by Kuhn is what Kitcher calls a "pure epistemic agent": one "for whom the primary goal is to reach an epistemically valuable state" (1993, p. 308).

Kuhn has good reasons to believe that non-cognitive values are mostly absent from scientists' evaluations. Epistemic values, as seen, define scientific disciplines, socially and semantically. It is expected then that scientists base their choices almost exclusively on these criteria, since appealing to them is part of the language-game of science (1977d, pp. 336-37). As Kuhn explains, scientists' decisions "may not be merely personal but must instead be accepted as solutions by many" (1962a, p. 167). The demand for public justification subordinates the personal interests of the scientist to epistemic considerations (Longino 1990, p. 76; Van Fraassen 2006, p. 28).

This conclusion seems surprising, considering the quite common reading of *SSR* as one of the major exponents of an externalist tradition in history and philosophy of science. Nonetheless, such a reading, I believe, is the most consistent with Kuhn's ideas.[5] Of course, even from an internalist perspective, some caveats have to be made. First, values should be understood as reasons, and not as motivations or causes of scientists' actions. They do not have to be actually present in the mind of the individuals when a choice is made, but need only to be presented as justifications when demanded, be it during social exchanges or in written publications (Glock 1998, entry "causation").[6] Moreover, epistemic values should not be understood as universal rules. Scientists can disregard such criteria and often do so. That does not necessarily invalidate Kuhn's observations, since his aim is simply to present a general model capable of accurately describing the motivations that are behind scientists' choices most of the time.

Also, as I hope to make clear in the second part of this article, the idea that scientists choose theories (almost always) based exclusively on cognitive values does not imply that agents are not subject to external pressures and influences. However, non-cognitive factors are not part of the evaluation itself. Instead, they act indirectly, determining the interpretation of values. Ultimately, we will see, this produces a gap between individual rationality and communitarian consensus.

---

[5] Not only in spirit, but also in its letter, *SSR* is mostly an internalist description of science. At least, according to Kuhn himself (1962a, p. xliv; 1997, pp. 287-88).

[6] The fact that scientists need to argue extensively and deeply (and often, for a long time) is what makes it unlikely that these values will be thought of as *mere* rationalizations. And even if, for a specific individual, these criteria are indeed rationalizations, carries no real force in her choice, the social demand for public justifications gives these values a real power in convincing other people during scientific debates.

Expectations

The claim that scientific theories are evaluated solely through cognitive values is not a perfect representation of Kuhn's position, though — and not for the reasons mentioned above. According to him, epistemic values, despite being the main element shaping evaluations, do not constitute "the unique or an unequivocal basis for paradigm choice" (1962a: 168). Scientists' considerations, he believes, are not restricted to what theories can accomplish at the moment of decision. Scientists are concerned also with the results they think theories might produce in subsequent research (1962, pp. 156-57, 200).

The idea that a belief in the future ability of theories to produce knowledge is a key element in scientists' choices was the root of many objections and criticisms of *SSR* (e.g. Scheffler 1982; Shapere 1964; Lakatos 1970; Laudan 1978). The notion of "faith," that Kuhn made use of in *SSR*, seemed to impose, just as that of "conversion", a mystical-religious, irrational character to decision-making. There is no doubt that this word choice caused interpretative difficulties that resulted in persistent criticisms. However, the idea that agents partially base their theory-choices on predictions about the future should not sound strange or uncommon. After all, decisions involving prospective considerations are consistently studied in several branches of the social sciences, such as economics and political science. In most of these fields, however, such a faith in future events is called "expectation."[7]

[7] Kuhn speaks of "expectations" in *SSR*, but in a different sense. There, it means the type of empirical and theoretical responses scientists are conditioned to expect from paradigms (anomalies, for example, are nothing more than failures of expectations (1962a, p.

Some of the sources for these expectations in the future performance of theories are analyzed by Kuhn. One of them, for example, is the subjective perception that a theory possesses some kind of intrinsic harmony or beauty — the theory is "neater," "more suitable" or "simpler" than its competitors (1962a, p. 154). Kuhn also mentions things such as religious propensities, personality, and professional trajectory (1962a, pp. 151-52) as possible causes of confidence in particular theories.

But not all expectations are fruit of personal and arbitrary elements such as these. Two other, more standard, factors involved in scientists' prospective evaluations are discussed by Laudan (1978, pp. 106-08).[8] They are both inductions that predict the future achievements of a theory based on its track record. The first of these element is the *general progress* of a theory, its growth in cognitive capacity over time. In a formal way, we could represent this as

$$\frac{\partial f_s\left(\partial v_{s,1}(t), \partial v_{s,2}(t), \dots, \partial v_{s,n}(t)\right)}{\partial s}$$

in which *s* represents a particular instant in time; $v_{s,i}(t)$ the value of theory *t* according to v_i in time *s*; and f_s the final evaluation of theory *t* in time *s*.

The second kind of measure considered by Laudan is the *rate of progress*. This is the acceleration or deceleration of a theory's problem-solving ability, or, in short, the variation of

xliii)). In this article, I use "expectation" as meaning how well a scientist thinks a theory will do in the future.

[8] Laudan (1978) is in fact discussing the measures for evaluating research traditions. I consider, however, that his ideas, when slightly modified, illuminate important aspects of the scientists' expectations.

progress. In an analogous way, it can be understood as the second derivative of the evaluation function for the variable time, or

$$\frac{\partial^2 f_s\left(\partial^2 v_{s,1}(t), \partial^2 v_{s,2}(t), \dots, \partial^2 v_{s,n}(t)\right)}{\partial^2 s}$$

