



Justifying method choice: a heuristic-instrumentalist account of scientific methodology

Till Grüne-Yanoff¹ 

Received: 30 November 2018 / Accepted: 17 November 2020 / Published online: 11 December 2020
© The Author(s) 2020

Abstract

Scientific methods are heuristic in nature. Heuristics are simplifying, incomplete, underdetermined and fallible problem-solving rules that can nevertheless serve certain goals in certain contexts better than truth-preserving algorithms. Because of their goal- and context-dependence, a framework is needed for systematic choosing between them. This is the domain of scientific methodology. Such a methodology, I argue, relies on a form of instrumental rationality. Three challenges to such an instrumentalist account are addressed. First, some authors have argued that the rational choice of at least some methods, namely those supporting belief formation, is not goal-dependent. Second, some authors have observed that some method choices seem intuitively rational, even though relevant goals are lacking. Thirdly, some authors have argued that instrumental rationality itself depends on a goal-independent form of rationality. It is the heuristic nature of scientific methods that affords me the arguments against these challenges. This heuristic-instrumentalist account provides the means-ends analysis needed to evaluate heuristic method choice. The paper thus offers the conceptual basis for a scientific methodology that is both compatible with the heuristic nature of actual scientific practice and also normatively relevant for assessing method choice.

Keywords Methodology · Heuristics · Instrumental rationality · Epistemology

1 Introduction

Scientific methods are those techniques, approaches and strategies that scientists employ to perform their research. These methods can serve different epistemic goals, for example different versions of explanation, prediction or design. An instrumentalist methodology describes, compares and evaluates these methods with respect to their

✉ Till Grüne-Yanoff
gryne@kth.se

¹ Division of Philosophy, Royal Institute of Technology (KTH), Teknikringen 76, 10044 Stockholm, Sweden

ability to further relevant epistemic goals. Through that it supports rational method choice. Specifically, such an instrumentalist position makes two claims. First, it claims that the sources of normativity for prescribing the choice of certain methods are located in the instrumentality of these methods for certain epistemic goals. Second, it claims that a systematic prescription of method choice can be derived from these instrumentalist considerations. An instrumentalist account of scientific methodology is desirable because it relies on a normatively unproblematic notion of rationality that nevertheless offers a powerful tool for evaluating and performing method choice.

Yet this instrumentalist perspective is contested in at least three ways. First, certain authors claim that the rational choice of at least some methods, namely those supporting belief formation, is not goal-dependent. Second, others have observed that some method choices seem intuitively rational, even though relevant goals are lacking. Thirdly, some have argued that instrumental rationality itself depends on a goal-independent form of rationality. Each of the challenges point to some non-instrumental, not goal-dependent form of rationality that underlies the normativity of methodology and drives systematic prescription of method choice. The challenges, if successful, would thus undermine the instrumental account.

In this paper, I defend the instrumentalist methodology account against these criticisms. Following Wimsatt, I argue that scientific methods are heuristics: they are incomplete in that they might not identify all the best solutions; they are underdetermined in that they require further judgments and decisions to be applicable; they are fallible in that they produce a correct answer with less than certainty; and they are biased in that they systematically produce error in certain contexts. Discussing recent work by Hey (2016), I then argue that an instrumental perspective offers a systematic account how to rationally choose between these scientific heuristics. Based on this account, I argue against the three challenges. First, I show that choice between all heuristics, including inferential ones, require a goal-dependent justification. Second, I demonstrate how the instrumental account can explain goal-less choices of methods without normatively vindicating them. Third, I argue that evidential justification of instrumental claims itself is purpose-dependent and instrumentally justified. Each of these arguments takes recourse to the notion of heuristics. My argument shows how the heuristic nature of methods strengthens the instrumentalist case, which I therefore call the heuristic-instrumentalist perspective on scientific methodology.

The paper begins by presenting the heuristic nature of scientific methods in Sect. 2. Section 3 discusses the need for a systematic framework of choosing between these heuristics and argues that such a framework is governed by instrumental rationality. Sections 4–6 address the three challenges to the instrumentalist account. Section 7 concludes.

2 Scientific methods as heuristic procedures

Every scientific discipline has its stock of procedures for performing research. These methods are employed for the recognition and formulation of a problem, the production and collection of evidence, and the formulation and epistemic appraisal of hypotheses. Some of these methods are employed across many disciplines. For example, all sci-

ences use similar strategies to operationalize not directly observable properties in order to make them measurable. All sciences that handle quantitative data employ statistical techniques, and many of them use a simple p value testing procedure. Other methods are particular to specific disciplines. Astronomers, for example, do not experiment. Chromatography methods (for separating a mixture) are ubiquitous in chemistry, but rarely found elsewhere. And only those scientists who investigate human attitudes—like social scientists, health researchers and engineers—ever collect data with the help of questionnaires. Often, a method developed in one discipline then spreads to others. Random assignment to treatment and control groups in an experiment first was developed in 19th century parapsychological research (Hacking 1988), and from there entered other disciplines like biology, medicine and sociology. Inversely, obsolete and outdated methods—like certain software used for genomic analysis (Wadi et al. 2016)—often continue to be used even though more powerful alternative are available. And some methods that were once widely used, like armchair speculation, today enjoy only a marginal existence in specific sub-disciplines.

Scientific methods take a prominent place in science. Descriptions of methods are the most highly cited papers in all of science.¹ Most disciplines have a canon of methods, typically connected to its main discoveries. This canon is taught to students, and because the competent use of many methods often involves tacit knowledge, this teaching often takes the form of labs and tutorials. Yet this canon is not rigid: development of new methods is a highly rewarded part of science, and older methods are regularly phased out as obsolete.

Despite this disciplinary regimentation of the method canon, the ultimate arbiters are the scientists doing the research: they *choose* which method to employ when pursuing a particular research project. In their presentations and publications they are also expected to defend this choice. Authors' choice of method—its consistency with the canon, its appropriateness for the research goal, its innovativeness and the competence in handling it—is for example one of the main considerations in the peer review process (Thrower 2012; Zwaaf 2013). Quite obviously, scientists are not just choosing their methods, but they are also held responsible for the legitimacy of this choice: they are expected to justify it.

