Synthese (2015) 192:3877–3913
DOI 10.1007/s11229-015-0870-6

S.I. : UNDERSTANDING THROUGH MODELING

On the role of explanatory and systematic power in scientific reasoning

Peter Brössel[1]

Received: 22 January 2014 / Accepted: 21 August 2015 / Published online: 14 September 2015
© Springer Science+Business Media Dordrecht 2015

Abstract The paper investigates measures of explanatory power and how to define the inference schema "Inference to the Best Explanation" (IBE). It argues that these measures can also be used to quantify the systematic power of a hypothesis and defines the inference schema "Inference to the Best Systematization" (IBS). It demonstrates that systematic power is a fruitful criterion for theory choice and that IBS is truth-conducive. It also shows that even radical Bayesians must admit that systematic power is an integral component of Bayesian reasoning. Finally, the paper puts the achieved results in perspective with van Fraassen's famous criticism of IBE.

1 Introduction

Arguably since Peirce, but at least since Harman (1965), philosophers have been debating whether *Inference to the Best Explanation* (IBE) is a legitimate form of rational inference. Harman describes this form of inference as follows: "one infers, from the premise that a given hypothesis would provide a 'better' explanation for the evidence than would any other hypothesis, to the conclusion that the given hypothesis is true" (Harman 1965, p. 89). If we introduce the background knowledge into the picture, as is nowadays standard, the resulting inference schema (IBE) can be portrayed as follows[1]:

[1] Some philosophers argue that one must relativize the third premise to a set of available hypotheses. Whether some hypothesis is the best explanation can only be evaluated in contrast to other hypotheses, so the argument goes.

✉ Peter Brössel
peter.broessel@rub.de

[1] Department of Philosophy and Center for Mind, Brain, and Cognitive Evolution, Ruhr-University Bochum, Bochum, Germany

Inference to The Best Explanation (IBE)

$$\frac{\begin{array}{c} B \\ E \\ H \textit{ is the best available explanation for } E \textit{ in the light of } B \\ \textit{(with respect to some set of available hypotheses)} \end{array}}{\textit{Therefore: } H}$$

Legend: B = background knowledge (henceforth, mention of background knowledge is suppressed), E = the evidence, H = a hypothesis.

For McMullin (1992), this is not just one inference schema among many others: for him IBE is the "The Inference That Makes Science." It is easy to provide examples from the history of science that can be taken to show that IBE is indeed crucial to scientific reasoning. Harman (1965) alludes to the following example by Darwin, who argues along these lines that the theory of evolution must be true:

> it can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified. It has recently been objected that this is an unsafe method of arguing; but it is a method used in judging of the common events of life, and has often been used by the greatest natural philosophers. (Darwin 1872, p. 421)

IBE plays an important role outside of science too. In particular, IBE plays an important role in many philosophical debates. As examples, one can cite the debates about scientific realism (for discussion see van Fraassen 1980, p. 19ff.) and external world skepticism (most explicitly Vogel 1990). With respect to the first debate, some philosophers argue that we are justified in believing that the theoretical terms of our best scientific hypotheses refer to objects in the world, since this is the best explanation for the apparent success of science. With respect to the second debate, other philosophers argue that we are justified in believing that there is an external world, since the existence of external world objects is the best explanation for our perceptions. For detailed information concerning the ubiquity of IBE in science and philosophy, see Douven (2011, Sect. 1.2).

Not all philosophers are so enthusiastic about the inference schema IBE, however. Van Fraassen, for example, often considered to be one of the fiercest opponents of IBE, writes the following:

> There are many charges to be laid against the epistemological scheme of Inference to the Best Explanation. One is that it pretends to be something other than it is. Another is that it is supported by bad arguments. A third is that it conflicts with other forms of change of opinion, that we accept as rational. (van Fraassen 1989, p. 142)

However, even van Fraassen, the figurehead of the anti-IBE camp, acknowledges the following:

> explanatory power is certainly *one* criterion of theory choice. When we decide to choose among a range of hypotheses, or between preferred theories, we evaluate

> each for how well it explains the available evidence. I am not sure that this evaluation will always decide the matter, but it may be decisive, in which case we choose to accept that theory which is the best explanation. But, I add, the decision to accept is a decision to accept as empirically adequate. (van Fraassen 1980, p. 71)

If even the fiercest opponents of the inference schema IBE accept the claim that explanatory power is one important criterion of theory choice, the only task remaining seems to be that of spelling out the vindication for that latter claim. Similarly, we have to ask whether the inference schema IBE as described by Harman (1965) can itself be vindicated. In this paper we investigate whether it is possible to provide such a vindication of explanatory power as a criterion of theory choice and of the inference schema IBE within Bayesian philosophy of science.

