Title: A History of Metaethics and Values in Science

Forthcoming in: Tsou, Jonathan Y., Shaw, Jamie, and Fehr, Carla (eds.), *Values, Pluralism, and Pragmatism: Themes from the Work of Matthew J. Brown*. Boston Studies in the Philosophy and History of Science. Cham: Springer.

Author: Paul L. Franco

Abstract: I pursue Matt Brown's (2020) and James Leach's (1968a) historical suggestions that there is a relationship between metaethical noncognitivism and certain arguments for the value neutrality of the sciences. In particular, I relate C.L. Stevenson's emotivism to arguments by Herbert Feigl, Carl Hempel, and Ernest Nagel for the sciences' value neutrality. I also consider whether these arguments can be disentangled from their controversial metaethical claims by looking at Robert Alexander's (1974) account of value neutrality based on the view that a scientist's aims are discharged by making empirical statements. Drawing upon Leach's (1968a; 1969b) defense of the argument from inductive risk, I argue Alexander fails to offer a metaethically neutral version of the value neutrality of the sciences. Though I do not explicitly explore this, I think the history I sketch is relevant to recent calls for philosophers of science to more fully characterize the 'values' in 'values and science.'

Acknowledgments: Versions of this paper were presented at the Philosophy of Science Association's 2022 poster session, Cal State-Long Beach's Philosophy Day!, and UW Philosophy's works-in-progress series. Thanks to those audiences for discussion and encouragement. Thanks also to two anonymous reviewers for helpful suggestions, to the students in my Autumn 2023 seminar for working through some of the history of science and values with me, and to Matt Brown for his support over the years.

1. Introduction

In his recent book, Matt Brown claims the worry that "let[ting] values into science…will lead to bias, subjectivism, and wishful thinking…is based on an implicit view that values themselves have no cognitive status or somehow systematically inferior cognitive status to other components of inquiry, such as data, experiments, hypotheses, logic, and reasoning" (2020, pp. 89–90). He further claims, "The implicit model of values that most philosophers of science work with, whether or not they think that science ought to be value-free, is generally an emotivist or at least noncognitivist one" (p. 103). Roughly, metaethical noncognitivism holds that value judgments are not truth-apt. One way to be a noncognitivist is to be an emotivist who holds that value judgments *express* one's attitudes or emotions in ways aimed at influencing others, but do not *state* anything true or false. Brown suggests this implicit model philosophers of science work with has roots in logical empiricism's outsized influence in the development of philosophy of science (p. 101).

Reaching slightly further back in history, James Leach says, "Following the progressive elimination of teleology from the province of the scientist in the seventeenth century, the domains of science and values seemed finally isolated and immune from the other" (1968a, p. 93). As a result, "the very notion of scientifically supported values [was] exploded as a 'naturalistic fallacy'" (p. 93). Finally, on Leach's telling, "the convergence in value theory of…emotivists and oxonians, conjoined with the success and influence of a positivist philosophy of science, consummated the separation of science and value" (p. 93). To unpack the last of Leach's claims: Emotivists, like C.L. Stevenson, and Oxonians, like R.M. Hare, converged on noncognitivism, as did (some of) the logical empiricists.[1] Eliding important differences, on these accounts, value judgments are distinct from empirical statements in that they express emotions, commend or condemn, or issue imperatives rather than describe the world. For Leach, these developments "enthroned…the value neutrality thesis of science…which prohibits the scientist from making value judgments" (p. 93).

In this paper, I pursue Brown's and Leach's historical suggestions that there is a relationship between metaethical noncognitivism and certain arguments for the value neutrality of the sciences. In particular, I look at Stevenson's influential noncognitivist view developed contemporaneous with logical empiricism and relate it to Herbert Feigl's, Carl Hempel's, and Ernest Nagel's respective arguments in the mid-20th-century for the value neutrality of science. I also consider whether these arguments can be disentangled from their controversial metaethical claims by looking at Robert Alexander's (1974) account of value neutrality based on the view that a scientist's aims are

[1]See Anne Sigetsleitner (2022) for an overview of the noncognitivism of some logical empiricists.

discharged by making empirical statements. Drawing upon Leach's (1968a; 1969b) defense of a version of the argument from inductive risk, I argue Alexander fails to offer a metaethically neutral version of the value neutrality of the sciences.

I sketch this history for two reasons. First, it highlights one facet of the conceptual development of the value-free ideal in parts of philosophy of science that turns on a particular view of value judgments as neither true nor false.[2] Second, though I do not explore it here, I think the history is potentially relevant to calls for philosophers of science to more fully characterize the 'values' in 'values and science' rather than relying on pre-theoretical notions of values and their roles in science (Scriven 1972; Brown 2020; Ward 2021; Wieten 2021).

1.1 Plan for the paper

In §2, I review early analytic metaethical arguments for a strong distinction between empirical statements and value judgments, focusing especially on Stevenson's emotivism. In §3, I show how metaethical commitments similar to Stevenson's inform accounts of the value neutrality of the sciences by Feigl, Hempel, and Nagel. In §4, I detail Alexander's (1974) argument that the value neutrality of the sciences can be formulated independent from the controversial metaethical commitments taken on by Feigl, Hempel, and Nagel. In §5, in response to Alexander, I use Leach (1968a; 1968b) to draw connections between metaethics and the argument from inductive risk. §6 concludes.

2. C.L. Stevenson's noncognitivist emotivism

While it is well-known that A.J. Ayer (1946 [1952], Chap. VI) outlines an emotivist ethics and some logical empiricists develop related accounts (Brown 2020, pp. 101–103; Sigetsleitner 2022), I foreground Stevenson's emotivism, as his version was more developed and influential in analytic metaethics.[3] Moreover, Feigl endorses Stevenson's 1944 *Ethics and Language* as "one of the very few and exceptional books on ethics in which a sound scientific attitude combined with a logical analysis of great subtlety and thorough penetration is applied toward the clarification of moral discourse" (1946, p. 80). Here, I focus on a 1937 article that provides an overview of Stevenson's

[2] For other aspects of this development, some related to the themes of this paper, see Don Howard (2003), Heather Douglas (2009, Chap. 3), and Krist Vaesen and Joel Katzav (2019).

