



Holism about fact and value

Kenneth Walden

Department of Philosophy, Dartmouth College, Hanover, NH, USA

ABSTRACT

This paper argues for confirmational holism about facts and values. This position is similar to one defended by (among others) Hilary Putnam, but the argument is importantly different. Whereas Putnam et al. rely on examples of the putative entanglement of facts and values – a strategy which I suggest is vulnerable to parrying – my argument proceeds at a more general level. I argue that the explanation of action can not be separated from our practical reasoning, and for this reason, the 'webs' of value and fact judgments are joined in the same way that Quine holds the judgments of mathematics and natural science to be.

ARTICLE HISTORY Received 29 September 2021; Accepted 25 July 2022

KEYWORDS Metaethics; facts; values; holism; action explanation; Putnam

'Two dogmas of empiricism' has two projects, one negative and one positive. The negative project is an attack on Carnap's distinction between internal and external questions, which Quine mounts with an argument against the analytic/synthetic distinction. The positive project is to sketch an alternative picture of the dynamics of inquiry, a picture of science without Carnap's dichotomy. In place of Carnap's bifurcation of science into fact and convention, Quine gives us a holistic science suffused with both. Knowledge, he says elsewhere, is

a fabric of sentences [which] develops and changes, through more or less arbitrary and deliberate revisions and additions of our own, more or less directly occasioned by the continuing stimulation of our sense organs. It is a pale grey lore, black with fact and white with convention. But I have found no substantial reasons for concluding that there are any quite black threads in it, or any white ones.¹

When joined with a mild form of methodological naturalism, this picture of inquiry leads to a distinctive approach to metaphysics, epistemology, and semantics. It does not entail particular first-order doctrines in these areas so much as prescribe a uniformity of regard. Once we appreciate this holism, the thought goes, we see that there are no good grounds for distinguishing between different parts of the 'web of belief' – for saying that sector A of the web is substantive, while sector B is merely conventional, that sector C aims to represent the world while sector D is merely expressive, or that the objects posited by sector E really exist, while those of F are mere props.

The goal of this paper is to argue for holism about fact and value – to argue that judgments of fact and judgments of value are part of the same 'web' in more or less the sense advanced in 'Two dogmas' and for this reason deserve the same uniformity of regard.

1. Fact/value holism

The crucial concept in Quine's argument is a kind of holism about confirmation. I will understand Quine's notion of confirmation as an epistemic notion – as a property about what we ought to believe.² The basic idea is simple enough. Someone is murdered, and we start concocting hypotheses about how the deed was done. These hypotheses are complicated in part because they have many component parts that stand in a network of complicated dependence relationships.

The bloody glove in Jones's glovebox is evidence that Jones did it, unless the murderer used poison. The vial at the scene is evidence the murderer used poison unless it was left as a red herring. It's unlikely that the vial was left as a red herring unless the maid is in on it. But the maid is innocent unless she's not who she claims to be.

Taken together these connections suggest that our theory of the crime should constitute a holistic web of beliefs: a set of hypotheses that mutually support each other in such a way that none of them is warranted individually but only as part of a 'corporate body'. So holism amounts to a kind of epistemic interdependence. If we entertain talk about credence or degrees of belief, we can put the idea thus: X is holistic if and only if for each p in X, what credence one ought to assign p depends on what credence one assigns to every other member of X.

²And thus not a semantic view like the thesis that we cannot distinguish fact judgments from value judgments in respect of meaning in the style discussed by Fodor and Lepore (1992, 37-59).

Ouine holds that the natural sciences are holistic in this sense. Our credence in a proposition about an electron in a bubble chamber will depend on what we believe about numerous parts of the natural world because that world is an integrated whole, and the bubble chamber is just one small part of it. Even if a particular proposition is not likely to make a significant difference to our credence in the proposition about the electron – say, because it's about sperm whale migratory patterns – it nonetheless matters in principle because the two propositions are part of the same comprehensive explanation. We can imagine, for example, that a recalcitrant result about sperm whale migration leads us to rethink our theory of sperm whale metabolism, which in turn leads us to rethink our theory about the chemistry of large organic molecules, which in turn affects our theory of chemical bonding, which then leads us to refashion our standards for electron detection. That we should not altogether exclude such a possibility means that we should concede the judgments about bubble chambers and sperm whales are in principle part of the same holistic web.

There is a natural way of understanding the fact/value dichotomy in terms of holism. To state it we need to generalize our definition a little. We need to generalize it so that it includes not just beliefs but other ways of judging as well. The generalization looks like this: X is holistic if and only if for all p in X, the degree to which one should judge that p depends on the degree to which one judges every other member of X. Because I want to be as ecumenical as possible about what a practical, normative, or value judgment consists in, this definition is designed to accommodate different views on that subject. If you think that a value judgment like 'Humphrey is the best choice' or 'I ought to campaign for Humphrey' is a belief, then the degree of judgment invoked in the definition is just a credence. If, however, you think it's a conative state, then it may be something like a strength of desire or degree of intention.³ I want to be neutral between these possibilities. The definition can also accommodate those who doubt the very idea of partial judgment. These people just believe that there are two possibilities: judging that one ought to help Humphrey and failing to so judge. The definition does make one important assumption: insofar as it invokes a claim about what one should judge it assumes that this attitude of judgment is itself evaluable.

³E.g. in the sense Holton (2009, ch. 2).

According to the conception of the fact/value dichotomy I have in mind, there is a body of 'pure' practical, value, or normative judgments V and a body of theoretical, factual, or non-normative judgments F that are independent of each other in a way that amounts to the denial of the holism of F + V. That is, the question of to what degree we should judge a member of V does not depend on our judgments concerning any member of F nor vice versa. (Assume for the sake of argument that we know how to identify the members of V and F, and bracket questions about just which judgments they are. Later I will suggest that even this preliminary sorting may prove impossible.) If we assume that these two bodies of judgments are themselves holistic, then we have two holistic webs, a value web and a fact web. Importantly, each of these webs may have downstream implications for a common set of hybrid judgments. Combining 'I ought to avoid danger' and 'radioactive cockroaches are dangerous' yields the hybrid judgment, 'I ought to avoid radioactive cockroaches'. Nonetheless, F and V themselves are isolated from each other because the entailment that produces 'I ought to avoid radioactive cockroaches' only goes in one direction: whether one should judge this depends on one's judgments about F and V, but not vice versa. On this picture, no facts affect which 'pure' value judgments we accept - or rather, the degree to which we judge them – nor vice versa.⁴