That scientists have a choice of methods, and that their research is evaluated by this choice of methods, implies that there are alternative methods to choose from, which at least *prima facie* offer themselves as equally appropriate for doing the research. The examples are legions. The choice might be formulated between alternative kinds—for example, ‘Experiment or observational study?’, ‘Field or lab experiment?’, ‘Experiment or simulation?’; or it might concern alternative specifications of the same kind—for example, ‘What factors to eliminate in an experiment?’; ‘Which ones to hold constant?’; ‘Which idealizations to accept in your model?’; ‘Which parameterization of p -value test?’. Finally, it might concern the concrete implementation

¹ The vast majority of the Thomson Reuter's Web of Science top 100 papers, published until 2014, describe experimental methods, software packages, statistical technics, and psychological inventories (van Noorden et al. 2014). This is explained by the fact that scientists reference the methods they use in their study reports. The high citation rates show that many methods are highly canonical: standardized in publicly accessible documents, shared by many scientists, and used over and over again.

of a specific method—e.g. which software package or which brand of measurement instrument to use.

Why do substantially different methods *prima facie* offer themselves as equally appropriate for a given goal? If methods produced results with certainty, such *prima facie* equivalence would rarely arise. Instead, method outcomes typically are uncertain (and this uncertainty is typically not fully specified, e.g. through a probability distribution) so that two or more methods will often seem to advance a given goal equally well. Because of this uncertainty and vagueness, some authors have argued that scientific methods are *heuristic* procedures.² In particular, William Wimsatt (2007, pp. 76–77) proposes four characteristics for heuristics:

1. They don't guarantee producing the correct solution
2. Their use is motivated by their cost-effectiveness in terms of demands on memory, computation or other limited resources
3. They are systematically biased, thus allowing at least in principle the identification of the conditions under which they fail and the direction of the error they produce
4. They transform the original problem into a "non-equivalent but related problem".

Examples of scientific methods that Wimsatt explicitly identifies as heuristics include holding environmental variables constant in experiments and simplifying representations of environmental factors in modeling (Wimsatt 2007, p. 83). Holding (known) background variables constant, for example, (1) does not guarantee a valid inference about effect size from an experiment, due to possible unknown confounders. Yet (2) such a method is widely seen as reasonable experimental method (in contrast to e.g. a potentially limitless search attempting to identify all confounders before making the inference). In principle, (3) the systematic bias of such unknown confounders could be identified. "Hold known environmental factors constant" thus (4) transforms a deductively valid inference scheme (as e.g. laid out in Mill's *Method of Difference*) into a fallible but practicable inference procedure whose quality depends on the available knowledge of relevant background factors.

Wimsatt contrasts heuristics with what he calls "truth-preserving algorithms" (Wimsatt 2007, p. 76). Such algorithms, he argues, when correctly applied to true premises, must produce a correct solution (ibid. 346). But they require a problem space whose structure is well-defined (ibid. 9)—a property many actual problems do not exhibit. More specifically, such algorithms are characterized by least three properties. First, for each such algorithm, a class of problems is specified for which it is effective. This specification must be exact enough to distinguish this target class from other classes of problems for which the method is not effective. Second, these algorithms consist of a finite number of exact, finite instructions. In particular, the exactness condition requires that each instruction is non-ambiguous, can be followed rigorously, and demands neither additional judgments nor decisions. Third, the application of the algorithm to its target class always produces one and only one solution

² The notion of heuristic has been widely used in cognitive science and decision theory contexts, starting with Simon in the 1950s, and gaining wider recognition with the work of Tversky, Kahneman and Gigerenzer, amongst others (for overviews, see Gigerenzer et al. 2011; Chow 2014). But their research focusses largely on heuristics outside of scientific contexts; in this paper I will instead focus on the more specific discussion of their use in the sciences.

after a finite number of steps (cf. Cleland 2002 for an in-depth analysis of algorithms, which he calls “effective procedures”).

If scientific methods were algorithms in this sense, choosing between them would be comparatively easy. As long as a given goal were part of its problem class, there would be at least one algorithm that guaranteed achieving this goal in a finite number of steps. If the goal were part of more than one method’s problem class, one only needed to compare their respective economies to choose between them. Such an account is not, I believe, an accurate description of scientific method choice. Instead, I consider uncertainty and vagueness central features of scientific methods, and therefore follow Wimsatt in characterizing them as heuristics.

If methods are heuristic procedures in Wimsatt’s sense, then the method choice problem that scientists face becomes more complicated. According to Wimsatt’s characteristic (iii), heuristics at least under some conditions systematically fail. And more generally, according to condition (i), they do not guarantee a correct solution. Such characteristics make it difficult for scientists to justify their method choice. They now need to argue that for their specific goal and problem, the heuristic biases are not relevant and the uncertainty of the heuristic result does not pose a problem. In the next section, I discuss possible ways how one could systematically justify choices between thus-characterized heuristic methods in science.

3 Meta-heuristics as instrumentalist methodology

Scientists not only have the privilege of choosing between different methods, they also *must* make a choice between the available methods in order to pursue their research—even when no optimal method is available, or not enough information is at hand to determine what the optimum could be. This justification of method choice is the domain of methodology. Methodology is distinct from methods: it concerns the justification of method choice, and thus the evaluation of methods, but it does not concern the development, description or application of methods themselves.

Albeit it is uncontroversial that methodology concerns the justification of method choice, it *is* controversial how such justification should proceed. In this section, I discuss the recent proposal of a systematic methodology for scientific heuristics in form of a set of meta-heuristics (Hey 2016). While I am congenial to the basics of Hey’s account, it leaves open a number of questions. In this section, I discuss these questions and show how they can be answered by re-casting Hey’s account as an instance of an instrumental methodology.

Hey follows Wimsatt in stressing the importance of heuristics in science. In particular, he emphasizes the uncertainty and context-dependence of the success of heuristics in scientific research. For these reasons, he concludes, a systematic framework is needed for deciding “when and how these heuristics can be applied prudently” (Hey 2016, p. 475). Wimsatt, according to Hey, has not provided such a systematic framework, but only offered non-generalizable recommendations when *not to use* specific heuristics. Hey now argues that for every heuristic, there exists a meta-heuristic that regulates its use.