[3] In the second edition of *Language, Truth and Logic*, Ayer says his emotive theory "needs to be supported by a more detailed analysis" (1946 [1952], p. 20). In a footnote, he says, "this deficiency has been made good by C.L. Stevenson in…*Ethics and Language*" (p. 20, n.3).

emotivism, "The Emotive Theory of Ethical Terms." Despite Stevenson not being a logical empiricist, this article was reprinted in Ayer (1959) alongside canonical logical empiricist texts.

Some brief background for Stevenson's view will be helpful. Much of twentieth-century analytic metaethics is framed in relation to G.E. Moore's *Principia Ethica*. For Moore, the primary subject matter of ethics is the property good. Moore understands good to be a simple, unanalyzable, and nonnatural property. He claims, "'good' is a simple notion, just as 'yellow' is a simple notion…just as you cannot, by any manner of means, explain to any one who does not already know it, what yellow is, so you cannot explain what good is" (Moore 1903 [1959], p. 7). Accordingly, any account of good that defines it in terms of or reduces it to other properties misunderstands its very nature. Moore is particularly concerned with avoiding the so-called naturalistic fallacy. Ethical views make this mistake when they aim at "substituting for 'good' some one property of a natural object or of a collection of natural objects; and in thus replacing Ethics by some one of the natural sciences" (p. 40). Ethical judgments stand independent of empirical statements, and cannot be reduced to them, because for any natural property put forward as identical with good it always remains an open question whether that property is actually good. This is Moore's open-question argument. Consider hedonists who analyze goodness in terms of pleasure. Such views commit the naturalistic fallacy by reducing good to a natural property, namely the psychological state of pleasure, of which it remains an open to question to ask, without contradiction, if pleasure is actually good. As Stevenson understands Moore, "No matter what set of scientifically knowable properties a thing may have (says Moore, in effect), you will find, on careful introspection, that it is an open question to ask whether anything having these properties is *good*….That is, the scientific method is not sufficient for ethics" (1937, p. 18). If good is a nonnatural property not reducible to empirically discoverable natural properties, then science and ethics are separate domains.

It is important to note that two of Moore's reasons for separating the domain of ethics from the domain of the natural sciences would not be acceptable to philosophers of science who adopt an empiricist criterion of meaning even if they similarly claim that the ethical and the scientific are independent. First, Moore takes the subject matter of ethics to be an unanalyzable, nonnatural property irreducible to any natural property. Second, he thinks ethical statements are truth-evaluable and known intuitively, rather than empirically.

However, two decades later, C.K. Ogden and I.A. Richards offer a linguistic formulation of Moore's separation of science and ethics that resonates with Stevenson and later accounts from philosophers of science. In *The Meaning of Meaning*, Ogden and Richards divide "the functions of language into two groups, the symbolic and the emotive" (1923 [1945], p. viii). They understand

Moore's claim that "'good' is…a unique unalysable concept" to suggest the "peculiar ethical use of 'good' is…a purely emotive use" (p. 125). For Ogden and Richards, on one hand, "The symbolic use of words is *statement*; the recording, the support, the organization and the communication of references" (p. 149). The scientific use of words is restricted to the statement (p. 124). On the other hand, "The emotive use of words is a more simple matter, it is the use of words to express or excite feelings and attitudes" (p. 149).[4] Ogden and Richards think 'good' has no referential function; it does not pick out any property. Instead, "it serves only as an emotive sign expressing our attitude…and perhaps evoking similar attitudes in other persons, or inciting them to actions" (p. 125). So, on this view, 'good' is unanalyzable because it is used to express attitudes and influence others rather than to refer to and make statements about properties in the world.

Stevenson credits Ogden and Richards's interpretation of Moore with inspiring his emotivist metaethics (1937, p. 23 n1). Following them, Stevenson claims, "Broadly speaking, there are two different *purposes* which lead us to use language" (p. 21). The descriptive use of language is "(as in science) to record, clarify and communicate *beliefs*"; the dynamic use of language is "to give vent to our feelings (interjections), or to create moods (poetry), or to incite people to actions or attitudes (oratory)" (p. 21).[5, 6] Stevenson understands the distinction between the descriptive and dynamic use of words pragmatically, in terms of the aims of the speaker, which the audience infers with the help of a speaker's "tone of voice…gestures, the general circumstances…and so on" (p. 22).

The emotive meaning of some words, understood semantically, makes them especially suited for dynamic uses. On Stevenson's view, "The emotive meaning of a word is a tendency of a word, arising through the history of its usage, to produce (result from) *affective* responses in people. It is the immediate aura of feeling which hovers about a word" (p. 23). For him, "there is an important contingent relation between emotive meaning and dynamic purpose: the former assists the latter" (p. 24). Ethical discourse is aimed at directing people's interests to get them to act using words with strong emotive meaning that make them well-suited for dynamic purposes. For example, Stevenson

[4] Susan Stebbing, another contemporary and interlocutor of the logical empiricists, makes a similar distinction (1930 [1942], pp. 16–17).

[5] Stevenson recognizes a speaker might use words with both purposes (1937, p. 21).