Bringing this all together, we can detect three broad causes of expectations: the past successes of the theory; its recent performance; and the further, subjective motivations described by Kuhn. Thus, expectation regarding the problem-solving capacity of a theory in the future, y^e, can be formalized as

$$y^e = h\big(p_s(t), z_s(t), \dots, m(t)\big)$$

in which *p* is the total progress of theory *t* until time *s*; *z*, the rate of progress of theory *t* until time *s*; and *m*, the support that further motivations give to theory *t*.[9] Arbitrarily, we can define that

$$y^e \in [0,1].$$

For Kuhn, these expectations are one of the components of evaluation, alongside cognitive value-based appreciation. So, an evaluative function that includes all these factors can be (compactly) written as

$$F_s = (v_{s,1}(t), \dots, v_{s,n}(t), y^e) = w_z$$

[9] I did not use a parameter for time in the case of *m*, because it is not obvious that these subjective motivations are dependent on time as are the first two parameters. However, nothing would change if time was also considered here.

in which F_s is the outcome of evaluation at time *s*, and arbitrarily defined as

$$w_z \in [0,1].$$

Undoubtedly, the introduction of expectations as an element present in the evaluations of scientists brings considerable difficulties. Two problems are particularly pressing. First, can we still talk about rational choices when next to the traditional epistemic values (accuracy, simplicity, etc.), scientists also choose theories based on what they think the theory will do in the future? The discussion of different causes of expectations aims to balance this problem. Certainly, there is no justification for expectations *stricto sensus*: induction remains an unsolved (or unsolvable?) problem. But that does not mean that all kinds of expectations are equally reasonable. In particular, estimates based on the prior performance of a theory seem quite different and much more acceptable than expectations originating from less palpable aspects, such as aesthetic preferences and religious inclinations. Scientists can, for instance, debate about the theories' merits and compare their evolution over time. Further, ultimately, all these estimates, even subjective ones, have an *epistemic* concern: they are all concerned with the theories' capacity for producing knowledge. So, in this sense, expectations are similar to traditional cognitive values.

A second problem is related to the coexistence of expectations with properly epistemic values. It is obvious that, with expectation, a theory that is currently not the most adequate can be chosen (Laudan 1981, p. 152). This seems to make cognitive values dispensable, for in an extreme case, high expectations could make any theory preferable. Kuhn does not deny that possibility: scientists might indeed choose a theory solely for the hope they place on its future success, while assigning little or no weight to its current effectiveness. In fact, Kuhn thinks this is not only a theoretical possibility, but a situation that once in a while occurs in science. Einstein's rejection of quantum mechanics is a proof, for him, that even the greatest scientists may be unmoved by the current results demonstrated by theories (1977d, p. 337). Nonetheless, despite being a recurring historical phenomenon, Kuhn considers that epistemic values tend to be preponderant upon scientists' choices. Generally, he says, only a few scientists tend to be enthusiastic about a paradigm before "hardheaded arguments can be produced and multiplied" (1962a, p. 157; see also 1969c, pp. 342-43). The impact of extravagant expectations on the number of scientists who accept a theory is in practice quite limited — the number of scientists who have high expectations is usually small.

Aggregation of Values

I mentioned above a considerable limitation of this notion of rationality. The comparison of theories is limited to the case in which the *ceteris paribus* clause can be applied. That is, a theory is considered better than another when it is at least as good in every respect and better in at least one.

This is not a problem in simple situations. Take as an example, a scientist who has to choose among three theories, while making use of five epistemic criteria — accuracy, scope, simplicity, fruitfulness and consistency. After some examination, she ends up with the following ordering of the theories according to each value (Table 1):[10]

---

[10] I had defined the outcome of evaluation as a real number. For the case I am treating here, however, the only important thing is

	Accuracy	Scope	Simplicity	Fruitfulness	Consistency
Theory 1	1º	1º	1º	1º	1º
Theory 2	2º	2º	2º	2º	2º
Theory 3	3º	3º	3º	3º	3º

Theory t_1 seems the obvious choice here, since it is superior in all aspects. t_1 is what D'Agostino (2003, p. 13) calls "dominant".

Unfortunately, this is a limiting case. In more complex situations, values may, and probably will, differ in their respective orderings. Consider, for instance, the following evaluation (Table 2):

	Accuracy	Scope	Simplicity	Fruitfulness	Consistency
Theory 1	1º	2º	3º	3º	1º
Theory 2	2º	1º	1º	1º	2º
Theory 3	3º	3º	2º	2º	3º

Which theory should a scientist choose now? Given that no theory is better over all, the current criterion of rationality does not favor any of the alternatives. According to D'Agostino (2003, p. 91ff.), two main strategies could be adopted in order to cope with this limitation. The first strategy is the eliminationist one. As the name suggests, it

their relative positions. For the sake of simplicity, therefore, I opted to use ordinal numbers.

involves removing some of the multiple dimensions of the problem, and reducing the number of available theories or values. This brings the situation back to the simplest case, in which one of the theories is at least as good as any other theory in every aspect. So, for example, if the criteria of scope, simplicity and consistency were removed from Table 2, t_1 would become again dominant over its competitors.

The problem with the eliminationist strategy, as D'Agostino (2003, p. 29) warns, is that it involves a serious "moral risk." Elimination requires discarding a value or theory that, in principle, should have been taken into account. By implementing this solution, we may end up with theories that "are not as good, all relevant matters considered, as they might have been if we had considered them in their full complexity."[11]

Fortunately, the eliminationist strategy does not account for all possible types of resolutions of conflicts among values. A second type of approach described by D'Agostino involves, instead, the aggregation of values. According to him, this aggregation can be done through the establishment of a "rate of substitution": a measure that determines the increase in a value a person is willing to accept in exchange for a decrease in another (2003, p. 28).