The paper has the following structure: Sect. 2 introduces various probabilistic measures of explanatory power that have been discussed in the literature. These measures enable us to compare the strength of the different explanations provided by different hypotheses and thus to define IBE. Common features of all these measures are (i) that they presuppose a notion of explanation and (ii) that the application of these measures for quantifying the level of explanatory power presupposes that the hypothesis in question is indeed an explanation for the evidence. In the context of theory choice, this presupposition is usually not satisfied. Even though the hypothesis might explain parts of the (total) evidence, almost no hypothesis explains all the evidence. Thus, Sect. 2 discusses what the proposed measures quantify, if the above presupposition is not satisfied, i.e., if the hypothesis does not explain the evidence. It is argued that in this case they are measuring the systematic power of the hypothesis with respect to the evidence. It is also argued that in the context of theory choice, we should take into account the entire systematic power of the hypotheses, not just their explanatory power. In addition, the corresponding inference schema *Inference to the Best Systematization* (IBS) is defined. Section 3 investigates whether it is possible to provide a vindication of systematic power as a criterion of theory choice and of the inference schema IBS. I argue that this is indeed the case. In particular, Sect. 3 demonstrates that in science, systematic power is a very fruitful criterion for theory choice: after finitely many pieces of evidence, and for every piece of evidence thereafter, (i) true hypotheses display a higher systematic power than false hypotheses, and (ii) logically stronger true hypotheses display a higher systematic power than logically weaker true hypotheses. In Sect. 3 I also demonstrate that the inference schema IBS is a fruitful inference schema in science, because it directs one to accept the logically strongest true hypotheses among the hypotheses available. The reason why we cannot achieve similar results for explanatory power as a criterion of theory choice and the inference schema IBE is also discussed in Sect. 3. Roughly, the reason is that our hypotheses cannot usually be considered an explanation of the total evidence available to us. Section 4 discusses how to reconcile consideration of explanatory and systematic power with Bayes' rule. More specifically, for radical Bayesians even though IBE and IBS are based on a probabilistic measure of systematic power, strictly speaking they are not Bayesian at heart, as they force agents either to accept a hypothesis or reject it for another. The fundamental epistemic norms of Bayesian epistemology recommend

assigning probabilistic degrees of belief to the hypotheses under consideration and to update these degrees of belief via Bayes' Rule. Radical Bayesians think there are no valid epistemic principles over and above the fundamental epistemic norms of Bayesian epistemology. However, Sect. 4 shows that Bayes' rule can be reformulated in such a way that one can see how explanatory power and systematic power both inform the degrees of belief of Bayesian agents. Section 5 discusses the results achieved in Sects. 3 and 4 and puts them in perspective with van Fraassen's famous criticism of IBE. Section 6 summarizes the findings.

2 From explanatory power to systematic power

2.1 Explanatory power

Suppose the hypothesis H provides an explanation for the evidence E. The core idea of all measures of explanatory power that have been proposed so far is this: how well a hypothesis explains the evidence depends on how much the hypothesis increases the probability of the evidence (i.e., our expectation of the observational data). This core idea encapsulates three minimal requirements on measures of explanatory power.

First, measures of explanatory power are defined in terms of probabilities. Formally, this requirement demands that measures of explanatory power be functions that assign a real number to each triple consisting of a probability function and a pair of propositions out of some algebra $\mathcal{A}$ of propositions (the first one: the hypothesis, the second one: the evidence).