[6] An anonymous reviewer points out Carnap also distinguishes descriptions of facts and expressions of attitudes in his criticism of metaphysics. Carnap: "the (pseudo) statements of metaphysics do not serve for the *description of states of affairs*… They serve for the *expression of the general attitude of a person towards life*" (1932 [1959], p. 78). Such expressions are not truth-evaluable. Carnap (1935) argues moral claims are imperatives, which cannot be true or false and have no empirical consequences. Reichenbach (1951, p. 280) makes a similar point. More on Carnap's noncognitivism in note 9.

says, "The word 'good' has a pleasing emotive meaning which fits it especially for the dynamic use of suggesting favourable interest" (p. 25). When I call some action or state of character good, I express my favorable attitude towards it and suggest my audience take the same attitude. Given the strong feelings and interests associated with terms expressing moral approval and disapproval (p. 26), moral terms like 'good' are especially useful for influencing others via the dynamic use of language.

Stevenson thinks his emotive theory satisfies three requirements a metaethical theory must meet to capture the "vital" sense of 'good' (p. 16). Here, I highlight his first and third requirements. The first requirement is that "we must be able sensibly to *disagree* about whether something is 'good'"; the third requirement is related to the naturalistic fallacy and Moore's open-question argument, namely, that "the 'goodness' of anything must not be verifiable solely by use of the scientific method" (p. 16). Consider his first requirement. Stevenson argues, "We must distinguish 'disagreement in belief' (typical of the sciences) and 'disagreement in interest'" (p. 27).[7] For Stevenson, the latter "occurs when A has a favourable interest in X, when B has an unfavourable one in it, and when neither is content to let the other's interest remain unchanged" (p. 27). Disagreement in interest is the type of disagreement characteristic of ethics. Resolving ethical disagreement starts with "the suggestive force of [an] ethical sentence," i.e., a sentence with emotive terms used dynamically (p. 28). This use "exerts enough pressure to start…trains of reasons" aimed at giving "a more thorough account of the object of interest" (p. 28). This more thorough account may be descriptive, e.g., it might consist in specifying the likely consequences of some course of action relative to another. In this way, "the empirical method is relevant to ethics simply because our knowledge of the world is a determining factor to our interests. But," Stevenson continues, "note that empirical facts are not inductive grounds from which the ethical judgment problematically follows" (p. 29). Empirical statements made in attempts to resolve disagreement in interest are in the service of influence and persuasion; they are not logical grounds from which ethical conclusions follow. This relates to Stevenson's third requirement. Recalling Moore, Stevenson says that a factual description of a thing's natural properties leaves open the question of whether those properties are good; this is because "to ask whether it is good is to ask for influence. And whatever I may know about an object, I can still ask, quite pertinently, to be influenced with regard to my interest in it" (p. 30). Factual statements aimed at saying something true do not entail ethical judgments aimed at influence.

I turn now to showing how this claim about the logical relationship between ethics and the sciences, and the related distinction between the purposes of scientific language—to state facts—and

[7] For more on this distinction, see Stevenson (1948 [1963]).

ethical language—to express approval or disapproval and influence others—is central to arguments for the value neutrality of the sciences advanced by Feigl, Hempel, and Nagel.[8, 9]

3. Arguments for the value neutrality of the sciences

I start with Feigl's 1950 paper, "The Difference Between Knowledge and Valuation."[10] Of interest for our purposes is his characterization of disagreement in the sciences and disagreement in ethics. Regarding disagreement in the sciences, Feigl says, "One of the characteristic features of science…is the possibility of removing disagreements concerning matters of fact. Observation, measurement, experimentation and statistical analysis, together with generally well agreed upon methods of interpretation and inference, provide at least in principle the means for the adjudication of scientific disputes" (1950, p. 40). In principle, "Everyone, equipped with the requisite intelligence and instruments, is in a position to examine the validity of assertions or hypotheses regarding matters of fact. This is the essential meaning of the *objectivity* (better perhaps, 'intersubjectivity') of the scientific method" (p. 40). It is important to note that claims about the relationships between means and ends are straightforwardly empirical for Feigl. With regards to practical reasoning about the means necessary to achieve ends, the objectivity characteristic of the sciences can be achieved. So long as there is agreement on ends or shared purposes and interests, the empirical methods of the sciences are relevant to successful means-end reasoning.

Note, though, that agreement on ends is needed to "evaluate means (actions and instrumentalities)" (p. 41). Someone might accept an empirical claim that a particular course of action achieves some end while rejecting the end itself. Indeed, Feigl thinks this type of disagreement characterizes ethical discourse. When it comes to the interests and purposes that set our ends, "we cannot ignore the actuality (or at least the possibility) of cognitively irremediable disagreements in regard to purposes" (p. 41). When choosing ends themselves, "it remains possible that no matter how

[8] See Douglas (2009, Chap. 3) for more on Hempel's and Nagel's arguments as they relate to inductive risk. Both Douglas and Vaesen and Katzav (2019) explore the relationship between Reichenbach's (1951) noncognitivism and claims about the value neutrality of the sciences, complementing themes in §3.

[9] The same anonymous reviewer in note 7 helpfully suggests Carnap as another source of influence for the views in §3. I agree, but focus on Stevenson given Feigl's praise of Stevenson's book and because Stevenson develops a fuller account than Carnap. Christian Damböck notes Carnap did not publish a developed account of noncognitivism, which makes it difficult to trace the reception of his views (2022, p. 498). See Damböck for more on connections between Reichenbach's and Carnap's noncognitivism, and Carnap's appreciation for Stevenson. Thanks to the same reviewer for suggesting Damböck.

[10] See Matthias Neuber (2022, §5) for an overview of Feigl on value judgments. See also Feigl (1952).

complete the information of each party to a dispute over *ends*, no agreement on these ends may be attainable" (p. 41). Unlike disagreement in the sciences, there is no agreed-upon method for settling disagreement in ethics.[11] For Feigl, "This implies that there is nothing in the nature of evaluation that ensures its intersubjectivity;" like Stevenson, he thinks ethical disagreement is settled by "persuasion, compromise, separation or coercion" rather than scientifically (p. 41).