If heuristics are the rough-and-ready rules describing scientific activities— such as, “describe p as type Q ”, “simplify X before Y ”, “keep variables $e_1 \dots e_n$ constant”—then meta-heuristics are the rough-and-ready strategies that specify the conditions under which these activities ought to be used (or not used). (Hey 2016, p. 478)

Hey’s meta-heuristics are methodological rules—rules regulating how to choose between methods. Examples of such methodological rules include: “avoid ad hoc modifications to theories”; “prefer double-blind over single- or zero-blind experiments”; “reject unfalsifiable theories”; “postulate the same kind of cause for the same kind of effect, as far as possible”; and so on.

But what regulates these meta-heuristics in turn? Hey suggests that they are grounded in what he calls a “problem type”, which “describes the context for the heuristics and meta-heuristics, and in so doing, grounds the problem space in the particular domain of scientific interest” (Hey 2016, p. 480). For example, a physicist engaging in a modeling project should start by asking “What do we want to explain?” One answer might be “A universal phenomenon”; another “A particular phenomenon instance”. These constitute two separate “regions of problem space” (Hey 2016, p. 486), that each have their necessary meta-heuristics. Explanation of a universal phenomenon requires a variable reduction strategy that removes irrelevant details of token events; explanation of a particular phenomenon requires introduction of additional variables “until the match between model and reality is sufficient to our needs” (Hey 2016, p. 487). Each of these meta-heuristics then allow a number of possible heuristics to be applied to these respective problem-types. Hey’s general framework thus consists in a three-level hierarchy (problem-type; meta-heuristic; heuristic) that “provides a working methodological map to ensure (as far as possible) that errors from the inappropriate use of heuristics can be avoided” (Hey 2016, p. 480).

While Hey’s account correctly stresses the importance of a general framework for the prudent choice of heuristics in science, at least three features of this framework remain underdeveloped. First, what is a problem type, and what are its characteristics? Second, must the meta-heuristics be necessitated by the problem-type? And thirdly, must one in all cases apply a meta-heuristic in order to justify the choice of a method? As Hey leaves these questions largely unanswered, I now argue that an instrumentalist account of methodology can answer all three in a satisfactory way.

An instrumentalist methodology relies on instrumental rationality as the only source of normativity in justifying method choice.³ It prescribes what properties a scientific method ought to have, so that it satisfies best a certain given goal in certain kinds of context. The reasons for the goal (its normative force) transmit to the prescribed means, *if* the efficacy of the means in producing the goal is sufficiently established, and if the choice criterion (what is meant by “satisfies best”) is determined. While the latter depends on a proper analysis of the goal, the efficacy issue requires an empirical inquiry. The methodological issue thus reduces to conceptual analysis of the goal, and empirical support for the efficacy of the means-end relationship. The advantage of such an instrumental methodology is that it does have normative content, because method choices are justified through their instrumental value for scientists’ epistemic

³ It thus is entirely separate from the Instrumentalism-Realism debate about theories.

goals: certain elements from the set X of all available methods will achieve these goals best. Methodology identifies these means-ends relations and prescribes them as hypothetical imperatives of the form: “If one’s goal is y , then one ought to do $x \in X$ ”.

With this sketch of instrumentalism in mind, let me now address the three questions from above. First, what are the core characteristics of a problem type? Hey did not give a general characterization, but only three illustrating examples: Patient assignment in a hospital, explaining universal patterns or singular instances, and modeling ecosystems. These cases remain vague; for example, Hey leaves it open whether epistemic goals—like explaining or predicting—must be part of the problem-type description. Here is an argument for why it must. In one of his examples, Hey argues that a particular degree of model-reality match (“sufficient for our needs”, Hey 2016, p. 487) must be reached for this problem type. But such a degree is dependent on the particular explanation one aims at, with a particular contrast and a particular precision specified. If the modeling were to serve another epistemic goal, e.g. prediction, then the required model-reality match would likely be lower, or at least different. More generally, heuristics are always only useful for something—they are purpose-dependent (Wimsatt 2007, p. 3461). Consequently, the problem type must include a description of the epistemic goal if a systematic meta-heuristic is to be grounded in it.

But that is not the only core characteristic. Goals can be reached in many different ways, yet some means applied in some contexts do not reach a goal. Thus, the problem type also must specify the relevant context in which these heuristics methods are applied. Hey implicitly assumes this, when arguing that the meta-heuristics specify the conditions under which certain heuristics are to be used (or not used). But the meta-heuristic is just a rule prescribing use of the heuristic given certain conditions. How does one determine whether these conditions are indeed satisfied? For that one needs to consult the problem type, which needs to contain the relevant information.

Thus, a problem type must describe both a researcher’s goal, as well as the relevant context in which the planned research is supposed to reach this goal. And a meta-heuristic prescribes which heuristics to use, conditional on the information provided by the problem type. An instrumentalist account therefore answers what the core characteristics of the problem type are: that information (about the goals and the context) that allows for a rational choice of the means for satisfying these goals in that context.

Now to the second question: must the meta-heuristics be necessitated by the problem-type? Hey is again ambiguous about this: in the physics example, he argues that the meta-heuristics are indeed necessitated by the problem types, while in the other two cases, multiple meta-heuristics seems to be compatible with the same problem type. The instrumentalist account can solve this apparent contradiction. Heuristics are neither necessary nor sufficient for a goal: they are not sufficient, because many other conditions must be in place besides the application of the heuristic; and they are not necessary, because many alternative heuristics, under the right conditions, often can be means to reaching a goal. Instead, heuristics are necessary parts of a sufficient means of realizing a goal.⁴ Consider two different heuristics F and H . Both are

⁴ This is notion is derived from Mackie’s (1974) INUS condition (insufficient but non-redundant parts of a condition which is itself unnecessary but sufficient for the occurrence of the effect), which he employed to analyze the notion of “cause”, but which can be easily adapted for my purposes here.

candidates as means for reaching goal G —hence neither is necessary for reaching G . Furthermore, neither is sufficient to reach G —they require specific conditions in place to be effective, C_F and C_H , respectively. But $F \& C_F$ are sufficient for G , as are $H \& C_H$. And finally, because $C_F \neq C_H$, and F is the heuristic specifically adapted to conditions C_F for reaching G , F is a necessary part of the sufficient means $F \& C_F$ for reaching G . This holds for both heuristics and meta-heuristics. A heuristic like “hold known environmental variables in an experiment constant” fails to support valid inferences about effect size in contexts where relevant confounders are not known. A meta-heuristic like “avoid ad hoc modifications” (i.e. changes to a falsified theory that do not increase its falsifiability) fails to secure theoretical progress in contexts in which no theory change increases falsifiability. Consequently, problem types never imply a specific meta-heuristic with necessity—there will always be some contexts in which a meta-heuristic will fail to advance a problem type.