The rate of substitution can be better understood if we consider its relation to the choice function, $f(t)$. All of the

---

[11] Strictly speaking, the eliminationist strategy does not violate the first axiom of the minimum criterion of rationality, which assumes in advance a set of theories and values. It is possible to say, though, that it undermines a meta-criterion of science, that of trying to produce better descriptions and explanations of the world. (D'Agostino 2003, p. 94) discusses contexts in which this strategy may be valid, such as when the adoption of a new theory involves high costs.

points at which $f(t)$ has the same value y form, by definition, a level curve. The substitution rate is simply the set of all the modifications of the values v_i that produce a movement along this level curve. It represents how much one is willing to exchange of one value for another, so that the overall assessment remains the same. Another way of considering the substitution rate is as the solution set to the following equation:

(4) $\frac{\partial f}{\partial v_1} dv_1 + \frac{\partial f}{\partial v_2} dv_2 + \cdots + \frac{\partial f}{\partial v_n} dv_n = 0$

in which dv_1 is the amount of variation in the value v_i, or

$$\Delta v_i = v_i(t_1) - v_i(t_2)$$

To simplify, let us suppose that all the values are fixed, except for v_1 and v_2. Then, rearranging (4), we have that

$$\frac{dv_1}{dv_2} = -\frac{\frac{\partial f}{\partial v_2}}{\frac{\partial f}{\partial v_1}}$$

By the second axiom of rationality, (2), this implies that

$$\frac{\frac{\partial f}{\partial v_2}}{\frac{\partial f}{\partial v_1}} < 0$$

The rate of substitution is therefore a negative number. As we would intuitively expect, a decrease in one value is only

acceptable when offset by the appropriate increase of another one.[12]
An alternative way of thinking about this method of aggregation is as an assignment of weight. This is, in effect, the expression used by Kuhn (1977d, p. 324). The idea is that each value has a specific contribution to the overall assessment, indicating its relevance to the scientist's final choice (see also D'Agostino 2003, p. ch. 1). Mathematically, the weight m_i is the derivative of $f(t)$ with respect to v_i. This can be expressed as

$$m_i = \frac{\partial f(\partial v_1(t), \partial v_2(t), \ldots, \partial v_n(t))}{\partial v_i(t)}$$

Because of (2), the weight is always a positive number.

The substitution rate and the weight are directly connected: a substitution rate is simply the ratio of different weights. Hence, it is the establishment, either of the values' weights, or of substitution rates between values (which, after all, are different ways of expressing the same relations), that allows for the aggregation of multiple criteria.

To simplify, we will make a linearity assumption about the weights assigned to the values. This means that we will consider the weight of each value as a constant. Treating them as a constant seems to be, in fact, the most intuitive notion of weight, which refers to a weighted arithmetic mean. With this, the first axiom of rationality, (1), becomes simply

(5) $\sum_{i=1}^{n} m_i v_i(t) = y$

---

[12] More precisely, the substitution rate is simply the tangent line to a specific point of a particular level curve.

in which, by definition, the sum of weights is equal to 1, i.e.

$$\sum_{i=1}^{n} m_i = 1$$

It is also necessarily the case that $m_i > 0, \forall m_i$ since if $m_i = 0$, the value in question would not be taken into account (it would be a value that it is not a value).[13] The third axiom (3) becomes therefore,

$$\sum_{i=1}^{n} m_i v_i(t_1) > \sum_{i=1}^{n} m_i v_i(t_2) \leftrightarrow t_1 \succ t_2$$

Finally, defining $\boldsymbol{m}$ as the matrix derivative of weights m_i,

$$\boldsymbol{m} = \begin{pmatrix} \frac{\partial f(\partial v_1(t), \partial v_2(t), \dots, \partial v_n(t))}{\partial v_1(t)} \\ \frac{\partial f(\partial v_1(t), \partial v_2(t), \dots, \partial v_n(t))}{\partial v_2(t)} \\ \dots \\ \frac{\partial f(\partial v_1(t), \partial v_2(t), \dots, \partial v_n(t))}{\partial v_n(t)} \end{pmatrix}$$

[13] We could also consider the possibility of values with negative weight: things that scientists despise. However, they are already included, in a certain sense, among the values. The contrary of any negative value is a positive values. For example, if inconsistency is something that scientists want to avoid, it means that consistency is a value for them.

and $\boldsymbol{\Delta v}$ as the variation in the values v_i,

$$\Delta \boldsymbol{v} = \begin{pmatrix} v_1(t_1) - v_1(t_2) \\ v_2(t_1) - v_2(t_2) \\ \dots \\ v_n(t_1) - v_n(t_2) \end{pmatrix}$$

we have that

$$(\boldsymbol{m}^T . \Delta \boldsymbol{v} > 0) \leftrightarrow t_1 \succ t_2.$$

It is possible now to give an answer to the situation in which the application of values differs in the ordering of theories. The aggregation of values allows a scientist to rationally choose a theory even when values point in different directions. A choice is rational as long as the scientist selects the theory which is superior to all values considered (even if it is inferior in the realization of certain particular criteria). Provided that the gain in some of the values compensates the loss in other criteria, it is rational to choose that theory. The remarks presented above show how the aggregation procedure can be incorporated so as to make the minimum criteria of rationality applicable to a broader set of cases.

AN INDIVIDUAL THEORY OF RATIONALITY

The first part of this article sought to systematize Kuhn's main ideas concerning rationality. Scientists are said to be acting rationally when they choose the theory which is epistemically superior to the available alternatives (available at the time of decision and, hopefully, in the future), according to the values accepted by the community. Except for concerns regarding expectation in the future success of the theory, Kuhn's observations do not differ much from the ideas that preceded the historiographical revolution of the 1960s.

What explains then the criticisms directed at his remarks on theory-choice (e.g., Shapere 1964; Scheffler 1982; Laudan 1978), as well as Kuhn's recurrent defenses of his theses against the accusations of irrationalism and relativism (1977d; 1970b, p. 259-66; 1970c, sec. 2, 5)? In order to understand how Kuhn's apparently traditional theory of rationality could be seen as having such radical consequences, we need to examine his remarks on the acquisition of values.

Scientific Pedagogy

As noted by Mody & Kaiser (2008, p. 378), pedagogy is a central analytical category in Kuhn's work, who often emphasizes the "special nature of scientific education" (1970c, p. 208; see also 1977d, p. 327). The pedagogy explains, for him, numerous aspects of scientists' behavior and scientific practice, including the nature of epistemic criteria.[14]

According to Kuhn, values are transmitted to students by more experienced members of the community, already familiarized with the paradigms currently accepted in the

[14] According to Laudan (1985, p. 286), "far more than most writers on these subjects, he [Kuhn] has tended to stress the importance of community and socialization processes in understanding scientific enterprise." Other important aspects of science explained by Kuhn through pedagogy are: the acquisition of paradigms, the practice of normal science, the selection of research problems, the acquisition of language, the impossibility of full communication among communities of experts (semantic incommensurability), and the standard cumulative view of scientific development.

field. Through theoretical and practical examples, these older practicians show them how to evaluate theories, indicating why some may be considered superior to others. Later, with time, the community's values are adopted by the apprentice (1970b, p. 275).