In part, this is due to a more fundamental difference Feigl sees between the sciences and ethics. He argues, "scientific knowledge consists of *information*…Scientific statements can by their very nature not contain such words as 'advisable,' 'desirable,' 'should,' 'ought,' or 'must'" (pp. 41–42). Here, again, is a linguistic gloss on the differences between the sciences and ethics. The sciences describe, whereas ethics provides "The stamp of approval, the accent of endorsement, the force of encouragement, the factor of exhortation" in light of "the needs and interests of…agents" (p. 41). Scientific statements are silent with regards to approval, endorsement, encouragement, or exhortation. As such, "Scientific knowledge, even if it concerns the conditions and consequences of valuations, is in a perfectly clear sense value-free or value-neutral" (p. 41). Empirical information does not entail ethical judgments about purposes, interests, or ends.

In "Science and Human Values," Hempel also distinguishes two types of value judgments when considering the relationship between the scientific method and ethics. The first type, "a *relative*, or *instrumental, judgment of value*…states that a certain kind of action, *M*, is good (or that it is better than a given alternative M_1) *if* a specified goal is to be attained; or more accurately, that *M* is good, or appropriate, for the attainment of goal *G*" (1960 [1965], p. 84). Like Feigl, Hempel thinks the empirical methods of the sciences are relevant to instrumental judgments of value. Indeed, "a relative, or instrumental, judgment of value can be reformulated as a statement which expresses a universal or a probabilistic kind of means-end relationship, and which contains no terms of moral discourse…at all" (p. 85). Still, recalling Feigl's view, such hypothetical judgments do not require anyone adopt the particular ends in question.

Hempel's second type of value judgment relates to ends themselves, i.e., "an *absolute, or categorical, judgment of value* to the effect that a certain state of affairs (which may have been approved as a goal or end) is good, or that it is better than some specified alternative" (p. 85). Consider a categorical judgment like 'Killing is evil'; Hempel claims "it serves to express a standard

[11] An anonymous reviewer asks if Feigl considers the possibility of scientific disagreement that cannot be settled. Feigl notes "the notorious practical difficulties of the decision between rival hypotheses," but seems to think these difficulties are in-principle solvable by shared methods (1950, p. 40).

for moral appraisal or a norm of conduct" (p. 85). Recalling noncognitivist accounts of ethical language, categorical judgments of value express approval or commitment to certain standards and norms, but do not state anything like instrumental judgments (p. 85). Thus, Hempel claims, "Categorical judgments of value, then, are not amenable to scientific test and confirmation or disconfirmation; for they do not express assertions but standards or norms for conduct" (p. 86). For Hempel, judgments about ends do not describe, and so do not have observational consequences and are not truth-apt (p. 85).

Still, when it comes to an instrumental judgment of value, "Science can render an indispensable service by providing us with increasingly extensive and reliable information relevant to our purpose" (p. 89). However, the methods and results of science do not speak for or against categorical judgments of values typical of ethical discourse. For Hempel, "no set of scientific statements logically implies an unconditional judgment of value" (p. 91). Given the lack of logical connection between descriptive statements and categorical judgments of value, "it remains for us to *evaluate* the various probable sets of consequences of the alternative choices under consideration. And this requires the adoption of pertinent valuational standards which are not objectively determined by the empirical facts" (p. 89). Moreover, like Feigl and Stevenson, Hempel countenances the possibility of cognitively irremediable disagreement over the standards needed to evaluate the choices and consequences relevant to ethical decisions.

Themes and distinctions similar to Feigl's and Hempel's account of the value neutrality of the sciences appear in Nagel's discussion of the social sciences in *The Structure of Science*.[12] There, Nagel distinguishes two types of value judgments. The first is "the sense in which a value judgment expresses *approval or disapproval* either of some moral (or social) ideal, or of some action (or institution) because of a commitment to such an ideal" (1961, p. 492). This recalls Hempel's characterization of categorical value judgment, the function of which is to express commitment to some value or approval or disapproval of an action. Nagel calls these "'appraising value judgments'" which hold that "some envisaged or actual state of affairs is worthy of approval or disapproval" (p. 493).

Nagel's second sense of value judgment is related to Feigl's and Hempel's views on means-end judgments. This is "the sense in which a value judgment expresses *an estimate* of the degree to which some commonly recognized (and more or less clearly defined) type of action, object, or institution is embodied in a given instance" (p. 492). Nagel uses 'anemia' as an example. He says,

[12] See Anna Alexandrova (2018) for recent work critically engaging Nagel.

"although the meaning…can be made quite clear, it is not in fact defined with complete precision; for example, the notion of a 'normal' number of red corpuscles that enters into the definition of the term is itself somewhat vague, since this number varies with the individual members of a species as well as with the state of a given individual at different times" (p. 492). On Nagel's account, given the vagueness associated with the term 'anemia,' a scientist "must judge whether the available evidence *warrants* the conclusion that the specimen is anemic" by using "the measure provided by…[an] assumed standard" (p. 492). He calls these "'characterizing value judgments'" and they consist in "evaluations of the evidence, which conclude that a given characteristic is in some degree present (or absent) in a given instance" (p. 492). On my reading, characterizing value judgments single out a property normally presumed to be desirable or related to something like normal functioning, and are made when that property is determined to be present relative to some antecedently accepted standard.

Having distinguished two senses of value judgments, Nagel asks whether the social sciences can be value-neutral in light of arguments that the social scientist uses thick terms in the explanation of human behavior, e.g., in characterizing survey responses as "uninformed, deceitful, or irrational" (p. 494). Nagel grants the social scientist uses thick terms, which "as commonly used have a widely recognized pejorative overtone" (p. 494); or, to draw upon Stevenson, they use words with emotive meanings. Nagel even grants that "anyone who employs such terms to characterize human behavior can normally be assumed to be stating his disapprobation of that behavior…and not simply characterizing it" (p. 494). Nevertheless, he maintains social scientists using such terms need only assume the standard necessary to make a characterizing value judgment, and need not take the additional step of making an appraising value judgment.