This leaves the final question: is a meta-heuristic required to justify method choice? Hey discusses this question as part of his answer to a potential regress counterargument:

... if heuristics need meta-heuristics, then perhaps meta-heuristics also need meta–meta-heuristics and so on—giving rise to an infinite regress of methodological justification. But this worry is resolved by the top-most level of my hierarchical model being a “problem-type” rather than a heuristic. The problem-type describes the context for the heuristics and meta-heuristics, and in so doing, it grounds the problem space in the particular domain of scientific interest—effectively closing the loop of heuristic justification. (Hey 2016, p. 480)

I have two problems with this line of reasoning. First, I doubt that “heuristics *need* meta-heuristics”, if this means that heuristic choice can only be justified through meta-heuristics. Second, I fail to see how the problem-type “grounding” prevents the regress.

Hey claims that his problem space concept cuts off the regress by taking the problem description as its entry point for analysis. Because the problem is a given, the methodological analysis starts from there. Yet in my understanding of Hey’s account, the problem type only identifies the conditions under which heuristics are applied, but does not provide evidence that the application indeed is justified. That’s what the meta-heuristics are supposed to do, at least in Hey’s account. If, as Hey seems to think, heuristic choice requires meta-heuristics to be justified, then it seems arbitrary to stop at this level, instead of also requiring meta–meta-heuristics, etc.

An instrumentalist account of methodology can dispel these two connected issues. First, there are different ways of justifying method choice. Token justifications provide instrumental reasons for choosing method M over others, based on localized knowledge—e.g. immediate experience of working with these method alternatives in this context. Rules play no role in this justification. Rule justification, in contrast, refer to generalizing principles about which methods best satisfy which rules in a variety of contexts. Hey’s proposal of meta-heuristics is one variant of such rule-justification. Either of these kinds might provide good justification. Sometimes, e.g. when a researcher choosing between methods has little local knowledge, a rule-based justification might be preferable. But in no way is referring to rules the only or neces-

sary mode of justification of method choice—meta-heuristics are not “needed” in this way.

Second, both justifications are instrumental, depending on goals and contexts. The normativity of instrumental reasoning is widely seen as unproblematic. No a priori arguments, pre-analytic intuitions or questionable conventions need be appealed to. Instead, the instrumental rationality of the hypothetical imperative is well known from practical reasoning, and widely accepted there. Engineers and economists employ it in their reasoning. Hence this form of normativity should be acceptable for scientists, who indeed might well recognize it as the kind of normativity they themselves appeal to in their own methodological discussions. The fallible part of any such justification lies in the empirical claim that a certain method furthers most the goal in the given context. Neither the token nor the rule approach is generally better suited to provide this justification. Rather, which one is more suitable depends on the form of the available information (e.g. either as local knowledge or as generalizing rules). Meta-heuristics are thus not the *only* way of justifying method choice. One can always ask “what justifies the justifier?”—that is just the problem of any inductive inference, and not a special problem of either heuristic or meta-heuristic choice. Irrespective of such regress concerns, researchers are often pragmatically justified to believe in (fallible) means-ends claims, and this in turn justifies their method choice. The hierarchy of Hey’s account thus gives unnecessary weight to such regress concerns. If instead one sees that meta-heuristics are not required for justification of method choice, but sometimes pragmatically preferable, the problem disappears.

To conclude, I have offered a re-interpretation of Hey’s meta-heuristic account. In particular, firstly, I have argued that Hey’s talk of “problem space” should be made more precise by making explicit the ends and success conditions for a method choice. Secondly, I have shown that problem types does not necessitate specific meta-heuristics. And thirdly, I argued that meta-heuristics are just one way out of many to justify a methodological choice, thus reducing the urgency of regress worries. Thus re-interpreted, an instrumental account offers a useful way to perform the means-ends analysis needed to evaluate scientists’ heuristic method choices.

4 First challenge: goal independence of rational belief formation

The heuristic-instrumentalist perspective on scientific methodology presented so far assumes that the success of different methods depends both on purpose and context, and that therefore it is (instrumentally) rational to choose that method that one has reason to believe will be most effective for one’s purpose given the context of application. But this goal-dependence of what is methodological rational has been challenged with at least three arguments. First, some authors have argued that the rational choice of at least some methods, namely those supporting belief formation, is not goal-dependent. Second, some authors have observed that some method choices seem intuitively rational, even though relevant goals are lacking. Thirdly, some authors have argued that instrumental rationality itself depends on a goal-independent form of rationality. For each of these reasons, the instrumentalist perspective is said to fail. But if it failed, the heuristic account would need to be rewritten, too. Methods could then not be

considered the most rational choice for certain purposes in certain contexts. Rather, there would be a universally rational method, and the only problem would consist in obtaining the relevant information for its identification.

In the following three sections, I argue against each of these claims in turn. I begin with the claim that the rational choice of at least some methods, namely those supporting belief formation, is not goal-dependent (Siegel 1996; Kelly 2003, 2007).

Rational reasons for belief, these authors claim, are constituted by evidence, independent of what goals the agent may have. They support this claim by pointing to the apparent categorical acceptance of evidence. If queried “why do you believe p ?”, people typically answer: “because of evidence e ”—rather than “because of evidence e , given my epistemic goal g ”. People seem to treat e as categorical reasons for believing or not believing p —categorical in the sense that they are independent of what epistemic goals the epistemic agent might have.

If both of us know that all of the many previously-observed emeralds have been green, then both of us have a strong reason to believe that the next emerald to be observed will be green, regardless of any differences which might exist in our respective goals. Similarly, in arguing for my conclusions in this paper, I think of myself as attempting to provide strong reasons for believing my conclusions, and not as attempting to provide strong reasons for believing my conclusions for those who happen to possess goals of the right sort. (Kelly 2003, p. 621)

Kelly concludes from this that evidential support itself is normative: it determines what a (rational) person ought to believe: “there is no gap between possessing evidence that some proposition is true and possessing reasons to think that that proposition is true” (Kelly 2007, p. 468). Where there is no gap, no choice of method needs to or even can fill in. Instead, the inference from evidence to beliefs is supported by a categorical and universal epistemic rationality—“the kind of rationality which one displays when one believes propositions that are strongly supported by one’s evidence, and refrains from believing propositions that are improbable given one’s evidence” (Kelly 2003, p. 612). It is rational to believe proposition p if and only if one has strong evidence for p . Differing epistemic goals play no role for this rationality.