However, as Kuhn stresses, the examples presented in this process are seldom accurate descriptions of actual conflicts (1977d, p. 328; see 1962a, ch. 11, for the role of textbooks in shaping scientists' perspective of the history of science). Instead, they are stylized and biased expositions that simplify real choice situations, presenting them in as unproblematic. The qualities of the winning theory tend to be highlighted, its weakness underplayed, and the strengths of the rejected alternatives diminished or simply ignored. Strictly speaking, these situations do not truly recount past scientific controversies. Students are simply learning what it means for one theory to be better than another. The theoretical justifications of theories, found in textbooks, are, to a large extent, just pedagogical instruments for teaching theories and values.

This pedagogical simplification has important consequences for the application of values. The cases presented to students provide immediate choices only because they are erected from idealizations made with didactical purposes. Things get more complicated, though, when scientists need to apply these same values to concrete cases in which there is no pre-established response — cases, as Kuhn says, in which "there are always at least some good reasons for each possible choice" (1977d, p. 328).

When facing these unprecedented scenarios, scientists then have to extend the use of values of textbook applications to new problem-situations. It is here that learning exerts its force. In spite of the force and strictness of scientific pedagogy, the way a scientist understands epistemic values is ultimately dependent on her particular experiences. Consciously or not, her previous experience in the field, the examples and theories to which she has been exposed, education, personality, cultural context, and propensity to take risks, are the material on which a scientist bases her interpretation to extend the use of values to other cases.[15] Idiosyncrasies like that are the repository that feeds an individual's view on scientific criteria (1977d, p. 325, 329).[16] For this reason, the application of values to new cases inevitably varies from individual to individual (1970b, p. 262; 1962a, p. 4).

Kuhn classifies the elements influencing the interpretations given by scientists to epistemic values in three groups: the previous experience of an individual as a scientist, factors outside of science, and personality (1977d, p. 325; Hoyningen-Huene 1993, pp. 150-51). But he does not examine these classes in any more detail. There are two reasons that explain his attitude. The first is the infinitude of possible causes, that would prevent any complete analysis. In theory, any personal experience can contribute to shape the

[15] Kuhn talk of "interpretation" without necessarily connecting the term to propositional contents. He means by it the process through which an individual grasps or introjects a shared value.

[16] Two observations need to be made. Kuhn does not use "interpretation" as necessarily involving a conscious reflection on values. For him, interpretation means simply the way an individual scientist applies the communitarian values that, in principle, are the same for all people. Second, it is important to note that idiosyncratic factors determine the evaluation formulas, and not choices themselves. The latter depends, not on social factors, but on the degree that a theory manifests epistemic values. Sometimes, Kuhn himself seems to have made the mistake of thinking that social factors directly determine a scientist's choice, when they only affect it indirectly, through shaping his evaluative formula (see, for example, 1977d, p. 329).

interpretation of values. A second reason resembles that which led Kuhn to avoid discussions on the precise set of scientific values. For his intentions — understanding the nature of values and its consequences to theory choice and rationality — it is enough to analyze the implications of pedagogy to the application of values.

This does not imply that he considers all the factors involved in the interpretation of values as being of equal relevance. In *SSR*, for example, Kuhn emphasizes the time and status in the field of research as the most decisive element for explaining the position a person adopts in a controversy. This is the famous "Planck's Principle": older and more experienced scientists tend to be more resistant to revolutionary changes, while new scientists, eager to create a name of their own, naturally lean to new alternatives (1962a, p. 150).

Beyond Objectivity and Subjectivity

The idea that an individual's prior experiences contribute to determine the evaluation of scientific theories may lead to some misunderstandings. The most common is thinking that an individual's evaluation formula is "determined by a mix of cognitive and social factors" (Argamakova 2019: 47; see also Laudan 1985: 283; Wray 2011: 160-64).

Kuhn, however, expresses his discomfort in expressing things in this manner (1977d, p. 325). In the first place, the notion of "subjectivity" applies poorly to these idiosyncrasies. Like any other social phenomenon, the elements that constitute the formulas of evaluation may, in theory, be identified and described (1970c, p. 191). Neither individual factors replace epistemic criteria in choices, substituting "actual facts" for "bias and personal likes or dislikes" (1977d, p. 357). Being science an activity governed by a series of characteristic values, scientists must always expose their decisions in a manner consistent with this disciplinary vocabulary. Moreover, "objective" values are as much socially determined as "subjective" elements: they are not universally shared, but are part instead of a group's culture and practice.

Nonetheless, there is a third, and more insidious, conception about the subjective nature of the personal and professional factors shaping a scientist's assessment. One may think that these elements would distort the impartial use of objective criteria, modifying the application of values in their pure form.

Trying to formalize this view, $f_u(t)$ would be an objective formula of evaluation. It would be based solely on a pure application of epistemic values. The formulas nurtured by individual scientists, by its turn, would be a mixture of the objective application of values with personal factors, such that for every scientist $\{j\} \in J$ (where J is the set of scientists in a community),

$$f_j(t) = f_u(t) + \varepsilon$$

in which ε would represent the influence of subjective factors, or the "error" in relation to the objective evaluation.