I think this is one reasonable gloss on the fact/value dichotomy. It codifies the intuition that facts and values occupy separate, autonomous realms, and spells out what we mean when we say that what goes on in one realm does not affect the other. It is not the only gloss, of course. We could, for example, construe the dichotomy logically, as a claim about the impossibility of certain inferences between claims that do and do not contain characteristically normative vocabulary. Or we could understand it metaphysically, as a claim about the ultimate 'grounds' of different propositions. One advantage of the epistemic version is that it depends on a relatively familiar notion that we have plentiful experience with. We have lots of practice thinking about what matters in deciding whether to form a certain belief or adopt a certain value judgment. On the other hand, metaphysical versions of the dichotomy depend on a more recondite relation (for instance, grounding) whose very well-

⁴Cohen (2003) for a view similar to this version of the dichotomy, though note that Cohen concedes that if a general epistemic picture like the one Quine champions is true, then one of the premises for the strongest version of this view doesn't hold (224).

⁵As Hume famously does in the *Treatise* III.1.1.27.

⁶Maguire (2015).

foundedness is a matter of controversy. And debates about the version of the dichotomy built upon the alleged inferential gap between different kinds of propositions usually end up being more about the nature of inference and synonymy – and how to formulate the essential 'gap' claim in a way immune to obvious counterexamples – than about any crucial differences between facts and values or the normative and non-normative. Moreover, we might think that whether or not we should accept one of these other versions of the dichotomy is not obvious, but something we discover in the midst of our other inquiries, and how those inquiries turn out will depend, in part, on whether they have the kind of structure the epistemic version of the dichotomy holds. For instance, we might believe that if the epistemic version of the dichotomy fails, then we could not be justified in believing that a metaphysical version holds. These are not decisive reasons to favor the epistemic version by any means, but I hope they justify me in making it my target.

My argument against this dichotomy is modeled on one of Quine's arguments. Not the argument of 'Two dogmas', but on Quine's argument for holism about mathematics and natural science. This argument is a bit scattered and obscure, so I am going to take an interpretative liberty and simply assert that Quine's argument for this holistic thesis is an instance of the following schema. It is the instance for A = physical science, B = mathematics, and P = general empirical phenomena.⁸

- (1) A-judgments are indispensable to explaining phenomena P. (First indispensability premise.)
- (2) B-judgments are also indispensable to explaining phenomena P. (Second indispensability premise.)
- (3) These explanations are holistic: which A-judgments will contribute to the best overall explanation of P depends on which B-judgments we include in our explanation and vice versa. (Holism of explanation.)
- (4) A judgment's being included in the best overall explanation of a phenomenon should affect the degree to which we make that judgment. (Inference to the best explanation.)
- (5) Therefore, the degree to which we make A-judgments depends in part on the degree to which we make B-judgments, and vice versa. (From the holism of explanation plus inference to the best explanation.)

⁷E.g. Restall and Russell (2010).

⁸A better elaboration and defense of this argument can be found in Colyvan (2001).

(6) Therefore, A and B constitute a holistic web.

In this paper, I present another instance of this argument schema. In my argument A is natural science, understood broadly to include the human sciences and much of what we call common sense. A will therefore consist of descriptive, theoretical claims about the world. B will be what we could variously call *normative* or *practical* or *value* claims – claims that do not just describe the world but evaluate things and prescribe actions. And P is the particular phenomenon of human action.

2. The context of the argument

I have some antecedents in this campaign, and so before setting out on my argument, I owe a brief explanation of what I find wanting in their accounts and why I think it is worth giving mine.

Reductionist naturalism entails holism about fact and value. If all value properties are identical to natural properties, which are presumably at home in our web of factual judgments, then value judgments are also part of that web. But the argument for this view works rather differently from mine insofar as it depends on the specification of reductions of value properties to particular natural properties through, for example, Railton's 'reforming naturalistic definitions of non-moral goodness and moral rightness' or Boyd's synthetic identities. The argument for holism does not depend on such reductions or identifications. The argument is not that value properties are identical to recognized natural properties, but that our best criterion of naturalness – inclusion in our most explanatory theory – *already* countenances value properties. As a result, the present argument eludes some of the classic problems facing this reductionist naturalism.

Hilary Putnam offers an argument more in the spirit of the present paper, but it isn't very compelling. He begins with a genealogy of the fact/value dichotomy. Our intuitive understanding of the theoretical/practical or fact/value dichotomy is inherited from Hume's idea that facts must be intimately connected to sense experience. More proximately it derives from the modern elaboration of the Humean idea by logical empiricists like Carnap. But Carnap's program proved unworkable. Quine's argument against the analytic/synthetic distinction showed that we could not maintain a rigid separation between the content of a theory and its

⁹Railton (2003, 32) and Boyd (1988).

background framework, and the underdetermination of theory by evidence shows that something intuitively 'non-factual' will inevitably find its way into our theories.¹⁰

Of course, these reflections do not show that the dichotomy is untenable, just that one motivation for it fails. We need something more for an argument against the dichotomy itself. Here Putnam and his allies turn to examples of the 'entanglement' of fact and value. There are many of these. Helen Longino, for example, argues that there is no value-free notion of 'evidence' that is capable of performing the methodological tasks we put to it, so we have to rely on values in constructing theories. Nicholas Sturgeon offers more specific examples of explanations with moral terms in them (opposition to slavery arose earliest in the Angloand Francophone worlds because slavery in these areas was worse than in others) which can be used to support at least one direction of entanglement. Ruth Anna Putnam points to the value-ladenness of social science in general, while Elizabeth Anderson cites recent research into divorce in particular. Such research, Anderson says, is more successful than its rivals because it presupposes different value judgments about the family, and the success of these new theories of divorce amounts to a kind of confirmation of those values. Putnam himself offers an argument with a similar structure, only his example is the value-laden economics of Amartya Sen. Catherine Elgin says that the dichotomy founders on the indispensability of thick concepts in our projects, both ethical and scientific. And Morton White argues that there is no good reason to exclude emotional reactions from the range of things contributing to the acceptance of a theory, and once we allow this, we can justify the inclusion of value judgments in the holistic web on the ground that theories including them enjoy are better by this liberal standard than those without.¹¹

I am sympathetic to all of these entanglement arguments, and I believe they lead us to more or less the same holism that I favor. Nonetheless, I don't think they are fully convincing on their own. They offer what we might call a 'low road' to holism. They marshal examples to show that there is entanglement between theory and practice, fact and value, but they do not give deep reasons for why there must be this kind of entanglement. This makes them vulnerable to parrying. If we reconstruct these examples, critics may say, we find that the practical content we thought was important is mere gilding. For example, they may say that Sturgeon's

¹⁰This is the argument of the first chapter of Putnam (2002).