This is because of an important difference between the functions of the types of value judgments. Appraising value judgments express approval or disapproval or commitment to some norm, whereas characterizing value judgments make an empirical claim relative to an assumed standard, which someone may or may not approve or disapprove. For Nagel, "an investigator making a characterizing value judgment is not…logically bound to affirm or deny a corresponding appraising evaluation" (p. 493). While "it is not always easy to make the distinction formally explicit in the social sciences" (p. 494), keeping separate the two judgments in the practice of science and in a philosophical account of science "are essentially practical rather than theoretical problems" (p. 495). In other words, the use of thick terms by the social scientist does not show fact and value are logically intertwined in a way rendering the value neutrality of the social sciences impossible. We can make appropriate, logical distinctions between value judgments, even if it is practically difficult to do so in practice.

Yet again, like Feigl and Hempel, Nagel advances an argument for the value neutrality of science turning on the claim that value judgments about ends are not logically entailed by factual statements, even if means-end value judgments are empirical in nature. And one reason Nagel offers in support of this conclusion concerns a difference in the linguistic functions of scientific statements and ethical judgments. As I have been suggesting, these claims are central to the metaethical noncognitivism advanced by Stevenson decades earlier.

4. Is value neutrality metaethics free?

Nagel (1961, p. 493) maintains his formulation of the value neutrality of the sciences does not assume any particular metaethical view. Alexander says this is "clearly incorrect" (1974, p. 394). As he makes clear, one way to understand Nagel's view—and *mutatis mutandis* Feigl's and Hempel's views—is that "*characterizations do not entail any appraisals*" (p. 394). This holds only if empirical statements are not inductive or deductive grounds from which appraising value judgments expressing approval or disapproval follow. However, as Alexander argues, "If one is a naturalist…the meaning of appraisals is rendered transparent by virtue of definitions composed entirely of characterizations. Thus, since the appraisal is ultimately nothing more than a characterization for naturalism, it is an obvious truth that the appraisals are entailed by characterizations on this metaethical view" (p. 394). An example will illustrate Alexander's point.

Take Philippa Foot, who advances the sort of naturalist thesis Alexander has in mind. Foot is concerned with whether an evaluative conclusion can follow from nonevaluative premises (1958, pp. 506–509). To tackle this question, she considers the thick term 'rude.' According to her, 'rude' "expresses disapproval, is meant to be used when action is to be discouraged, implies that other things being equal the behavior to which it is applied will be avoided by the speaker, and so on" (p. 507). So, the judgment, e.g., 'Interrupting someone who is speaking is rude,' expresses one's disapproval of interrupting and is meant to discourage that behavior. In this way, uses of 'rude' have the same features as Stevenson's dynamic uses of words with emotive meaning, Hempel's categorical judgments of value, and Nagel's appraising value judgments. Yet, any application of 'rude' must meet factual criteria to be used correctly. In calling an action 'rude,' we are not only expressing a personal attitude, but claiming certain facts hold. We do not call just any action rude if we disapprove of it; 'rude' is applicable in cases where "behavior causes offence by indicating lack of respect" and inapplicable otherwise (p. 507). Now, whether an instance of interrupting someone causes offense in a way that indicates a lack of respect is an empirical question. For Foot, one cannot

make such a factual determination and deny the action was rude (p. 509). This holds for all competent speakers of a language; one cannot unilaterally reject such a definition simply because they do not have the requisite attitude.[13] In Alexander's terms, the meaning of the appraisal, 'That act of interrupting was rude,' is rendered transparent by the characterization, 'That act of interrupting caused offense in a way indicating lack of respect.' Where the latter empirical statement holds, the former value judgment follows.[14]

As such, arguments for the value neutrality of the sciences that claim empirical statements do not entail value judgments make metaethical assumptions a naturalist would reject. However, according to Alexander, there is a formulation of the value neutrality thesis that does not beg the question against naturalistic metaethics. In considering whether or not the sciences are value-neutral, Alexander thinks "we also have to consider whether the scientist is trying to achieve something with his descriptions or explanations which would make appraisals relevant and necessary to his job, and consequently, something he should include if his product is to be adequate and complete" (1974, pp. 394–395). So, he asks: "what is the guiding conception of the empirical scientist's job?" (p. 395). For Alexander, the answer to this question will tell us if the scientist is simply out to make descriptive statements, or if they must also make something like appraising value judgments to achieve their aims (p. 395).

Alexander's answer is that the job of the scientist is "'to explore, to describe, to explain, and to predict the occurrences of the world we live in'" (p. 395; quoting Hempel 1966, p. 1). In light of this, he thinks we can grant metaethical naturalism, and still hold fast to value neutrality. Even if appraising value judgments follow directly from facts, given the task of the scientist, "the entailed appraisals are not *independently* relevant for scientific confirmation…since they are either epistemologically posterior to characterizations or strictly redundant" (Alexander 1974, p. 397). On one hand, if an appraising value judgment adds something on top of the characterizing judgments it logically follows from, e.g., some expression of approval, then a scientist need not make it; their job of describing, explaining, or predicting is discharged by the characterizing, or factual judgment since approval is not relevant to confirmation. On the other hand, if, as the naturalist claims, "an appraisal is simply some characterization…the appraisal can be ignored as redundant in favor of its definitionally equivalent characterization" (p. 398). Alexander argues, then, that no matter one's

[13] On Foot's view, "whether a man is speaking of behaviour as rude or not rude, he must use the same criteria as anyone else" (p. 509).