No philosopher or historian would deny that, at least sometimes, things like individual prejudices or institutional affiliations play a role in scientists' actual choices. The relevant matter is the role assigned to these elements. For a certain philosophical tradition, which we might call objectivist, there would be "always an objective ordering of the available theories" (Worrall 1990, p. 332) — even if unknown to scientists. Personal factors, at best, would serve to fill the gaps in the knowledge of the objective evaluation formula. They would allow scientists to choose between theories in the beginning of controversies — periods in which "it was less apparent which theory was superior" (Wray 2011, pp. 160-64). In most of the cases, however, these particularities would simply distort the objective judgment of the scientists. Then, "as evidence increases with the passage of time", we would expect "the algorithms of different individuals converge to the algorithm of objective choice" (Kuhn 1977d, p. 329).

Contrary to that, Kuhn sustains that "there is no neutral algorithm for theory-choice" (1970c, p. 198), and no formula like $f_u(t)$ exists. A number of arguments is provided by him to deny the existence of a general, impersonal application of epistemic criteria. First, Kuhn considers the repeated failures to find this objective formula as pointing to the implausibility of the project (1977d, p. 325ff.).

He then makes a genealogy of the social and psychological sources of this conception. For him, the responsibility for such a belief in the existence of a universal algorithm of choice is due to distortions caused by textbook literature (1962a, ch. 11). By assigning a role to crucial experiments that they have not had historically, presenting solely the arguments of the winning theory, textbook historiography gives the impression that there are clear and definitive procedures, valid in all times, for justifying and evaluating theories. Kuhn also considers that the objectivist conception is based on the false assumption that the production of consensus necessarily indicates an approximation of the formulas of evaluation — ignoring that scientists may accept the same theory for different reasons (1977d, pp. 328-29; see also D'Agostino 2010).

Another argument given by Kuhn is based on the nature of values, discussed below. Values are transmitted through the display of paradigmatic cases. So, when scientists have to deal with real choices, they are urged to extend the use of these criteria to more complex situations. In this case, the resulting application is based on the only resource they posses: their previous experiences. Thus, the very idea of a universal evaluation formula, not needing a person producing it from the extrapolation of situations she has have had contact with, would show itself incoherent.

A final argument provided by Kuhn is normative in character. This is the so-called risk-dispersion argument. According to Kuhn, cognitive diversity have epistemological advantages over homogeneity. A community whose members can rationally disagree in their evaluations allows several paths to be pursued simultaneously. This raises the chances of someone solving the problems and consequently reduces the risk that the research, as a whole, fails. On the contrary, in a community whose members possessed an objective algorithm of choice, all individuals would choose the exact same theory. If just one route is explored, the failure of the project becomes more likely. Hence, the effectiveness of science in producing of knowledge is better explained by Kuhn's hypothesis that scientists differ in their assessments of theories (1970b, pp. 241, 248-49, 262; 1970c, pp. 185-86; 1977d, p. 332; see also D'Agostino 2005, 2010; Kitcher 1993, ch. 8).

"Objectivity" and "subjectivity", in sum, are not the best words to describe Kuhn's ideas on values and rational choice. Scientists' evaluations are indeed dependent on factors that vary for each person. But considering their choices as subjective can obscure the fact that decisions need to be based on epistemic criteria shared by other members of the community. At the same time, if we say that the resulting evaluation is objective, then objectivity must be seen as allowing some degree of individual rational disagreement.

Instead of using the notions of "objective" and "subjective", Kuhn prefers to contrast values with rules. By a "rule", he understand an algorithm of choice, a set of procedures that leads every person to the same conclusions. A good example would be a computer algorithm, in which the same inputs always give the same outputs. A value, on the contrary, may be seen as a kind of "incomplete rule" (1977d, p. 331): something that influence choice, but is not as specific as to determine a single result. Using an analogy with the computer algorithm for the same inputs, the value would produce similar outputs, although not quite the same.

Kuhn is defending that scientific values are not very different of the concepts and terms we employ in natural language. Most of the words we use are not determined by rules, and do not have a fixed meaning (Wittgenstein 2009, 79§). Our concepts of "chair" or "couch", for example, are not supposed to work in every conceivable situation. Two people looking at a uniquely designed furniture would have a great chance of disagreeing whether it is a chair or not. However, we would hardly say they do not know what a chair is, or that their concept of "chair" involves an objective concept of chair couples with some subjective elements. We naturally assume that the use of concepts is open to a certain amount of divergence.

Epistemic values should be seen, for Kuhn, in the same perspective. Most of the time, scientists agree about their application. However, in some cases, divergences may arise. This happens because values do not determine a unique course of action. Instead, they provide general guidelines for scientists, highlighting the essential issues involved in the decision process and pointing "to the remaining aspects of the decision for which each individual must take responsibility himself" (1977d, p. 330). Despite being critically important to scientific behavior (scientists would behave very differently if they had other or no values), the nature of values makes them open to divergences in application.[17]

Individual Formulas of Evaluation

The variability in the interpretation of values has direct implication for the notion of rationality. Record the general parameters that define the choice of a theory as being rational: the scientist evaluates available theories according to the values accepted by the community; prefers theories that manifest these criteria in a higher degree; and chooses the theory that is superior overall. When broadly conceived, these postulates are valid for every individual. Nonetheless, since values may be interpreted differently, their particular instantiations will vary from person to person.

Kuhn indicates two main ways in which individuals' evaluations could differ (1977d; this was already present in 1970c, sec. 3; see also Laudan 1985). First, scientists may disagree on how to apply specific values to concrete cases (1977d, p. 322; Laudan 1985, p. 284; D'Agostino 2005, p. 202). Kuhn mentions a couple of examples of scientific

[17] For Daston (2016), Kuhn would be constantly struggling with the idea that theory-choice is a "no-rule-governed reasoning". For one side, "the fertility of Kuhn's paradigms for scientific practice lay precisely in their being indefinite enough to point by analogy to kindred cases, specific enough to anchor problem-solving techniques, and open-ended enough to support research programs". At the same time, Kuhn would share with his critics "the conviction that only rules can offer the analytic clarity, reliability, and objectivity demanded of all kinds of reasoning worthy of the name, including scientific reasoning" (pp. 126-27). My solution to this difficulty is to consider values at two levels: they are not rule-governed at the community level, but are rule-governed at the individual level.

controversies that involved this sort of disagreement. The first was the dispute between the geocentrism and heliocentrism. According to Kuhn, prior to Kepler's innovations none of the theories presented a higher overall accuracy. His second examples is the debate between the oxygen and phlogiston theories. In this case, which Kuhn considers more typical, each theory was more accurate in a certain aspect, but not enough as to lead to an unequivocal choice. "To choose between them on the basis of accuracy," he concludes, "a scientist would need to decide the area in which accuracy was more significant" (1977d, p. 323).