¹¹Longino (1990), Sturgeon (1988), Anderson (2004), Putnam (1987), Elgin (1997), White (1981), and White (2002, 153-178).

'worse' is the most convenient way to pick out a state of affairs that is explanatorily relevant, but that this does not mean that normativity of the word itself explains anything. The evaluative content is a nomological dangler. They could charge that the examples of value-laden scientific work conflate the context of discovery and the context of justification. They could dispute the gloss that Anderson and Putnam put on their examples, saying that the success of, e.g. new research on divorce is due not to its countenancing different values but to some other innovative feature of that work. Or they could maintain that these value judgments are helpful auxiliaries for working scientists but not strictly speaking indispensable, and so should be paraphrased away in the final analysis. Critics may also note that thick concepts can change their valence or lose it altogether (as 'chaste' and 'pious' seem to have) and use this fact to argue that such concepts can always be factored into a fact component and a value component. And, finally, they can simply reject the sort of liberalization that White proposes. Maybe Quine had reason to accept such a liberalization, I can imagine them saying, but that doesn't mean we do. 12

I do not visit these debates here. What I want to offer is a supplement to this work, what amounts to a 'high road' argument for holism about fact and value. Rather than an argument proceeding from putative examples of entanglement, I offer an argument that shows that a pervasive entanglement must follow from the relationship between explanation and action. The basic thought behind the argument goes like this. Quine's argument about mathematics is premised on the idea that which judgments we should countenance depends on what we call on to explain the phenomena we encounter. But explanation is a situated undertaking. As Quine himself explains, 'science ventures its tentative answers in man-made concepts, perforce, couched in man-made language, but we can ask no better. The very notion of object, or of one and many, is indeed as parochially human as the parts of speech'. So what counts as a good explanation for us clearly depends on features of us – on our cognitive architecture, the beliefs, hypotheses, and background principles we have going into the project of explanation, the concepts we employ, maybe even the social contexts in which explanations are offered. Now, true enough, it is a goal of inquiry to make our explanations as unparochial as possible, to give explanations that would be valid for agents who were not situated as we are. But we should not

¹²For versions of these critiques see Smart (1999), Quinn (1986), Solomon (2012), and Blackburn (2011).

confuse this regulative ideal with a constitutive requirement. If we did, there would be no good explanations. Mathematics is indispensable to our scientific explanations, but probably not to those that we could share with Laplace's Demon. But this does not make them bad explanations or vitiate our commitment to mathematics. As Quine continues: 'to ask what reality is really like, however, apart from human categories, is self-stultifying. It is like asking how long the Nile really is, apart from parochial matters of miles or meters'. 13

One kind of situatedness plausibly lies behind the explanatory usefulness of mathematics. Our being so limited – computationally, spatially, temporally - means that any chance we have of understanding the world in any generality will depend on our ability to mark off patterns that emerge, and mathematics is the science of patterns. A different kind of situatedness leads to the holism of fact and value: explainers of actions are also performers of actions. Given that these two endeavors – performing actions and explaining them – concern exactly the same objects, our default assumption should be that these two projects will become bound up with each other. We should assume, that is, that thinking about how to act will affect our thinking about why agents acted as they did, and vice versa. Moreover, we should expect that severing them will make our explanations worse in roughly the same way (though not necessarily to the same degree) as denying ourselves mathematics in the natural sciences: they will be worse because what we will end up with will be less intelligible to the sorts of creatures we are than they could have been. This, anyway, is the inchoate idea motivating the argument that follows. Each of the following sections introduces and defends one of the premises in the argument summarized above.

3. Premise 1: the indispensability of fact judgments

The first premise of my argument is that any satisfactory and complete explanation of an action will inevitably include some judgments of fact - or, slightly differently, that there will always be some judgment of fact that is relevant to a question of the form, 'why did $A \phi$?' Let's work with an example. We have an agent; call him The Butler. One evening we see him rummaging through the papers on the desk of his master,

¹³Quine (1992).

Lord Musgrave. Why does this The Butler do this? And, more philosophically, what sorts of considerations are relevant to explaining this action?

The Butler's rummaging is a concrete event in a physical world ordered by causes and effects, so of course fact judgments about the character of that world and its causal order will enter into our best and fullest explanation of that rummaging. No conjunction of value judgments about what the Butler ought to do, what it would be sensible for him to do, or what it would be good for him to do will determine his doing a thing absent judgments linking these to concrete events in the world – absent fact judgments. So we will inevitably find such judgments in our explanation.

4. Premise 2: the indispensability of value judgments

The second premise is that value judgments are indispensable to action explanation in the same way. This is, by a wide margin, the more difficult and controversial of the two indispensability theses. To defend it, I want to borrow some tools from Donald Davidson's theory of interpretation. Davidson maintains that the interpretation of actors and actions is constrained by a methodological precept that has come to be called a principle of charity. Davidson's most dashing expression of the principle is this:

In our need to make him [the object of explanation] make sense, we will try for a theory that finds him consistent, a believer of truths, and a lover of the good (all by our own lights, it goes without saying).¹⁴

If Davidson is right in asserting this as a kind of constitutive principle of interpretation, then premise 2 follows. For if we must interpret The Butler as a 'lover of the good', then the fact that something is indeed good (by our own judgment, it goes without saying) will be part of the reason that we see The Butler's rummaging in one light rather than another – as one kind of thing rather than another. And so the relevant value judgment will be part of the comprehensive explanation we offer when putting The Butler's rummaging in this light.