This view of rational belief formation is at odds with scientific practice. I will illustrate this at the hand of statistical inference practices in science.⁵ In practically all scientific disciplines, multiple inferential methods are employed to form beliefs (or decide acceptance/rejection of a hypothesis) based on available evidence. Which of these methods to employ for which purposes under which conditions is often controversial, as recent debates about significance testing, replication or Bayesian inferences illustrate (see e.g. Open Science Collaboration 2015; Wasserstein and Lazar 2016). Furthermore, each of these types can cater to a continuum of different epistemic goals. Fisherian significance testing rejects the hypothesis at different error-rate thresholds; Neyman–Pearson hypothesis testing favors different hypotheses depending on which type of error is considered more momentous; Bayesian updating assigns probabilities depending on the prior probability, which includes various features of epistemic

⁵ I am focusing on inferential methods only, because Kelly (2003, p. 635) separates reasons for evidence acquisition and for belief formation, admitting that the former (but not the latter) might be instrumental reasons. Thus, I only disagree with Kelly on the nature of reasons for belief formation.

goals. Inferential statistics thus exemplifies the goal-dependence of evidence: what is considered evidence for a hypothesis under a goal with a certain error rate, might not be considered evidence for this hypothesis under a different error rate, or under some other prior. I believe this indicates an instrumental dependence of these methods (and therefore also of the reasons for beliefs) on scientists' goals, thus posing a problem for categorical accounts.

In order to support this claim, I draw on the contrast between heuristics and algorithms, as described in Sect. 2. Algorithms are characterized by three features: specified problem class, finite number of non-ambiguous instructions, and assignment of one and only one solution for each problem. By showing the statistical methods most widely used in science do not satisfy these characteristics, I show that they are not algorithms but heuristics.

First, statistical inference methods do not specify a class of problems for which they are effective. There is no explicit specification: statistical inference methods do not come with a description of which problems they may or may not be applied to. Could this instead be implicitly specified? For example, the different methods take different informational inputs: significance tests for example require the likelihood of the data given the hypothesis, while Bayesian updating requires also likelihood of the data given the falsity of the hypothesis as well as the prior probability of the hypothesis. Perhaps these different input formats identify the problem classes to which these methods are applicable? Two factors speak against input formats as such an implicit differentiation. First, this is often not a mutually exclusive differentiation, but a nesting relationship: Bayesian updating requires all information significance testing requires, and then some more. Second, not all statistical inference methods are thus differentiated: Fisherian significance testing and Neyman–Pearson hypothesis testing, for example, both require the same likelihood information. So, there is no indication that the inference methods carry implicit specification of their application classes. This corresponds to the observation that in scientific practice, these methods are seen as competitors for solving the same problems (sometimes yielding different results), not as complementary methods for distinct kinds of problems (Lehmann 1993; Lenhard 2006).

Second, statistical inference methods consist of a finite number of steps, but many of these steps are not precise. In particular, their execution often requires making decisions regarding the values of some parameter, which often requires further epistemic judgments. Take for example significance testing. This seemingly simple method tests a specific hypothesis with a particular set of data. Yet in order to arrive at the conclusion (“reject” or “don’t reject”) a lot of additional decisions and judgments must be made. When testing a hypothesis H , one needs to choose which statement to reject: either H or *not* H . Often, it is not clear from the formulation of the hypothesis, which statement to choose (in particular, the “no-effect” alternative implied by the (mis-)nomer “null hypothesis” often is not applicable). Furthermore, the test procedure requires determining the data generation process. This can be done in multiply different ways—for example, through experiment or observational studies, at different levels of background factor control, with different measurement instruments, and at different sample rates. How this is done affects what possible outcomes can be expected, and what probability distribution one should assume over these possibilities. The description of possible

outcomes, additionally to this question, also depends on how these possibilities are partitioned. Again, this is something that cannot be determined from the data or the hypothesis alone but requires further decisions. Moreover, the distribution over the test statistic, although influenced by the design of the data generating process, is not fully determined by it. Instead, it needs to be decided what assumptions to accept for the purpose of the test. Last, and perhaps most obviously, the threshold (the “significance level”) to which the calculated p-value is compared and that determines whether the hypothesis is rejected or not, is set by fiat, thus requiring yet another decision. Thus, even a seemingly simple method like significance testing consists of many steps for which no precise instructions are given; instead, they require further decisions and judgments.

Third, statistical inference methods do not produce one and only one solution for each problem. As I argued, it is often ambiguous to which problem classes a particular method is correctly applied. Moreover, there are many ambiguities in the individual steps that constitute the method. A method therefore might produce many different results, depending on how these ambiguities are decided. In scientific practice, this is indeed an often-observed result. For example, there is widespread evidence for “p-hacking”—i.e. scientists intentionally using the method’s ambiguities in order to make a non-significant result seem significant (Head et al. 2015; Open Science Collaboration 2015). Another example is the dependence of results from Bayesian updating on the subjectively set prior probability of the hypothesis. Agents with sufficiently different priors cannot possibly sample enough data to converge on a common probability in their lifetimes (Hesse 1975; Earman 1992, pp. 147–149). Instead, the choice of the prior poses a substantial influence on the outcome of updating.

Thus, I conclude that statistical inference methods are *not* algorithms. Instead, they are heuristics, characterized as incomplete, underdetermined and fallible (Wimsatt 2007). They are incomplete in that they might not identify all the best solutions, or not even the most optimal ones. They are underdetermined in that they require not only some input, but further judgments and decisions. They are fallible in that they might produce an incorrect answer. Scientists thus cannot rely on a universal and categorical epistemic rationality that leads them from evidence to beliefs. Rather, they must choose from a multitude of inferential heuristics: each with specific advantages and disadvantages, to be chosen in order to suit their respective epistemic goals best.