A second form of divergence is related to the weight assigned to values (1977d, p. 322). Kuhn illustrates this situation once more with the controversy between heliocentric and geocentric systems. Consistency with other theories, particularly regarding terrestrial phenomena, he argues, spoke unequivocally in favor of the geocentric tradition. On the other hand, simplicity hung slightly in favor of Copernicus — while Ptolemy needed two circumferences to explain the general qualitative aspects of the planets, Copernicus needed only one (1977d, pp. 323-24; see also Laudan 1985, p. 289). In order to choose one of them, a scientist had to decide the importance of each of these values for a theory.

This means that more than a set of functions, the minimum criteria of rationality should be seen as a set of function-sketches. For each scientist *j*, it gains body as an *individual criterion of rationality*.[18] Our original postulates become, then:

$$\sum_{i=1}^{n} m_{ji} v_{ji}(t) = y_j,$$
$$m_{ji} > 0, \forall m_{ji} \in \boldsymbol{m_j},$$

And,

$$(\boldsymbol{m_j}^T . \Delta \boldsymbol{v_j} > 0) \leftrightarrow t_1 \succ_j t_2$$

in which v_{ji} represents the application of value v_i by scientist *j*; m_{ji} is the weight that *j* assigns to value v_{ji}; $\boldsymbol{m_j}$ is the vector representing the weights of values for *j*; and $\succ_j$ is the relation of preference for scientist *j*. Also, $\Delta \boldsymbol{v_j}$ is the

[18] Salmon (1990) proposes an alternative way to understand Kuhn's remarks on theory-choice, using Bayes' theorem. He assumes a general criterion of rationality (the Bayes algorithm) that permits divergences among individuals (the values attributed to the *a priori* probabilities). I think the approach formulated here through functions incorporates more easily two factors: the multiplicity of values used by scientists and the possible kinds of disagreement. The advantage of the Bayesian approach, on the other hand, is to display more easily the conditionalization, which is necessary for explaining scientists' evaluation changes (see also Chen & Zhang 2006). Another alternative approach is Kuukkanen's (2007) coherentist interpretation of Kuhn. This reading is also compatible with mine, and, as Salmon's, helps to illuminate other interesting aspects of Kuhn's thought. Particularly, the coherentist approach can give the same treatment to cumulative and non-cumulative changes. At least for assessments of rationality, comparative evaluation is all that matters. In this sense, normal and revolutionary changes can be treated similarly (another matter is the implication of revolutions for communication and truth).

vector representing the difference in evaluations of t_1 and t_2, according to *j*, for all v_i, such that

$$\Delta \boldsymbol{v_j} = \begin{pmatrix} v_{j1}(t_1) - v_{j1}(t_2) \\ v_{j2}(t_1) - v_{j2}(t_2) \\ \dots \\ v_{jn}(t_1) - v_{jn}(t_2) \end{pmatrix}.$$

Finally, we can define arbitrarily that

$$v_{ji} \in [0,1],$$
$$y_j \in [0,1],$$

and,

$$\sum_{i=1}^{n} m_{ji} = 1.$$

Having defined how the general criterion of rationality is individually specified, let us analyze some situations that involve its use in the choice of theories. Let us first imagine a situation of disagreement as to the individual application of values. Two scientists, $\{j, k\} \in J$, evaluate two theories, $\{t_1, t_2\} \in T$, according to five epistemic criteria: accuracy, scope, simplicity, consistency, and fruitfulness. These individuals disagree, though, on how effective theories are in relation to the manifestation of each value. For the first scientist, *j*, we have the following evaluation (Table 3):

	Accuracy	Scope	Simplicity	Consistency	Fruitfulness
Theory 1	0.8	0.5	0.4	0.8	0.5
Theory 2	0.4	0.6	0.7	0.3	0.5

k, by its turn, posses the following appreciation (Table 4):

	Accuracy	Scope	Simplicity	Consistency	Fruitfulness
Theory 1	0.5	0.5	0.2	0.9	0.5
Theory 2	0.6	0.4	0.8	0.2	0.5

Suppose, in this case, they both agree on the attribution of weights, and that their matrix of weights is, for instance,

$$\boldsymbol{m} = \begin{pmatrix} 0.4 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.2 \end{pmatrix}.$$

Assuming the simple linear model, their overall evaluations (the best theory is highlighted in gray) become, respectively (Table 5):

	Scientist *j*	Scientist *k*
Theory 1	0.64	0.51
Theory 2	0.48	0.52

Therefore, for the first scientist, *j*, we have that

$$f_j(t_1) > f_j(t_2),$$

and thus, by the second postulate of the minimum criterion of rationality,

$$t_1 \succ_j t_2.$$

For the second scientist, *k*, we have that

$$f_k(t_1) < f_k(t_2),$$

and hence

$$t_2 \succ_k t_1.$$

In short, a divergence in the individual application of values results in distinct preferences for theories.

Consider now a second case, in which scientists agree on individual values, but disagree on the weights assigned to them. The weights attributed by *j* and *k* are, respectively,

$$\boldsymbol{m_j} = \begin{pmatrix} 0.2 \\ 0.3 \\ 0.1 \\ 0.3 \\ 0.1 \end{pmatrix}, \boldsymbol{m_k} = \begin{pmatrix} 0.5 \\ 0.1 \\ 0.1 \\ 0.1 \\ 0.2 \end{pmatrix}.$$

Also, suppose that their common table of values is the following (Table 6):

	Accuracy	Scope	Simplicity	Consistency	Fruitfulness
Theory 1	0.4	0.7	0.1	0.8	0.6
Theory 2	0.6	0.3	0.5	0.6	0.4

In this case, the overall evaluation becomes (Table 7):

	Scientist *j*	Scientist *k*
Theory 1	0.6	0.48
Theory 2	0.48	0.52

Another type of disagreement in the application of values, regarding weighting, leads again to the first scientist choosing t_1, and the second, t_2.