It is worth emphasizing how strong this claim is. A common thought against value-laden explanation is that while it may be that someone's holding a value judgment enters an explanation of her action, the value

 $^{^{14}}$ Davidson (2001, 222). This is just one formulation of the principle of charity. See Lepore and Ludwig (2007, 198ff) for a discussion of Davidson's different statements and defenses of the principle.

judgment itself does not. But what follows from the charity principle is the second thing: something's being good, sensible, or choiceworthy is part of our interpretation, not merely that the agent finds it so.

My argument for premise 2 is an argument that something like Davidson's charity principle is correct. (Though what I need and ultimately argue for is different and probably weaker than Davidson's own version of the principle.) I want to suggest that it is a stricture of the interpretation of agents that we try to render them, as far as possible, as acting for genuine reasons, responding to things of real value, and following binding rules. If this is right, then our judgments that certain things are indeed reasons, valuable, or binding will be part of our explanation.

I can't claim to fully understand Davidson's argument for the principle of charity, so I will try to cobble one together on my own. It goes like this.

- 2.1 We must adopt some methodological precepts in our explanatory activities to overcome the underdetermination of explanation by our evidence.
- 2.2 Insofar as we are trying to explain actions, these precepts should aim to render an event intelligible as an action.
- 2.3 What we find intelligible as an action will depend on our own theory of action.
- 2.4 This theory will include value judgments.
- 2.5 Therefore, our methodological precepts in explaining action will make reference to our own value judgments.
- 2.6 Therefore, such value claims are indispensable to the explanation of action.

For any body of evidence about someone's behavior there are innumerable competing explanations about that person and why they do what they do. If we are going to have any chance of understanding what an actor is up to, we need some principles for sorting through these rival explanations.

For all we are told about The Butler in my little vignette we could postulate explanations centered on such claims as these:

- (1) The Butler is searching for papers containing a clue about where a treasure is hidden.
- (2) The Butler wants to go up in the aliens' spaceship and believes that rummaging is a way to signal them.

(3) The Butler wants to know how many tokens of the letter 'k' are in Musgrave's papers and is counting them in an unconventional way.

Indeed, strictly speaking, absent principles guiding our explanation, we could offer a story like this to explain The Butler's rummaging:

(4) The Butler felt a bit ripe, and he thought a shower would fix him up nicely, so he decided to rummage through some papers.

These are all strictly compatible with the evidence we have about The Butler, as are a mountain of other stories. Some of them make The Butler look like rather an odd fellow, but that's exactly the point. We need some principles about the explanation of action that will motivate the demotion of these stories relative to more plausible ones. And this is what premise 2.1 claims.

Perhaps we need some a priori principles to overcome underdetermination, but why Davidson's charity principle in particular? Why suppose that when we interpret The Butler we ought to prefer explanations that portray him as a 'lover of the good', rather than a lover of the bad, the weird, or the Chicago Cubs?

When we are trying to explain something, we are trying to render it intelligible as the kind of thing it is, and this goal will be one source of principles of explanation like those we are interested in. For example, when I want to explain why carbonic acid results from carbon dioxide being dissolved in water, I try to render this event intelligible as a chemical reaction. There are certain things it takes to be a chemical reaction, and these standards will guide my attempts at explanation: chemical reactions involve reagents and products, the transfer of electrons, and the exchange of energy. Insofar as these are part of our very idea of a chemical reaction, they will function as methodological precepts of my explanatory endeavor. We may discover that some event simply doesn't fit the template and try to explain it in some other way, but so long as we are trying to explain it as a chemical reaction, we will be guided in that explanation by a conception of what it means to be a chemical reaction. So it should also be with the explanation of action. To explain an action is to render it intelligible as an action. This is what premise 2.2 claims.

But what does it take to render something intelligible as an action? At this point we could assert some essentialist claim about action, e.g. that action always aims at a particular end. We could then argue that this

feature should be reflected in our methodological precepts of interpretation. This isn't the way I want to go, though. Any such claim would be controversial, and I don't think we need it anyway. 15 We don't need it because each and every person already has a conception of what makes an action intelligible that is acquired from their understanding action from the inside. Creatures like us have an understanding of what action is, and thus of what it means for something to be intelligible as an action because we act. This understanding need not consist in knowledge of what is essential to or constitutive of action; it need not even be particularly systematic. Rather, it is a vast network of individual judgments made in the midst of practical reasoning. For example: Given the aim of finding the treasure and the possibility of being hindered in doing this if one is found out, one ought to search for the treasure in secret. There is really no good reason to be interested in counting the number of tokens of 'k' in a stack of papers. It would make no sense at all to go rummaging through papers because you have decided you need a shower. These are workaday value judgments, and they form part of our 'theory' of action in the simple sense that they are judgments about action.

Importantly, these judgments are both structural and substantive. That is, they include both schemas about how an agent's attitudes ought to fit together - whatever those attitudes happen to be - and judgments about which particular actions are appropriate when, which actions make sense under which particular circumstances, which particular ends are sensible and senseless, etc.¹⁶

Such judgments naturally inform what we find intelligible in the way of action. To see why, return to the analogy with chemistry. When we are explaining the formation of carbonic acid, we are trying to render this event intelligible as a chemical reaction. But we don't come to this conception through some clear and distinct idea about chemical reactions. It is an empirical standard that evolves alongside the rest of our chemical theory. The modern conception of a chemical reaction originated with Lavoisier and was formulated in response to the empirical failures of the previous conception, the idea that many substances gained weight when burned, that air is necessary to so many reactions, that air can be 'dephlogisticated'. As much as our conception of a chemical reaction

¹⁵For the controversy see Millgram (2010).

¹⁶See Williams (2018) for more on what constitutes 'substantive' radical interpretation as well as an argument that the stance can be used to explain the referential stability of the concept moral wrongness in a way that is neutral between metaethical theories.

guides our theorizing, that conception also depends on the judgments reached in theorizing – on the workaday results from the laboratory. So there is a reciprocal relationship between chemical theorizing and our conception of a chemical reaction. The same should be true mutatis mutandis for what we find intelligible in the way of action. And this is what 2.3 says: our conception of intelligible action will depend on the multitude of particular judgments – both structural and substantive – that we make about action, just as our conception of a chemical reaction will depend on what we find in the laboratory.