5 Second challenge: missing goals and instrumental errors

Non-instrumentalists have another challenge for the heuristic-instrumentalist perspective defended here. If scientific methods were rational to the extent to which they further the agent’s epistemic goals, then how come that people often seem to accept evidence as reasons for beliefs, even though they are lacking the relevant goals for such an inference? Kelly (2003) distinguishes two cases. In the first, people consider certain evidence as reasons for believing p , even though believing p is of complete epistemic indifference to them (e.g. “I don’t care whether Russell was left-handed. But if you show me a photo of him writing with his left hand, that constitutes a reason for me to believe it”). In the second, people intentionally avoid the acquisition of evidence

because they do *not* want to acquire certain beliefs (e.g. “making a conscious effort to avoid learning the ending of a movie that one plans to watch in the future”).

Parallels to Kelly’s cases can easily be found in scientific practice. Even for experienced model users, it is often tempting to draw inferences from the model that go beyond the inferences required for the modeling purpose at hand. For example, for predictive purposes it often suffices that a model delivers valid output values from accurate input values. Yet the model of course employs some computational or causal process to produce outputs from inputs; and model users—although solely aiming to predict with the model—often are nevertheless tempted to infer from the model process to the underlying causal structure of the target. Independent of whether such an inference is justified (as it might well be), it goes beyond what is needed for the purpose of prediction. So, they are considering the model as evidence for certain beliefs, even though they lack the epistemic goals for forming these beliefs.

Furthermore, scientists make systematic efforts to intentionally avoid the acquisition of evidence, for example when they implement single blinding in experiments. In a single-blind study, the experimenters does not observe who is receiving a particular treatment. Thus, the experimenter avoids acquiring evidence because she does *not* want to acquire certain beliefs.

Both of these cases challenge the instrumentalist account. In the first, the absence of an epistemic goal prevents the agent having an instrumental reason; yet the agent appears to have a normatively valid reason, nevertheless. So, the normative force must arise from something else than the goal, a conclusion that strongly contradicts instrumentalism. In the second case, avoiding the acquisition of evidence would be an unnecessarily cumbersome way of avoiding beliefs if instrumentalism were true: under instrumentalism people could simply view the evidence and determine that, given their goals, they have no reason to form such a belief. The seeming non-availability of this option, Kelly argues, makes the instrumental account implausible.

But what would the alternative be? I have already argued that Kelly’s proposal of a universal and categorical epistemic rationality does not square with scientific inferential practices. Another option is one that Kelly himself considers, namely that there are universal epistemic goals. If there were such universal epistemic goals, instrumentalism could explain both of Kelly’s cases. In case 1, the universal goal would provide an instrumental reason for forming the belief, after all; and in case 2, the universal goal would stand in conflict with the goal not to know, thus requiring strategies to deal with goal conflicts.

Unfortunately, this alternative does not seem promising as an account of scientific methodology. I agree with Kelly that the relevant epistemic goals are not universal, in particular not in science. Scientists pursue many different epistemic goals. For example, they might aim at predicting certain variables, while discounting accuracy of those variables unnecessary for such prediction. Or they might focus on explanation, under specific theories of what constitute explanation—for example knowledge of difference-making causes. Even when the question concerns specific beliefs, scientists distinguish between different error rates: some accept a hypothesis at a 95% confidence level, for example, while others will accept hypothesis only at a 99% confidence level. Consequently, the attempt to salvage an instrumentalist account of Kelly’s cases by presupposing universal epistemic goals is hopeless.

Instead, I answer Kelly's challenge by questioning the *normative* relevance of his two cases. I do not deny that the phenomena described in them requires explanation: people often treat certain evidence as if it gave them categorical reasons for certain beliefs, irrespective of their epistemic goals. But I deny that such behavior is normatively justified. What seems to be a reason for the person treating it that way in fact isn't—rather, it is just a psychologically explainable motivation that lacks normative substance. Thus, I am providing a non-vindicating explanation of the two cases.⁶ Such a psychological explanation then amounts to an error theory of the categorical character of epistemic reasons.

The starting point of such an error theory is again the heuristic character of scientific methods. Such heuristics are *rules* for evidence-gathering and evidence-analysis. Epistemic rationality in science is a matter of forming beliefs in accordance with a system of good enough rules whose consistent application would bring about the sufficient satisfaction of one's cognitive goals over time.⁷ As I argued in Sect. 2, heuristic method often have vague application targets, inexact execution instructions and ambiguous solutions. It is therefore plausible that such methods occasionally are *misapplied*. One type of misapplication occurs when an agent applies a heuristic (good enough for a certain epistemic goal) to a problem even though this agent actually does not have this goal. Such misapplications explain in a non-vindicating fashion why it might seem that the agent has a categorical reason to form certain beliefs.

Take the modeling case again. A model used for explanatory purposes must accurately represent at least some difference-making cause of the explanandum. Modelers regularly focusing on explanations therefore might adopt a methodological rule like this: "models must contain highly accurate representations of the model target's causal structure". Now imagine a scientist not interested in explanation but in prediction of some phenomenon. Because of the prevalence of the explanatory goal in her discipline, the above accuracy-of-assumptions rule will be ubiquitous, and she might well feel her reasons for forming predictive beliefs should only be based on models that satisfy this rule. Yet this is incorrect: the maximal accuracy of model assumption is not a necessary condition for the predictive success of a model, and in some instances even degrades predictive success (for an example, see Küppers and Lenhard 2005). Mindlessly applying a prevalent methodological rule misleads this scientist into taking this rule as reasons not to believe certain predictive statements.

Such misapplications of methods in science are common. Scientists often have acquired expertise in the use of a technically demanding method and apply this method to any problem they encounter. Sometimes scientists are confused about their epistemic goal. Sometimes they are blinkered by dominant methodological conventions. In each of these situations, scientists employ a method whose consistent application has brought about the sufficient satisfaction of certain epistemic goals. But it isn't *their* goal. Therefore, what seems to provide them with reasons to form certain beliefs (or with reasons not to form certain beliefs) actually does not do so. The apparent epistemic rationality of the thus formed beliefs is a mistake, based on the accordance with

⁶ For another non-vindicating account, albeit in a non-scientific context, see Sharadin (2018).

⁷ Leite (2007, p. 460) offers a rule-based account of epistemic rationality in response to Kelly, though without involving heuristics.

rules that do not correspond to the actual epistemic goals the agent has. This appearance explains these mistakes, but it does not vindicate them: they lack the normative force of actual reasons.