[14] Foot argues value judgments are derivable from empirical statements. Steven Diggin (2022) engages literature in science and values and argues empirical statements are derivable from ethical judgments.

metaethical view, "Not all of the logical consequences of an acceptable scientific proposition are scientifically relevant" in the sense of confirming or disconfirming the scientist's descriptions, explanations, or predictions (p. 399). Even when it comes to the act of publication, the scientist need not make explicit any appraisals that follow from characterizations since that would be taking on "the role of appraiser, and no longer simply…the role of a scientist" (p. 399).

Alexander's formulation of science's value neutrality does appear independent of controversial, noncognitivist glosses on ethical judgments. However, drawing upon Leach (1968a; 1968b), I argue in the next section that Alexander assumes a divide between the realms of theoretical and practical reason that the argument from inductive risk challenges.

5. Inductive risk, and the practical and theoretical

Alexander is right that securing the value neutrality of the sciences on the grounds that value judgments are not logically entailed by scientific statements because value judgments merely express approval or disapproval and scientific statements are factual is metaethically controversial. His attempt to formulate a metaethically neutral version of value neutrality appeals to the aims of the scientist, namely, to make well-supported, reliable claims and predictions about the world. Alexander's argument also relies on a distinction "between the *process* of science (the scientist's *activities* of choosing problems, forming concepts, validating hypotheses, etc.) and the *product* of that process (the *propositions* accepted into the body of scientific knowledge)" (1974, p. 391). He restricts his focus to the product of knowledge, arguing "the scientific product, at least in principle, is value neutral in the sense that these propositions need only contain factual claims, and not also value claims" (p. 391). He thinks even if value judgments are logically entailed by scientific statements, as ethical naturalists maintain, the scientist's job is discharged by making factual claims. Were the scientist to "use his scientific results for purposes of making practical appraisals," they would no longer be acting "*qua* scientist…but *qua* governmental consultant, concerned citizen, moral agent, or human being" (p. 400).

Importantly, Alexander thinks questions raised by the argument from inductive risk about the relationship between theoretical reason and practical reason are independent of his argument for the value neutrality of the product of science; considerations of inductive risk apply only to what he identifies as the process and not the product of scientific practice (p. 392). I think this is mistaken because his conception of the confirmation of empirical statements makes assumptions about the relationship between belief and action that some would find controversial.

Consider Leach's (1968a; 1968b) defense of the argument from inductive risk. Leach says arguments about the value neutrality of the sciences fail to specify "what sorts of value the scientist should avoid and hence what type of neutrality to maintain" (1968a, p. 94). Leach prefigures Alexander's focus on the scientist's role when he claims that one way to make value neutrality more specific is to look at "the goals of scientific inquiry" given that "the values ones advocates, seeks or avoids are correlated with goals or objectives" (p. 94). Leach considers three different ways to formulate the goals of science corresponding to stronger or weaker forms of value neutrality. The strongest version—"the 'odds maker' view of science" (p. 95)—holds that the scientist only "assigns degrees of confirmation or probability to hypotheses so that others, decision makers, can decide as to the truth or falsity of the beliefs" (94). A less strong version—"the 'truth seeker' view"—holds "that the scientist, though himself a decision maker, seeks only one goal, has only one overriding value: truth at any cost" (p. 95). Finally, the weakest version of value neutrality—the "'cognitive decision maker' version" (p. 95)—views "science as a quest not only for true hypotheses but also for ones with additional cognitive, theoretical, or epistemic values" (p. 95).

Leach thinks the cognitive decision maker version of the value neutrality thesis is the "most cogent" (p. 95), and it appears to be the version Alexander endorses insofar as he includes activities like exploration and explanation in his account of the aims of science. To argue against this view, Leach defends "the 'practical decision maker' view of science by arguing that the scientist in the logic of his inquiry also makes pragmatic value judgments in pursuit of practical goals" (p. 95). Leach's defense of the practical decision maker view turns on a reformulation of Richard Rudner's (1953) inductive risk argument in light of criticisms from Isaac Levi (1960). Levi argues Rudner's argument from inductive risk assumes too strong a connection between belief and action. Namely, Levi thinks the argument works only if accepting a hypothesis "is equivalent to choosing to act on the basis of H relative to some specific objective P" (p. 348). He questions this equivalence by providing cases where one might accept the truth of a hypothesis, but be unwilling to act on it. Importantly, his argument parallels arguments for value neutrality that assume value judgments are not entailed by empirical statements. In this case, Levi denies a connection between theoretical belief (or acceptance) and practical action.

However, Leach thinks the inductive risk argument need not assume *equivalence* between accepting a hypothesis and acting on the basis of that hypothesis. Instead, he claims, "The only conditions required for the argument are that belief be logically tied to action in a such a way that beliefs can be *partially justified* by practical considerations of action, and that different mistakes in action be taken with different degrees of seriousness" (1968a, pp. 100–101). Making this logical

connection involves positing a relationship between accepting a hypothesis, having a disposition to act on the basis of that hypothesis relative to some practical goal, and acting in the appropriate circumstances. For Leach, practical considerations related to whether one is warranted, in an ethical sense, in acting in the appropriate circumstances on the basis of a hypothesis are relevant to assessing whether that hypothesis is rationally acceptable. On his view, "the reasonableness or rational acceptability of H relative to P [rests] not only on whether H is true or has high epistemic utility, but also upon its pragmatic value, upon whether or not action based upon H is warranted" (p. 102). Thus, Leach posits an "epistemic interdependence of acts, dispositions and beliefs" that need not be as strong as the equivalence relation Levi thinks the inductive risk argument needs (p. 102).