Only here this multitude will include judgments about which actions make sense when, what it is reasonable to prefer, which rules are rational to follow, etc. It will therefore include value judgments that we would ordinarily associate with the faculty of practical reason. This is premise 2.4. The most obvious example of this influence involves the unintelligibility of actions premised on schemes of value that are radically out of sync with our own. This is why we find explanations like (3) and (4) inadequate. The idea that someone would see needing a shower as a reason to rummage through papers on a desk is so strange to us, so clearly nonsense that we cannot feel that we have rendered The Butler's rummaging intelligible as an action by giving such an explanation. By the same token, though perhaps to a lesser degree, by attributing such a bizarre fixation to The Butler as discovering the number of 'k' inscriptions in a stack of papers, we feel like we have not really made his activity intelligible – we judge that this fixation makes no sense - and that's why, other things being equal, we prefer a different explanation.¹⁷

Putting these premises together yields our conclusion. Particular value judgments made in the midst of our own practical deliberations are relevant to the explanation of action because they inform our conception of what is intelligible in the way of action. (In the same way, I have suggested, that individual judgments made in the midst of chemical theorizing will inform the chemist's ideas about what is intelligible as a chemical reaction.) And rendering an event intelligible as an action is a methodological precept that necessarily guides our explanation of action.

I just said that judgments about action made amidst our own deliberations will naturally inform our conception of what is intelligible as an action and will for this reason figure in our explanation of the actions of others. This does not mean that it is strictly impossible to quarantine

¹⁷There is a long line of philosophers who have recognized an evaluative dimension to rendering an agent intelligible, including Anscombe ([1957] 2000, §§5-11), McDowell (1998, 389), Nagel (1986, 142-143), and Dancy (2000, 94-97).

one's own value judgments from the business of explaining others' actions. I don't think it is impossible. But there is a substantial cost to doing so, and to complete the argument I want to give a sense of what that cost would be.

One potential consequence of the imagined severance would be the inability to eliminate zany-yet-coherent alternative explanations of an agent's action. For most explainers, value judgments are an essential point of friction for overcoming the underdetermination of our theory of behavior by the evidence available. Thus in The Butler's case we set aside potential explanations like (3) and (4), or at least demote them relative to others, because they do a poor job of making The Butler intelligible. We do this because they portray him as acting in blatant violation of fundamental norms of rationality or valuing the valueless. (Similarly, we set aside explanation (2) since it involves an uncharitable construal of The Butler's beliefs.) If we did not rely on such value judgments, we could not do this (not absent some other source of empirical friction), and the underdetermination we face in explaining action would be all the more acute. More generally, the costs of quarantine will approximate those of segregating the judgments of two different branches of science, for example refusing to employ what we know about chemistry in human physiology or mathematics in physics. Physiology would not grind to a halt without chemistry. And yet the additional leverage in our explanations provided by acknowledging that our bodies are made up of the same stuff that chemists study has paid tremendous dividends. The manifest phenomena of inheritance are compatible with numerous different theories of how traits are transmitted from parent to child. But only some of these are realizable by the substances that make up our bodies. Fewer still are compatible with what we know about nucleic acids in particular. What chemistry offers physiologists, then, is an increase in the points of theoretical friction a theory faces across the board. I have argued that relying on value judgments in the explanation of action offers the same benefit, and refusing to do so is not only ad hoc, it would entail a corresponding cost.

Before moving to the remaining premises of my overall argument, I want to say a few words in response to potential objections to premise 2.

First objection. Are value judgments the only thing that can do the work I have put to them? Mightn't we set explanations like (3) and (4) aside on other grounds, for instance, that they are statistically unlikely: people tend not to care about the number of 'k's printed on a page? My reply is that statistical regularities will often fail to provide completely satisfactory

explanations. I ignite borax on fire and it creates a green flame. Why does it do that? Noting that 98% of borax samples burn green doesn't help. We want to know what it is about borax that creates green flames (assuming there is such a property) and recording the statistical correlation doesn't tell us that. What we would find more satisfactory is an explanation that renders the emission of green light intelligible as part of an endothermic chemical reaction. The chemical explanation is better because it orients the reaction in a broader and more systematic picture of the world. By the same token, noting that 98% of people behave thus in these circumstances does not necessarily make that behavior intelligible as an action, or at least not as intelligible as we would like.

Second objection. Doesn't the argument falsely entail that it is impossible to explain agents as acting contrary to our (qua interpreters) values? No, it doesn't. The methodological goal of charity must be balanced against other constraints on the explanation of action: capturing the evidence, charity principles concerning true belief, and general explanatory desiderata like simplicity, fecundity, and the avoidance of ad hoc stipulations. So the argument does not entail that we cannot depict people as doing wicked things for crummy reasons; sometimes the weight of evidence will mandate such explanations. What the argument does entail is that, other things being equal, we should make the errors in valuing that we attribute to other people as shallow and localized as possible. And this seems appropriate. We will better understand the racist, the bully, and the petty criminal – we will be more apt to find what they do intelligible – if we see their vices as exceptions to a general pattern of overlapping values than we will if we jump to the conclusion that they are Satanic – that they have decided that evil is good and good is evil. Thus it is false that premise 2 forbids us from positing explanations involving errant valuing, and the role it does play in nudging us toward more charitable interpretations seems perfectly appropriate given the aims of explanation.

Third objection. I have been speaking quite generally of value judgments and their contributions to rendering an action intelligible. But one might worry that in doing this I am eliding over important distinctions within the web of value. For example, one might object that we can render an event intelligible as an action by showing why the agent who performed it would do such a thing given her beliefs, desires, and intentions. If this right, then charity only demands a preference for interpretations that render an agent structurally rational; it does not demand that we make substantive judgments about what it is good,

right, or sensible to value. (Similar objections could be made concerning different subsets of our value judgments. My reply to these will be similar.) To this objection, I have three replies. First, this objection only impedes my overall argument if we think that judgments about structural rationality are 'isolated', in the sense of our understanding of holism, from the web of value judgments. Second, there are simple cases where we are well-advised to portray an actor as structurally irrational if it allows us to avoid attributing some particularly ridiculous belief or odious desire. Fred is an ordinary fellow in every way, but one day he accepts a series of bets that turn out to be a Dutch book. Should we interpret Fred as disliking money? Enjoying the thrill of gambling even when he is guaranteed to lose? Of course not. We should say that Fred accepts the bet because suffers from a mild and limited bit of structural irrationality. Similar examples involving akrasia, compulsion, addiction, and all of humanity's other foibles are easy to find. These are reasons for rejecting the suggestion that the imposition of value judgments in our explanation of action is restricted to judgments of structural rationality. It is both impossible given the way that interpretation works and substantively misguided. Third, and most fundamentally, the objection assumes that we have already ascribed this battery of attitudes to the agent before we embark on explanation. But that is false. I ascribe the belief, 'I will get a ticket if I don't feed the meter' to Jones because I am trying to explain his meter-feeding. The explanation of action and the ascription of intentional attitudes are part of the very same interpretative project. It is true that for the explanation of individual actions we might treat certain intentional attitudes as fixed points, but this is a matter of relative entrenchment, not absolute priority. The predicament of the interpreter is never one of making sense of why someone did something *given* a comprehensive picture of her psychology. It couldn't be since that very picture is worked out in a tapestry of overlapping explanations of the actor's behavior. This means that the kind of charity that we employ in action explanation cannot be limited to structural rationality.¹⁸