Requirement 1 (measures of explanatory power 1) A measure of explanatory power is a function $\mathfrak{ep} : \Pr \times \mathcal{A} \times \mathcal{A} \to \mathbb{R}$. (In the following we write: $\mathfrak{ep}_{\Pr} : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$)

Probability theory is a branch of mathematics. Its axiomatic foundation was laid down by Kolmogorov in his (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*. The following definition of 'is a probability function' defines it over a set of possibilities, as is standard in contemporary epistemology.

Definition 1 Let W be a set of possibilities (e.g., possible worlds) and let $\mathcal{A}$ be a σ-algebra of subsets over W. A function $\Pr : \mathcal{A} \to \mathbb{R}$ is a probability function on $\mathcal{A}$ if and only if for all $A, B \in \mathcal{A}$:

1. $\Pr(A) \geq 0$
2. If $A = W$, then $\Pr(A) = 1$
3. $\Pr(A \cup B) = \Pr(A) + \Pr(B)$, if $(A \cap B) = \varnothing$

This definition of probability functions has to be supplemented by the definition of conditional probability.

Definition 2 If $\Pr(B) > 0$, then $\Pr(A|B) = \Pr(A \cap B)/\Pr(B)$

The reason why measures of explanatory power are explicitly relativized to probability functions is that we want to allow for different agents using the same measure of explanatory power but apply it to different probability functions. Which probability

functions are admissible and how these probability functions should be interpreted—as an agent's subjective or objective degree of belief function or as the objective chance function—, are interesting questions. In his 2009 paper "Locating IBE in the Bayesian Framework", Weisberg does not aim at defining explanatory power and IBE in terms of probabilities, but discusses whether IBE can be made compatible with Bayesian theories of rational reasoning. The conclusion Weisberg reaches concerning our latter question is this:

> What I have been trying to show is that compatibility with subjective Bayesianism infects IBE with the same limitations and counter-intuitiveness. My hope is that, having seen this, explanationists who are not already committed subjective Bayesians will appreciate the limitations of compatibilism. I hope that their explanationist inclinations will then compel them to reject Subjectivist Conditionalization, and to embrace a more full-blooded understanding of IBE. [...] If we grant IBE primacy instead, and use it to shape a more objective Bayesianism that rejects Subjectivist Conditionalization, we are in a position to develop a Bayesianism that is free of the limitations that come with Subjectivist Conditionalization. (Weisberg 2009, p. 136)

Thus, keeping our presupposition that explanatory power and IBE should be defined in terms of probabilities, we learn from Weisberg that we can define explanatory power and IBE in terms of subjective Bayesian probabilities; however, the resulting notions of explanatory power and IBE inherit the same limitations of such a definition. Instead, and still following Weisberg (2009), we should define explanatory power in objective Bayesian terms, and the probabilities should be interpreted as the "objectively correct distribution of 'a priori' probabilities" as Weisberg (2009, p. 137) puts it. For the remainder of this paper we do not take a stand on whether the respective probabilities underlying the definition of explanatory power and IBE should be interpreted as subjective or objective Bayesian probabilities or even as objective chances. For a detailed exposition of how our inductive probabilities should be interpreted see Brössel (2012). Brössel and Eder (2014) also address this topic very briefly.

Now let us return to the minimal requirements on measures of explanatory power. The second minimal requirement is that measures of explanatory power are measures of probabilistic relevance. More formally[2]:

[2] Note that the following Requirement 2 seems to admit of the possibility that some hypotheses can explain the evidence while actually lowering its probability. (The requirement specifies how a measure of explanatory power behaves in case $\Pr(E|H) < \Pr(E)$, even though we do not know whether the assumption that the hypothesis actually explains the evidence is compatible with the case of $\Pr(E|H) < \Pr(E)$. A Gricean implicature of this specification is that advocates of this requirement are ready to admit the possibility that a hypothesis explains the evidence even though $\Pr(E|H) < \Pr(E)$. If they do not admit that this is a possible case, why would they specify a necessary requirement for it?). Thus, it is an interesting question whether this is possible, and the answer certainly depends on the notion of explanation presupposed. For example, probabilistic theories of causation allow for the possibility of probability-lowering causes. Thus, given an understanding of explanation as casual explanation it seems possible to give an explanation of some observational fact by citing a cause that decreases the probability. For a discussion of the formal requirement in the context of confirmation theory, see Crupi (2007).