How is this related to metaethics and Alexander's argument? First, it pushes back against Alexander's claim that value judgments "are not *independently* relevant for scientific confirmation" (1974, p. 397). If there are connections or an interdependence between beliefs, dispositions to act, and actions, then value judgments are relevant to confirmation. Second, for Leach, the argument from inductive risk "forces a reconsideration…of the relations between theoretical, technological and policy making aspects of rational inquiry" (1968a, p. 107). This raises questions about the connection between belief and action related to metaethical disputes about the nature of fact, value, and the scopes of theoretical reason and practical reason. When it comes to the picture of rational inquiry advanced by proponents of the argument from inductive risk, Leach thinks "No less at issue are the embarrassingly complex conceptual interrelationships between 'belief', 'value', 'action', 'accepting a belief or hypothesis', 'acting upon a belief' and 'reasonable belief'" (1968b, p. 366). And for him, it is "Almost inconceivable…that an adequate analysis of 'belief' could be provided which is logically *un*related to 'action'. For there would then seem to be no way to account for the important influence of beliefs on action" (1968a, p. 102). Thus, Leach posits a connection between belief in a hypothesis and acting on that hypothesis that Alexander rejects. Value judgments are relevant to the aims of the scientist to describe and explain because descriptions and explanations are connected to action, and value judgments are needed to determine whether or not an action is reasonable. In turn, judging an action based on some hypothesis reasonable or unreasonable is relevant to confirming said hypothesis, e.g., in determining the amount of evidence we demand to accept it.

In effect, Leach's defense of the argument from inductive risk suggests Alexander cannot focus on the *product* of scientific investigation without also looking at the *process*. If the argument from inductive risk goes through, value judgments relative to the costs of acting on an erroneous hypothesis are relevant to whether one's acceptance of that hypothesis is rational. I follow Leach in thinking that rejecting this connection requires substantive metaethical assumptions about the

connection between the domains of theoretical reason (belief) and practical reason (action). So, while Alexander's argument for the value neutrality of the sciences does not follow Hempel, Feigl, and Nagel in taking a controversial stance on the cognitive status of value judgments that recalls Stevenson's emotivism, it is not completely metaethically innocent.

6. Conclusion

The history I sketched supports Brown's (2020) and Leach's (1968a) historical suggestions about the relationship between metaethical noncognitivism—especially as formulated by Stevenson—and certain arguments for the value neutrality of the sciences—especially Feigl's, Hempel's, and Nagel's. I also think Alexander's (1974) attempt to formulate a role-based, value neutrality thesis independent of metaethical claims prompts us to consider connections between the argument from inductive risk and metaethical questions. Drawing upon Leach (1968a; 1968b), I argued that Alexander's intervention raises further questions about the connections between theoretical reason and the generation of belief and practical reason and action. Importantly, Brown pursues related questions about the empirical status of value judgments, as well as the relationship between epistemic values, belief formation, and acts of assertion. In doing this, he explicitly eschews any "deep metaethics" (2020, p. 90 n.1). Yet, I think the metaethical issues touched on in the history I sketched, as well as other related issues, merit revisiting in light of recent calls and attempts to more fully characterize the nature and role of values in 'values in science.'[15]

In closing, I note the narrow focus of my historical sketch. Metaethical noncognitivism stands in stark contrast to the pragmatist tradition Leach (1968a, pp. 107–108) hints at adopting and Brown (2020) explicitly works within. In this tradition, given the tight connection between knowledge, valuation, and action, there is no hard distinction drawn between fact and value like there is for Stevenson. For many pragmatists, value judgments are a type of empirical statement. This is also true of some work in the feminist, pragmatist tradition Brown draws upon in developing his view, e.g., Lynn Hankinson Nelson (1990), Elizabeth Anderson (2004), and Sharyn Clough (2011). Note, too, that the metaethical characterization of value judgments I examined are distinct from Thomas Kuhn's discussion of value judgments and theory choice. Kuhn (1977) glosses a value

[15] See also Scriven (1972), Dan Hicks (2014), Brown (2020), Ward (2021), and Wieten (2021). An anonymous reviewer also mentions historical antecedents and related issues beyond the scope of this paper. These include David Hume's claim that passions, rather than belief, are needed to motivate action, and whether actions, which are nonpropositional, can ever be conclusions of arguments. I agree it could be fruitful to look into how these issues relate to arguments for and against the value-free ideal.

judgment as a nonalgorithmic decision made on the basis of commitments to certain theoretical virtues.[16,] While scientists share the same commitments, different scientists might reasonably disagree on how to bring them to bear on theory choice. Given Kuhn's influence, as well as recent interest in pragmatist philosophies of science, a more complete history of the 'values' in 'values in science' would compare and contrast the noncognitivist, pragmatist, feminist, and Kuhnian treatments of value judgments.[17]

[16] Thanks to P.D. Magnus for a discussion of this point.

[17] Of particular interest to a more complete history is the robust debate in feminist philosophy concerning the empirical status of value judgments. For a recent overview of this debate as it relates to science and values, see Christopher ChoGlueck and Elisabeth Lloyd (2023). See Audrey Yap (2016) for critical engagement with Nelson, Anderson, and Clough on the empirical status of value judgments. Kristen Intemann (2001) is also of interest since she pushes back against those who would conflate 'empirical' and 'descriptive.' For Intemann, value judgments might be empirical even if not descriptive. Thanks to Jer Steeger for helpful conversations related to points in this note.

References

Alexander, James. 1974. "Metaethics and Value Neutrality in Science." *Philosophical Studies*. Vol. 25, No. 6: 391-401.

Alexandrova, Anna. 2018. "Can the Science of Well-Being Be Objective?" *British Journal for the Philosophy of Science*. 69: 421–445.

Anderson, Elizabeth. 2004. "Uses of Value Judgments in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce." *Hypatia* 19, no. 1:1–24.

Ayer, A.J. 1946/1952. *Language, Truth and Logic*. New York: Dover.