This non-vindicating explanation handles scientific versions of Kelly's first case well: it makes it understandable—but does not justify—why scientists, based on the use of certain methods, appear to have reasons to form or not form certain beliefs, even though they are indifferent about the epistemically relevant goals. It also explains scientific versions of the second case, like single blinding. Here scientists themselves have become aware that the application of inferential heuristics can be detrimental to their overall epistemic goals. They have insight into the erroneous, non-normative nature of these heuristics, but also understand their causal role in belief formation. Therefore, instead of trying to inhibit the application of these heuristics directly, they prevent the acquisition of evidence that would trigger these heuristics.

The proposed error theory of epistemic rationality has important methodological consequences. Scientists might err in employing heuristics to justify inferences from evidence to belief in multiple ways. Therefore, an instrumentalist methodology must help in correcting these errors. In the first place, it should help scientist uncover the epistemic goal of their respective projects. Such goals are data—they are not to be prescribed by the methodology, but rather taken as inputs elicited from scientists. Secondly, the methodology should help the scientist describe the contexts in which these goals are to be met—including available background knowledge and the kind of data that might function as evidence. Thirdly, based this context, the methodology identifies the heuristic methods that are sufficient to satisfy the elicited goal. This identification can have both a prescriptive function—what method the scientist ought to choose, given her goal and the context of her project—as well as a critical function—what method the scientists should have chosen for an already concluded project. Heuristic instrumentalism thus is a powerful prescriptive methodology for scientific research.

6 Third challenge: the foundation of instrumental rationality

The first two challenges directly questioned the goal-dependence of method choice. The third challenge instead grants that method choice might proceed along instrumentalist considerations, but then argues that instrumental rationality itself depends on a goal-independent (and thus non-instrumental) form of rationality. The instrumentalism defended here claims it is rational to choose a method *M* only if *M* is sufficiently efficacious in bringing about the desired goal *G* in the relevant contexts. The non-instrumentalist critic then asks: what justifies the instrumentalist to claim that “*M* is sufficiently efficacious in bringing about *G*”? Presumably, there is some evidence that the instrumentalist can point to in order to support this claim. But this relation between evidence and claim—and in particular the normative claim that it is rational to belief this claim given the evidence—is not itself an issue of instrumental rationality, so the non-instrumentalist insists.

We have lots of empirical evidence that double-blind experimental procedures control for a source of bias – the placebo effect – that single-blind procedures do

not control for. That evidence gives us good reason to believe that double-blind methodology is a better indicator of the actual medicinal properties of drugs than single-blind methodology. Does this evidence constitute good reason for that claim only relative to some particular end? What end would that be? ... there is no such end. (Siegel 1996, p. S122).⁸

Siegel here does not deny that the choice of blinding is based on instrumental considerations. He even might allow that patient blinding isn't universally better—i.e. that for certain goals (e.g. investigating placebo effects), one might prefer to not blind patients. Instead, what he denies is that the evidence for this instrumental claim is itself *instrumentally* justified. One might cite evidence E as justification for why M_1 is a better means to G than M_2 . The quality of E, however, is not itself goal-dependent, neither on G nor some other goal H.

I disagree. In science there are many cases of goal-dependent evaluation of instrumental evidence. This is particularly obvious where non-epistemic considerations, e.g. the ethical concerns regarding animal and human experiments, play a role. Consider the goal of testing the efficacy of a novel drug, and imagine that in similar previous cases, it has been observed that tests involving bacterial cultures are better means towards this end than simulation studies. Typically, such evidence is quite uncertain—there will be only a handful of studies involving either method, so the reason is not particularly strong. But for the sake of the argument, let's assume that this evidence is deemed sufficient in this case. Now imagine that for the same goal, the method choice is between a human experiment and a simulation study. If the quality of the evidence for the instrumentality of human experiments were similarly weak as in the bacterial culture vs. simulation choice, one might well conclude that for this purpose, the evidence is *not* strong enough to constitute a reason for choosing the human experiment. Instead, involved researchers might conclude that the evidence for the higher instrumental value is too weak to outweigh the threat of harm that comes with human experiments.

This scenario shows that the assessment of evidence for instrumental justification of method choice is itself goal-dependent. In scientific practice, such differences in the evaluation of instrumental evidence are deeply institutionalized: while scientists are free to evaluate reasons for choosing various experimental methods that do not involve animal or human experiments, what counts as a reason to choose the latter is more circumscribed.

Thus, scientific methods are heuristics, and so are methodological rules that help choosing between alternative methods for some goal. Because methodological rules are heuristics, they are biased and fallible, just as the methods that they help choose between. But if methodological rules are heuristics in this sense, then one also must evaluate their justification—i.e. the evidence for the instrumental betterness of some methods over others—dependent on one's purpose. The above human experiment example illustrated such a purpose-dependent meta-meta-heuristic. But I see no reason

⁸ The terminology is not entirely settled here—most often, single-blind studies blind patients to their treatment allocation, while double-blind studies blind both patients and researchers to treatment allocations. Placebo effect biases are prevented by patient blinding. Siegel here seems to employ the terms single and double the other way around. However, this does not affect his main argument.

why the buck would stop here. Instead, it's purpose-dependent heuristics all the way up, and on each level they are instrumentally justified.

This “heuristics all the way up” may provoke two responses. First, why bother with justification at all, if they are always fallible? This question, I believe, is prompted by confusing “fallible” with “is not performing better than its competitors”. This is too big an argument to unpack here, so let me simply point to the idea of *ecological rationality*, according to which using heuristics—and choosing heuristically between different heuristics—is not just a second-best strategy but might indeed constitute fully rational choice. The basic idea, as proposed by Gigerenzer and collaborators, is that ecological rationality facilitates choosing the right heuristic for the right purpose in the right environment. Notably, in this view, being purpose-dependent and context-sensitive is an advantage of heuristics over general purpose, truth-preserving algorithms, not a flaw.

Heuristics *can* lead to more accurate judgments than strategies using more information and computation, including optimization methods, if one takes into account the relation between a reasoner's heuristics and his or her environment. (Gigerenzer and Sturm 2012)

Note the qualified claim “heuristics *can* lead...”—the justification of heuristic method choice, just like the heuristics themselves, are fallible: the respective methods/heuristics work sometimes, under some conditions, and for certain purposes.