5. Premise 3: holism about explanation

I have so far argued that both fact judgments and value judgments are indispensable to the project of explaining action. This does not yet give

¹⁸Compare Lewis (1983), Lewis (1992), and Williams (2018, 48–49) on the need for substantive radical interpretation.

us fact/value holism, for we have not shown that these explanations are themselves holistic. To get this conclusion we need to show that there is a give and take between fact and value judgments in these explanations, that we might ditch one fact judgment from our explanation and take on another in light of some realization about value, and vice versa.

Showing that both kinds of judgment are indispensable does not yet get us this conclusion because it is possible that the judgments contribute not to a single explanation, but to two parallel explanations. Compare the following case. Suppose that on Tuesday at eight o'clock the following event occurs: a uranium atom in New Mexico decays and a sperm whale in Nantucket surfaces for air. An adequate explanation of this spatially disjoint event must cite two kinds of judgment, judgments about atomic decay and judgments about the respiration of sperm whales. But these indispensability claims do not by themselves entail any kind of holism between our theories of atomic decay and sperm whale respiration. Explaining the conjunctive event means explaining each of its conjuncts and conjoining the explanations. There needn't be, and indeed oughtn't be, any revision of our sperm whale theory in light of the half-life of uranium or vice versa. So when we explain this conjunctive event, we are explaining its conjuncts in parallel, not in concert, and so the two theories that contribute to this explanation can remain completely isolated from each other.

My argument must rule out the analogous possibility for action explanation, which amounts to something like the following thought. Value judgments contribute to the explanation of one aspect of our actions: their intentional aspect. They are done for reasons, as part of a plan, and so on. Fact judgments contribute to a different aspect: the physical aspect. They involve the moving of various bodies in space and time controlled by physical processes. But we need only explain these two aspects in parallel, not necessarily in concert. We rely on value judgments to produce an explanation of the intentional aspects of an action and fact judgments to produce an explanation of the physical aspects, but there is no reason that what goes on in one of these explanations should affect what goes on in the other. The overall explanation of an action is no more than the conjunction of two autonomous subexplanations.

My reply is that the picture of action that this possibility relies on is incompatible with a very basic postulate that I will call the unity of action. Unlike our uranium-whale event, an action is something that is at once intentional and physical, not something intentional conjoined

to something physical. That is, an action is not just an attempt coinciding with a movement. It is an insoluble unity. This is why it does not make sense to ask what is left over when we subtract the intention from the movement.¹⁹ An explanation of an action must capture this unity. It must make sense of this deliberation corresponding to that movement - not merely explain each as independent events.²⁰

Ordinary action explanation accomplishes this by going back and forth between intentional explanation and physical explanation with the goal of reaching equilibrium between the two. Here are two stylized examples. Suppose we observe The Butler kick Lord Musgrave in the shin. Our first thought will be that The Butler kicked Musgrave to hurt him. But we know antecedently that The Butler did not bear any ill will toward Musgrave, so The Butler's kicking Musgrave to hurt him wouldn't make sense. The rejection of our initial hypothesis on the grounds that it fails as an intentional explanation then leads us to an alternative hypothesis about the physical features of The Butler's kick. We might think that it was caused by his patellar reflex, and so was quite involuntary. But we check The Butler's knee with our best scientific methods and find that it was not. So we offer another hypothesis: perhaps The Butler kicked Musgrave out of play. We hypothesize that this would make sense if he wanted to flirt with Musgrave. Life in the countryside can be boring after all. But we return to our physical measurements and find that The Butler's kick was far too ferocious to be flirtation. (Another value-laden conclusion.) So we start over again and hypothesize that The Butler meant to disable Musgrave in order to escape. But this raises further practical questions. And so on. Now suppose we observe The Butler sneaking into Musgrave's study. We are poised to explain this by saying that he wanted to rob Musgrave to pay off his gambling debts. This does make some sense given what we know about The Butler. But this is hard to square with some of the physical aspects of The Butler's action, like his going for old family documents instead of money. So we go back and revise our hypothesis: The Butler snoops through Musgrave's papers out of a prurient fascination with the lineage of well-bred men. But when we check this hypothesis against the facts gleaned from our theoretical methods we are disappointed, for we find that The Butler's skin conductance doesn't increase while he paws through these documents - usually a sign of

¹⁹Jaeger (1973).

 $^{^{20}}$ I said earlier that my argument does not rely on strong claims about the constitutive nature of action. I now must qualify that claim: the unity of action is such a claim. But it is such a basic axiom about action, and so central to our thinking about action that I hope it's not a point of controversy.

arousal. So we ask once more when it would make sense for The Butler to snoop through Musgrave's documents in absence of any intrinsic interest in them. We then come up with the thought that he hopes to find clues to the location of the treasure. And so on.

In both of these examples, our attempts at explanation oscillate between fact and value judgments, between the intentional side of action and the physical, trying to reach a stable equilibrium between the two. (In a fashion quite unlike how our uranium-whale explanation would go.) We are trying to reach a state where the intentional and physical pictures we paint are mutually supportive. We offer an intentional account, revise it in the face of some recalcitrant fact judgment, find this revised judgment wanting by the lights of our intentional account (which is informed by our value judgments) and so revise it again. We repeat this process, going back and forth, until our intentional and physical explanations are brought into equilibrium.