The two cases illustrated above involve just one type of disagreement at a time: either in relation to the individual application of values or in relation to the relative weights. But since epistemic criteria are learned through examples and practice, and their interpretation depends on a scientist's particular traits and personal experiences, divergences are expected in both aspects — involving the v_i as well as the m_i.

For Kuhn, in sum, scientific disagreement can happen for purely epistemic matters (1962a, ch. 12). Rational scientists committed to the same values and in possession of the same body of evidences may, despite that, choose different theories (1977d, p. 324).

Other Sources of Variability

The variability in the application of values is, in reality, just one of the numerous possible sources of divergence in evaluation formulas.[19] It was the one chosen by Kuhn to discuss the problem in (1977d), possibly because it was not based on complex philosophical theses, as incommensurability, theory-ladenness of observation, or underdetermination of theories.

I would like to briefly mention some other factors that contribute to increase the chance of disagreement in judgments (1977d, p. 338). The most important of them is, certainly, the incommensurability of scientific theories (1962a, pp. 147-49). Scientists may disagree, for example, on which problems are more important or the way to solve them (Elliot 2017). Also, supporters of different paradigms are also "always slightly at cross-purposes" (1962a, p. 112), and their communication is "inevitably partial" (1962a, p. 148). And, in a certain sense, scientists from rival traditions live in worlds filled with different entities (1962a, ch. 10).

Agreement is also harder because tests never have a clear validity, being always subject to conflicting interpretations (1970a, sec. 3). The same can be said for evidence, which can invite alternative interpretations (Barnes & Bloor 1982, sec. 3; Shapin 1982). Finally, there is the fact that, in practice, scientists never fully share their bodies of evidences —

[19] And even in this case, we assumed a simple linear function of the values that reduces the possibility of disagreements. If scientists used other types of functions, the difference between evaluations would come not only from the values assigned to the individual criteria or the weights given to them, but also from the form of the function itself.

generally, they tend to be much more informed about things related to their own research fields.

All of these factors open up room for considerable more disagreement in evaluations than in the mere application of values. I will not explore them here, however, because the variability of values is enough for demonstrating the possibility of rational disagreement in Kuhn's approach. The only thing to bear in mind is that other elements make it even more difficult for scientists to agree on their choices.

CONCLUSION

It is obvious why Kuhn's concept of rationality led to so many criticisms and accusations of irrationalism. In the first place, it attributes a central role to non-epistemic factors in science. The latter not only determine the problems and fields that are explored and the rate with which knowledge is produced — they directly contribute to the appreciation and acceptance of theories. Further, this theory of rationality is incapable of explaining consensus formation by itself. Values are not sufficient to delimit a single choice for every individual. In consequence, scientists sharing the same values and evidences may choose distinct theories, by virtue of residual differences in their backgrounds. According to Kuhn, the canons of evaluation "are not by themselves sufficient to determine the decisions of individual scientists" (1977d, p. 325; 1962a, p. 4). "In matters of theory-choice", he says, "the force of logic and observation cannot in principle be compelling" (1970b, p. 126).

Because of this, many critics have concluded that, for Kuhn, scientific consensus has to be caused by non-epistemic forces that favor one theory over its alternatives. Since reason does not seem capable of producing a compelling agreement, consensus would have to be yielded "by factors that lie outside the 'normal' resources of scientific inquiry" (Laudan 1985, p. 294).

Answering this question lies beyond the scope of this paper, so I can only suggest the way Kuhn tries to close the gap between individual rationality and consensus formation. First, he stresses that not everything that is rational is necessarily "reasonable." Some interpretation of values can be rational (i.e. follow the requisites of rational action) and, despite that, not be considered by the community as serious alternatives (1962a, ch. 12). In second place, Kuhn thinks that scientific consensus can be produced despite divergent formulas of evaluations. This is where dominance enters: "Individual scientists," he claims, "embrace a new paradigm for all sorts of reasons and usually for several at once" (1962a, p. 151).

Lastly, Kuhn believes that scientific consensus is not engendered by the application of a set of methodological tools, but by a series of social-epistemic mechanisms that progressively persuades the members of the community.[21] The first, which, I already mentioned, is scientific pedagogy. Educational homogeneity restricts scientists' divergence in the application of values, facilitating future formation of consensus. Another mechanism is the continuous production of evidence and arguments in science. Scientists respond to the production of new evidence by constantly reevaluating their assessments. This may lead with time to the convergence of evaluations, changing the corresponding distribution of opinion in the community. The last mechanism is the restructuring of the community. When a group of scientists remains unconvinced of the validity of a theory, two things may happen. If they are few, the majority

[21] Kuhn (1962a, p. 151ff.) discusses some of the arguments that regularly convert scientists to a new paradigm.

may (intentionally or not) marginalize these holdouts. When the division among scientists is more profound, the community may split, generating new, independent disciplines (1962a, ch. 12). A detailed answer to this problem, however, is the subject for another article.[22]

REFERENCES

ARGAMAKOVA, A. "Modeling Scientific Development Lessons from Thomas Kuhn". In M. Mizrahi (ed.) (2019), pp. 45-59.

BARNES, B., BLOOR, D. "Relativism, rationalism and the sociology of knowledge". In M. Hollis and S. Lukes (eds.) (1982), pp. 201-60.

BLOOR, D. *Knowledge and social imagery*. Second edition. Chicago: University of Chicago Press, 1991.

CHEN, X., & JIANFENG, Z. "Bayesian Test and Kuhn's Paradigm". *Frontiers of Philosophy in China*, 1(3), pp. 491–505, 2006.

COHEN, R. S., LAUDAN, L. *Physics, Philosophy and Psychoanalysis*. Boston: Reidel, 1983.