The second response again invokes the regress worries that I discussed earlier. If the justification of method choice is fallible in this sense, then what justifies the method of justification, etc.? The worry is understandably exacerbated by the environment- and purpose-dependence of these justifications. Do scientists have to have at their ready a substantial collection of meta- and meta-meta etc. heuristics that justify their method choices dependent on which goals they have and in which environments they operate? That would be cumbersome, to say the least, and put in doubt my contention that a heuristic-instrumentalist account of scientific methodology is compatible with actual scientific practice.

I can answer this worry, again with reference to the heuristic nature of methods and the context-sensitivity of their success. If scientists indeed faced numerous different environments in random sequence, they perhaps would have to make do with such a cumbersome collection of meta-heuristics. But they don't. Science is a social enterprise, replete with institutional structure, including disciplinary divisions. Disciplines (at least the mature ones) provide relatively stable environments in which to do research in. They offer modes of problem representation and identify the menu of standard methods and goals, shared among members of a discipline. They also offer a shared history of heuristic use in environments that can be systematically compared to each other and to present contexts. This does not reduce research to the application of algorithms; environments still change, problems are still ill defined, methods are imprecise and don't always yield complete and unique solutions. But the stability of disciplinary division offers scientists ways of justifying their method choices—e.g. by relating to similar goals and by pointing to past successes in similar circumstances—without having to justify the meta-heuristic implicit in such moves. My argument from disciplinary stability then defangs the regress worries, without

abandoning my position that on all levels, justification is instrumental, context- and goal-dependent in nature.

7 Conclusion

I have argued for a heuristic-instrumental account of scientific methodology. I adopted the heuristic view of scientific methods (in the line of Wimsatt, Hey, etc.), extended it by grounding its epistemology in an instrumentalist account, then defended this epistemology against several lines of critique.

The heuristic view of methods afforded new arguments against three anti-instrumentalist claims. First, I showed the heuristic nature of many scientific inferential methods implies that they are indeed purpose-dependent, thus requiring a goal for their justification. Second, the heuristic view provided an error-theory of heuristics misapplication that explains, but does not vindicate, why some methods employed even though they do not serve any actual epistemic purpose. Third, I showed how “heuristics all the way up” is a viable alternative to more foundational sentiments, by relying on arguments from ecological rationality and disciplinary stability. Overall, the non-instrumentalist accounts of Kelly, Siegel and others fail both as descriptions of how scientists form beliefs as well as prescriptions of how scientists should choose their methods.

With these arguments I defended a heuristic-instrumental account of methodology against claims that parts of scientific methods must rely on non-instrumental, categorical epistemic rationality. This provides the conceptual basis for a powerful normative and prescriptive methodology, which justifies and critically assesses scientists’ method choice by linking them to their actual epistemic goals.

Acknowledgements I thank Caterina Marchionni and participants of a seminar at KTH for very instructive comments.

Funding Open access funding provided by Royal Institute of Technology.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Chow, S. J. (2014). Many meanings of ‘heuristic’. *The British Journal for the Philosophy of Science*, 66(4), 977–1016.
- Cleland, C. E. (2002). On effective procedures. *Minds and Machines*, 12(2), 159–179.
- Earman, J. (1992). *Bayes or bust*. Cambridge: MIT Press.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Heuristics: The foundations of adaptive behavior*. Oxford: Oxford University Press.

- Gigerenzer, G., & Sturm, T. (2012). How (far) can rationality be naturalized? *Synthese*, 187(1), 243–268.
- Hacking, I. (1988). Telepathy: Origins of randomization in experimental design. *Isis*, 79(3), 427–451.
- Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biology*, 13(3), e1002106.
- Hesse, M. (1975). Bayesian methods and the initial probability of theories. In G. Maxwell & R. M. Anderson (Eds.), *Induction, probability, and confirmation: Minnesota studies in the philosophy of science* (Vol. 6, pp. 50–105). Minneapolis: University of Minnesota Press.
- Hey, S. P. (2016). Heuristics and meta-heuristics in scientific judgement. *The British Journal for the Philosophy of Science*, 67(2), 471–495.
- Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research*, 66(3), 612–640.
- Kelly, T. (2007). Evidence and normativity: Reply to Leite. *Philosophy and Phenomenological Research*, 75(2), 465–474.
- Küppers, G., & Lenhard, J. (2005). Validation of simulation: Patterns in the social and natural sciences. *Journal of Artificial Societies and Social Simulation*, 8(4). Retrieved October 22, 2018, from <http://jasss.soc.surrey.ac.uk/8/4/3.html>.
- Lehmann, E. L. (1993). The Fisher, Neyman–Pearson Theories of testing hypotheses: One theory or two? *Journal of the American Statistical Association*, 88(424), 1242–1249.
- Leite, A. (2007). Epistemic instrumentalism and reasons for belief: A reply to Tom Kelly’s “Epistemic rationality as instrumental rationality: A critique”. *Philosophy and Phenomenological Research*, 75(2), 456–464.
- Lenhard, J. (2006). Models and statistical inference: The controversy between Fisher and Neyman–Pearson. *The British Journal for the Philosophy of Science*, 57(1), 69–91.
- Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford: Clarendon Press.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.
- Sharadin, N. (2018). Epistemic instrumentalism and the reason to believe in accord with the evidence. *Synthese*, 195(9), 3791–3809.
- Siegel, H. (1996). Instrumental rationality and naturalized philosophy of science. *Philosophy of Science*, 63, 117–124.
- Thrower, P. (2012). Eight reasons I rejected your article. A journal editor reveals the top reasons so many manuscripts don’t make it to the peer review process. Retrieved October 15, 2018, from <https://www.elsevier.com/connect/8-reasons-i-rejected-your-article>. Posted on 12 September 2012.
- Van Noorden, R., Maher, B., & Nuzzo, R. (2014). The top 100 papers. *Nature*, 514(7524), 550.
- Wadi, L., Meyer, M., Weiser, J., Stein, L. D., & Reimand, J. (2016). Impact of outdated gene annotations on pathway enrichment analysis. *Nature Methods*, 13(9), 705.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA’s statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133.
- Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Cambridge: Harvard University Press.
- Zwaaf, E. (2013). 8 reasons I accepted your article. Journal editors reveal the top reasons a manuscript gets published. Retrieved October 15, 2018, from <https://www.elsevier.com/connect/8-reasons-i-accepted-your-article>. Posted on 15 January 2013.