In ordinary action explanation (of which these are slightly laborious elaborations) we capture the unity of action by explaining holistically: by adjusting fact judgments in light of value judgments and vice versa until we reach a comprehensive equilibrium between both sorts of judgment that renders the event intelligible as a unified action. Here there is no special problem of the unity of action, since our explanatory stance never brooks the possibility of disunity. If we did not go about explanation in this way, however, unity would be a distinctive explanandum, and it is far from clear how we could explain it. Maybe we could offer a story about the pre-established harmony of the intentional and physical worlds or cook up a divine concurrence view. But these look like desperate and dark theories when compared to the ordinary holism of action explanation.

6. Premise 4: inference to the best explanation

There is one final gap between what we have shown and what we want to show. Holism of explanation does not necessarily entail the kind of holism we started with. To get that, we need a weak form of inference to the best explanation (IBE). To wit: p's figuring in our best explanation of some phenomenon should affect the degree to which we judge p. This is all I need for the argument to go through, but presumably the way in which this claim will be true is that p's being part of our best explanation will *increase* the degree to which we judge p (e.g. our credence).

I will assume that IBE is valid for fact judgments. This claim has its detractors, but these days I take it to be relatively uncontroversial.²¹ But why should we accept the principle for practical judgments about reasons, goods, and what is reasonable? We need to be careful about what form this resistance can take. No one who otherwise accepts IBE can seriously deny that giving a good explanation of The Butler that includes the claim that 'r is a reason to ϕ ' has some implications for what we judge about r and ϕ -ing. For instance, my judging that r is a reason to ϕ when explaining The Butler should make me more willing to cite that judgment when explaining the behavior of The Gardner in similar circumstances. In this respect, 'r is a reason to ϕ' is no different from an ordinary empirical fact. So the explanatoriness of the judgment is a reason to adopt the judgment that 'r is a reason to ϕ' – increase our degree of judgment – at least in the sense of being willing to deploy the judgment in other explanatory endeavors.

What a person can deny, however, is that this has any bearing on our practical reasoning about whether to ϕ . We might agree that a judgment's figuring in our best explanation should make us more confident in it in a 'theoretical' sense, but doubt that this bears on what we judge from a practical point of view. Why, such a critic would ask, should the explanatory utility of a claim like 'r is a reason to ϕ ' or 'q is good' or ' ϕ -ing is reasonable' have any effect on my taking r to be a reason to ϕ , q to be good, or ϕ -ing to be reasonable while deciding what to do?

This is the position I want to show is untenable. In effect, someone who wants to reject a practical version of IBE must intend to keep two books. For him there are 'theoretical' value judgments about what one ought to do, what is sensible, what is good, etc., that he puts to the theoretical purposes of categorization and the explanation. These are governed by IBE. And then there are 'practical' value judgments about what one ought to do, what is good, what is sensible, etc., that he comes to amidst deliberation. These guide his actions and are not governed by IBE. In a more tendentious tone, such a person might dub the first category *pseudo*-value judgments. They look value-laden because they use words like 'reason', but their function and the logic that governs their acceptance is entirely classificatory and explanatory, but not at all action-guiding. If this view is correct, then the fact/value holism I am recommending depends on a

²¹For a defense see Lipton (2004).

pun: slipping from a theoretical sense of 'reasonable' and 'good' useful for categorizing mental states and explaining behavior to a practical sense that has a constitutive role in our deliberations.²² According to the position, the judgments about reasons and goods that I make in deliberating and acting, and the judgments about reasons and goods that I cite in explanations of others are in fact different judgments belonging to different spheres. One justifies or recommends, the other explains. It follows that when I say that The Butler acts for a reason or in recognition of some good, the 'reasons' and 'goods' I invoke in this explanation do not refer to things I encounter in my own deliberations. One kind of reason justifies and recommends, the other explains.

This arrangement has the startling consequence that I cannot view other people as acting on bona fide normative reasons. And this is a form of practical solipsism: the view that there is something so significant about the fact that I am a particular individual that it makes a difference to whether the considerations on which that individual acts are normative reasons or mere explanatory props.²³ In fact, the consequences will be even worse, since we often take up the explanatory stance toward ourselves: we often decide what to do by deciding what it would make sense for us to do.²⁴ If the double-bookkeeping strategy is appropriate, then I must view my own actions as explained by 'pseudo'-value judgments, just as I would anyone else's. And this creates an intolerable alienation. When deliberating, I make practical value judgments about what it makes sense for me to do, and when these deliberations go well, I act on genuine reasons, but a moment later, if I stand back from myself and try to understand what I did, I will have to cite pseudo- versions of the same value judgments as the explanation. I will see myself not as acting for the reasons I recognized in deliberation, but for something else. The reasons I must see myself acting on (in my explanation book) and the reasons that I act upon (in my deliberation book) will be of completely different species.²⁵

Any view that entails both practical solipsism and this kind of self-alienation should be rejected. But rejecting these positions means agreeing

²²Here compare Timothy Schroeder's (2003) argument that Davidson's theory of mental content is, despite appearances, non-normative because even though it uses normative terms like 'rational' in its categorization scheme, the fact that these terms have a certain normative force makes no contribution.

²³Nagel (1970, 112ff).

²⁴Velleman (1989).

²⁵Dancy (2000, 98ff) also emphasizes the implications for the explanation of action of the axiom that a person must see herself as acting on genuine, normative reasons.

that the normative entities we cite when explaining an action - reasons, goods, etc. – are identical to the ones we engage with when we are deciding how to act. And that means that a given judgment's role in our explanations will have some effect on the extent to which it is appropriate to make that judgment in the context of deliberation – that is to how strongly we should take something to be good or bad, obligatory or forbidden when deciding whether to do it. This is what my final premise says: a judgment's appearing in our best explanation is relevant to whether – to what degree – we should make that judgment.

7. Conclusion

This completes my defense of each of the premises in the master argument for the thesis that our fact and value 'webs' are part of a single, holistic web, in the same sense that Quine suggests about the natural sciences and mathematics, and for similar reasons. As I said at the outset, I am not the first advance this thesis. What I take my argument to add to existing efforts is a 'high road' complement to more familiar 'low road' cases for such holism. The latter proffer examples of the 'entanglement' of fact and value judgments, usually in the form of ways that value judgments can be explanatorily useful. I find many of these examples convincing, but like all example-driven arguments they can be deflected: maybe it's not the value judgment itself that is explanatorily useful, but something else; maybe the relevant value judgments are useful, but they are not really indispensable; maybe a given judgment isn't really a value judgment in the appropriate sense; etc.