---- (ed.). 1959. *Logical Positivism*. New York: The Free Press.

Brown, Matthew J. 2020. *Science and Moral Imagination: A New Ideal for Values in Science*. Pittsburgh: Pittsburgh University Press.

Carnap, Rudolf. 1932/1959. "The Elimination of Metaphysics Through the Logical Analysis of Language." *Logical Positivism*. A.J. Ayer (ed.) New York: The Free Press.

----. 1935. *Philosophy and Logical Syntax*. London: Kegan Paul.

ChoGlueck, Christopher and Lloyd, Elisabeth. 2023. "Values as heuristics: a contextual empiricist account of assessing values scientifically." *Synthese*. 201, 220.

Clough, Sharyn. 2011. "Radical Interpretation, Feminism, and Science." *Dialogues with Davidson*, edited by Jeffrey Malpas. Cambridge, MA: 405–26.

Damböck, Christian. 2022. "The Politics of Carnap's Non-Cognitivism and the Scientific World-Conception of Left-Wing Logical Empiricism." *Perspectives on Science*. Vol.30, no. 4: 493–524.

Diggin, Steven. 2022. "Ethical Evidence." *Synthese*. 200:329: 1–24.

Douglas, Heather. 2009. *Science, Policy, and the Value-Free Ideal*. Pittsburgh: Pittsburgh University Press.

Feigl, Herbert. 1946. "Review of *Ethics and Language* by C.L. Stevenson." *Philosophy of Science*. Vol. 13: 1: 80.

----. 1950. "The Difference Between Knowledge and Valuation." *Journal of Social Issues*. Vol. 6: 4: 39-44.

----. 1952. "Validation and Vindication: An Analysis of the Nature and the Limits of Ethical Arguments." *Readings in Ethical Theory*. Wilfrid Sellars and John Hospers (eds.). New York: Appleton-Century-Crofts: 667–680.

Foot, Philippa. 1958. "Moral Arguments." *Mind*, Vol. 67, No. 268: 502-513

Hempel, Carl G. 1960/1965. "Science and Human Values." *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press.

----. 1966. *Philosophy of Natural Science*. Upper Saddle River, NJ: Prentice-Hall.

Howard, Don. 2003. "Two Left Turns Make a Right: On the Curious Political Career of North American Philosophy of Science at Midcentury." *Minnesota Studies in the Philosophy of Science Volume XVIII: Logical Empiricism in North America*. Gary L. Hardcastle and Alan W. Richardson (eds.). Minneapolis: University of Minnesota Press: 29–93.

Hicks, Daniel J. 2014. "A New Direction for Science and Values." *Synthese*. 191: 3271–3295.

Intemann, Kristen. 2001. "Science and Values: Are Value Judgments Always Irrelevant to the Justification of Scientific Claims?" *Philosophy of Science*. Vol. 68, No. 3: Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers: S506-S518

Kuhn, Thomas. 1977. "Objectivity, Value Judgment, and Theory Choice." *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press: 320–39.

Leach, James. 1968a. "Explanation and Value Neutrality." *The British Journal for Philosophy of Science*. Vol. 19, No. 2: 93-108.

—-. 1968b. "Historical objectivity and value neutrality." *Inquiry: An Interdisciplinary Journal of Philosophy*, 11:1-4: 349-367.

Levi, Isaac. 1960. "Must the Scientist Make Value Judgments?" *The Journal of Philosophy*. Vol. 57, No. 11: 345–357.

Moore, G.E. 1903/1959. *Principia Ethica*. Cambridge: Cambridge University Press.

Nagel, Ernest. 1961. *The Structure of Science: Problems in the Logic of Explanation*. New York: Harcourt, Brace, and World.

Nelson, Lynn Hankinson. 1990. *Who Knows: From Quine to a Feminist Empiricism*. Philadelphia: Temple University Press.

Neuber, Matthias. 2022. "Herbert Feigl." *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.): URL: https://plato.stanford.edu/archives/win2022/entries/feigl/

Ogden, C.K. and Richards, I.A. 1923/1945. *The Meaning of Meaning*. (7th edition). New York: Harcourt, Brace and Company.

Reichenbach, Hans. 1951. *The Rise of Scientific Philosophy*. Berkeley: University of California Press.

Rudner, Richard. 1953. “The Scientist Qua Scientist Makes Value Judgments.” *Philosophy of Science*. Vol. 20, No. 1: 1–6.

Scriven, Michael. 1972. “The Exact Role of Value Judgments in Science.” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*: 219-247.

Siegetsleitner, Anne. 2022. “Noncognitivism.” *The Routledge Handbook of Logical Empiricism.* Thomas Uebel and Christoph Limbeck-Lilienau (eds). London: Routledge: 168–175.

Stebbing, Susan. 1930/1942. *A Modern Introduction to Logic*. London: Methuen & Co.

Stevenson, C.L. 1937. “The Emotive Meaning of Ethical Terms.” *Mind*. Vol. 46, No. 181: 14-31.

—-. 1948/1963. “The Nature of Ethical Disagreement.” *Facts and Values: Studies in Ethical Analysis*. New Haven: Yale University Press: 1-9.

Vaesen Krist and Joel Katzav. 2019. “The national science foundation and philosophy of science's withdrawal from social concerns.” *Studies in History and Philosophy of Science*, 78: 73-82.

Ward, Zina. 2021. “On value-laden science.” *Studies in History and Philosophy of Science, Part A*. 85: 54–62.

Wieten, Sarah. 2021. “Review of *Science and Moral Imagination* by Matthew J. Brown: Implications for values in medicine.” *Studies in History and Philosophy of Science Part A*. 89: 299-300.

Yap, Audrey. 2016. “Feminist Radical Empiricism, Values, and Evidence.” *Hypatia*. 31, no. 1: 58-73.