D'AGOSTINO, F. *Incommensurability and Commensuration: The Common Denominator*. Hampshire: Ashgate, 2003.

_____ "Kuhn's Risk-Spreading Argument and the Organization of Scientific Communities". *Episteme*, 1(3), pp. 201-09, 2005.

[22] For two works that develop the mechanisms of formation of scientific consensus in a Kuhnian outlook, see D'Agostino (2010) and Kitcher (1993, ch. 8).

_____ *Naturalizing Epistemology: Thomas Kuhn and the "Essential Tension"*. London: Palgrave Macmillan, 2010.

DE LANGHE, R. "The Problem of Kuhnian Rationality". *Philosophica*, 86, pp. 11-31, 2012.

ELLIOTT, K. C. *A Tapestry of Values: An Introduction to Values in Science*. New York: Oxford University Press, 2017.

FAVRETTI, R. R., SANDRI, G., SCAZZIERI, R. *Incommensurability and Translation*. Cheltenham, UK: Edward Elgar, pp. 33-37, 1999.

GOLDMAN, A., WHITCOMB, D. *Social Epistemology: Essential Readings*. Oxford: Oxford University Press, 2011.

HACKETT, E. J., AMSTERDAMSKA, O., LYNCH, M., WAJCMAN, J. *The Handbook of Science and Technology Studies*. Third edition. Cambridge, MA: MIT Press, 2008.

HACKING, I. *Scientific Revolutions*. Oxford: Oxford University Press, 1981.

HEMPEL, C. "Turns in the Evolution of the Problem of Induction". *Synthese*, 46, pp. 389-404, 1981.

_____ "Valuation and Objectivity in Science". In R. S. Cohen & L. Laudan (eds.) (1983), pp. 73-100.

HOLLIS, M., LUKES, S. *Rationality and Relativism*. Oxford: Blackwell, 1982.

HOYNINGEN-HUENE, P. *Reconstructing Scientific Revolutions*. Chicago: Chicago University Press, 1993.

KITCHER, P. *The Advancement of Science*. Oxford: Oxford University Press, 1993.

KUHN, T. S. *The Structure of Scientific Revolutions*. Chicago: Chicago University Press, 1962a/2012.

______ "Comment on the Relations of Science and Art". In *The Essential Tension*. Chicago: CUP, 1969c/1977.

______ "Logic of Discovery or Psychology of Research". In *The Essential Tension*. Chicago: CUP, 1970a/1977.

______ "Reflections on my Critics". In *The Road since Structure*. Chicago: CUP, 1970b/2010.

______ "Postscript". In *The Structure of Scientific Revolutions*. Chicago: CUP, 1970c/2012.

______ "Preface". In *The Essential Tension*. Chicago: CUP, 1977b/1977.

______ "Objectivity, Value Judgment, and Theory Choice". In *The Essential Tension*. Chicago: CUP, 1977d/1977.

______ "Rationality and Theory Choice". In *The Road since Structure*. Chicago: CUP, 1983d/2010.

______ "The Road since *Structure*". In *The Road since Structure*. Chicago: CUP, 1991a/2010.

______ "Afterwords". In *The Road since Structure*. Chicago: CUP, 1993a/2010.

______ "A Physicist Who Became A Historian For Philosophical Purposes: A discussion between Thomas S. Kuhn and Aristides Baltas, Kostas Gavroglu, and Vassili Kindi". In *The Road since Structure*. Chicago: CUP, 1997/2010.

______ "Remarks on Incommensurability and Translation". In R. R. Favretti et al. (eds.) (1999), pp. 33-37.

KUUKKANEN, J-M. "Kuhn, the Correspondence Theory of Truth and Coherentist Epistemology". *Studies in History and Philosophy of Science Part A*, 38, pp. 555–566, 2007.

LACEY, H. "Is there a significant distinction between cognitive and social values?". In P. Machamer and G. Wolters (eds.) (2004), pp. 24-51.

LAKATOS, I. "Falsification and the Methodology of Scientific Research Programmes". In I. Lakatos and A. Musgrave, A. (eds.) (1970), pp. 91-196.

______ , MUSGRAVE, A. *Criticism and the Growth of Knowledge*. London: Cambridge University Press, 1970.

LAUDAN, L. *Progress and its Problems*. Berkeley: University of California Press, 1978.

______ "A Problem-Solving Approach to Scientific Progress". In I. Hacking (ed.) (1981), pp. 144-155.

______ "Kuhn's Critique of Methodology". In J. Pitt (ed.) (1985), pp. 283-300.

MACHAMER, P., WOLTERS, G. *Science, values and objectivity*. Pittsburgh: Pittsburgh University Press, 2004.

MIZRAHI, M. *The Kuhnian Image of Science*. London: Rowman & Littlefield, 2019.

MODY, C., KAISER, D. "Scientific Training and the Creation of Scientific Knowledge". In Hackett, E. J.; Amsterdamska, O.; Lynch, M.; Wajcman, J. (eds.) (2008), pp. 377-402.

PITT, J. *Change and Progress in Modern Science*. Reidel: Dordrecht, 1985.

SAVAGE, C. W. *Scientific Theories*. Minneapolis: University of Minnesota Press, 1990.

SCHEFFLER, I. *Science and Subjectivity*. Second edition. Indianapolis: Hackett Publishing Company, 1982.

SHAPERE, D. "The Structure of Scientific Revolutions". *The Philosophical Review*, 73(3), pp. 383-94, 1964.

SHAPIN, S. "History of Science and its Sociological Reconstructions". *History of Scienc*e, 20(3), pp. 157–211, 1982.

WITTGENSTEIN, L. *Philosophical Investigations*. Basil: Wiley-Blackwell, 2009.

WORRALL, J. "Scientific Revolutions and Scientific Rationality: The Case of the Elderly Holdout". C. W. Savage (ed.) (1990), pp. 319-54.

WRAY, B. K. *Kuhn's Evolutionary Social Epistemology*. Cambridge: Cambridge University Press, 2011.