What I have offered is a more abstract and principled argument as to why fact and value not only are entangled in the ways suggested by these examples, but why they *must be* holistically unified. Because of how we are situated as explainers – because we are both explainers and actors – there is a presumption in favor the relevance of our own value judgments to the business of explaining action. In each step of the argument, I have presented we see how a tempting way of overcoming this presumption – denying the indispensability of value judgments, denying the holism of explanation, denying the general validity of IBE – leads to untenable consequences. As a result, we should accept this presumption and agree that fact and value judgments ultimately inhabit a single holistic web. This argument works best, perhaps, as a supplement to 'low road' arguments, for it gives us reason to think that the examples marshaled by these arguments are not solitary, exceptional, or



anomalous, but rather reflect a pervasive and ineliminable feature of our situation as actor-explainers.

Acknowledgements

For comments and helpful discussion, I am grateful to Kate Elgin, Richard Holton, Rae Langton, Sam Levey, Alejandro Perez Carballo, Alice Phillips Walden, Judy Thomson, to the participants in a workshop at the University of Nottingham, and to anonymous referees and editors. This paper has had an unusually long period of development, so I fear I may be overlooking some of the help I received along the way. For this I apologize.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

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 (1): 1– 24.

Anscombe, G. E. M. (1957) 2000. Intention. Cambridge: Harvard University Press.

Blackburn, Simon. 2011. "Through Thick and Thin." In Practical Tortoise Raising and Other Essays, 129–146. New York: Oxford University Press.

Boyd, Richard. 1988. "How to be a Moral Realist." In Essays on Moral Realism, edited by Geoffrey Sayre-McCord, 181–228. Ithaca, NY: Cornell University Press.

Cohen, G. A. 2003. "Facts and Principles." Philosophy and Public Affairs 31 (3): 211–245. Colyvan, Mark. 2001. The Indispensability of Mathematics. New York: Oxford University Press.

Dancy, Jonathan. 2000. Practical Reality. New York: Oxford University Press.

Davidson, Donald. 2001. "Mental Events." In Essays on Actions and Events, 207-224. New York: Oxford University Press.

Elgin, Catherine Z. 1997. "The Relativity of Fact and the Objectivity of Value." In Between the Absolute and the Arbitrary, edited by Catherine Z. Elgin, 176-191. Ithaca, NY: Cornell University Press.

Fodor, Jerry, and Ernest Lepore. 1992. Holism: A Shopper's Guide. Cambridge: Blackwell. Holton, Richard. 2009. Willing, Wanting, Waiting. New York: Oxford University Press.

Jaeger, Robert. 1973. "Action and Subtraction." Philosophical Review 82 (3): 320-329.

Lepore, Ernest, and Kirk Ludwig. 2007. Davidson: Meaning, Truth, Language, and Reality. New York: Oxford University Press.

Lewis, David. 1983. "Radical Interpretation." In *Philosophical Papers Volume I*, 108–118. New York: Oxford University Press.

Lewis, David. 1992. "Meaning Without Use: Reply to Hawthorne." Australasian Journal of Philosophy 61 (4): 343-377.



Lipton, Peter. 2004. Inference to the Best Explanation. 2nd ed. London: Routledge.

Longino, Helen. 1990. Science as Social Knowledge. Princeton, NJ: Princeton University Press.

Maguire, Barry. 2015. "Grounding the Autonomy of Ethics." In Oxford Studies in Metaethics, edited by Russ Shafer-Landau, 188–215. New York: Oxford University Press.

McDowell, John. 1998. "Functionalism and Anomalous Monism." In Mind, Value, and Reality, edited by John McDowell, 325–340. Cambridge: Harvard University Press.

Millgram, Elijah. 2010. "Pluralism About Action." In A Companion to the Philosophy of Action, edited by Timothy O'Connor and Constantine Sandis, 90–96. Malden, MA: Wiley-Blackwell.

Nagel, Thomas. 1970. The Possibility of Altruism. New York: Oxford University Press.

Nagel, Thomas. 1986. The View from Nowhere. New York: Oxford University Press.

Putnam, Ruth Anna. 1987. "Weaving Seamless Webs." Philosophy (London, England) 52 (240): 207–220.

Putnam, Hilary. 2002. The Collapse of the Fact/Value Dichotomy. Cambridge: Harvard University Press.

Quine, W. V. 1953. "Two Dogmas of Empiricism." In From a Logical Point of View, 20–46. Cambridge: Harvard University Press.

Quine, W. V. 1976. "Carnap and Logical Truth." In The Ways of Paradox, 100–125. Cambridge: Harvard University Press.

Quine, W. V. 1992. "Structure and Nature." Journal of Philosophy 89 (1): 9.

Quinn, W. S. 1986. "Truth and Explanation in Ethics." Ethics 96 (3): 524-544.

Railton, Peter. 2003. "Moral Realism." In Facts, Values, and Norms, 3-42. New York: Cambridge University Press.

Restall, Greg, and Gillian Russell. 2010. "Barriers to Implication." In Hume on Is and Ought: New Essays, edited by Charles Pigden, 243–259. Basingstoke: Palgrave.

Schroeder, Timothy. 2003. "Donald Davidson's Theory of Mind is Non-Normative." Philosophers' Imprint 3 (1): 1–14.

Smart, J. J. C. 1999. "Ruth Anna Putnam and the Fact-Value Distinction." Philosophy (London, England) 74 (289): 431-437.

Solomon, Miriam. 2012. "The Web of Valief: An Assessment of Feminist Radical Empiricism." In Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy, edited by Sharon L. Crasnow and Anita M. Superson, 435-450. New York: Oxford University Press.

Sturgeon, Nicholas. 1988. "Moral Explanations." In Essays on Moral Realism, edited by Geoffrey Sayre-McCord, 229–255. Ithaca, NY: Cornell University Press.

Velleman, J. David. 1989. Practical Reflection. Princeton, NJ: Princeton University Press. White, Morton. 1981. What Is and What Ought to Be Done. New York: Oxford University

White, Morton. 2002. A Philosophy of Culture. Princeton, NJ: Princeton University Press. Williams, J. Robert G. 2018. "Normative Reference Magnets." Philosophical Review 127 (1): 41-70.