Requirement 2 (measures of explanatory power 2) A function[3] $\mathfrak{ep}_{\Pr}(\cdot,\cdot)$ is a measure of explanatory power relative to probability function Pr only if there is some $r \in \mathbb{R}$ such that: for all hypotheses H and evidence E if $1 > \Pr(E) > 0$ and $\Pr(H) > 0$, then

$$\mathfrak{ep}_{\Pr}(H,E)\begin{cases} > r, & \Pr(E|H) > \Pr(E) \\ = r, & \Pr(E|H) = \Pr(E) \\ < r, & \Pr(E|H) < \Pr(E) \end{cases}$$

The idea that the more a hypothesis increases the probability of the evidence, the better it explains the evidence, brings with it a third requirement. In particular, if two hypotheses H_1 and H_2 explain the evidence E, but E is more probable in the light of H_1 than in the light of H_2, then H_1 has more explanatory power with respect to E than H_2. More formally[4]:

Requirement 3 (measures of explanatory power 3) If function $\mathfrak{ep}_{\Pr}(\cdot,\cdot)$ is a measure of explanatory power relative to probability function Pr and $\Pr(E|H_1) > \Pr(E|H_2)$, then $\mathfrak{ep}_{\Pr}(H_1,E) > \mathfrak{ep}_{\Pr}(H_2,E)$.

An important advantage of this last requirement is that it links measures of explanatory power with the likelihood of the evidence in the light of the hypothesis. More specifically, many, if not all, Bayesians believe that the likelihood $\Pr(E|H)$ is less subjective than the posterior $\Pr(H|E)$ and the prior probabilities $\Pr(E)$ and $\Pr(H)$. Thus, Requirement 3 ensures that comparative judgments of the form 'with respect to the evidence E hypothesis H_1 displays a higher explanatory power than hypothesis H_2' are less subjective than judgments of the form 'in the light of the evidence E hypothesis H_1 displays a higher posterior probability than hypothesis H_2'.

In the spirit of these three requirements, various measures of explanatory power have been suggested. The three most popular measures of explanatory power will now be introduced and then be shown to satisfy requirements 1–3. Popper (1959) introduced the first measure of explanatory power, which is ordinally equivalent to the following one, ep^1, by Good (1960)[5]:

Definition 3 (*explanatory power 1*) If hypothesis H explains evidence E, then the explanatory power of H regarding E with respect to probability function Pr is:

$$ep^1_{\Pr}(H,E) = \frac{\Pr(E|H)}{\Pr(E)}$$

if $\Pr(H) > 0$ and $\Pr(E) > 0$.

[3] I adopt the following notational convention with respect to function terms: I use gothic fonts for function variables and normal calligraphic fonts for function constants.

[4] A closely related requirement can already be found in the work of Harman (1967) in connection with the inference schema IBE (even though it is restricted to statistical probabilities and statistical hypotheses). Harman calls it the *generalized maximum likelihood condition*. Crupi calls a related requirement for measures of confirmation the *Final Probability* requirement. For a discussion of the latter requirement see Crupi (2013).

[5] The original formulation of Popper's (1959) measure of explanatory power is this: $ep^{Popper}_{\Pr}(E,H) = \frac{\Pr(E|H)-\Pr(H)}{\Pr(E|H)+\Pr(H)}$.

Recently, Schupbach and Sprenger (2011) and Crupi and Tentori (2012) have suggested alternative measures. More specifically, for quantifying the explanatory power provided by a hypothesis regarding some evidence, Schupbach and Sprenger propose to employ the following measure, ep^2:

Definition 4 (*explanatory power 2*) If hypothesis H explains evidence E, then the explanatory power of H regarding E with respect to probability function Pr is:

$$ep^2_{\mathrm{Pr}}(H, E) = \left[\frac{\mathrm{Pr}(H|E) - \mathrm{Pr}(H|\neg E)}{\mathrm{Pr}(H|E) + \mathrm{Pr}(H|\neg E)}\right]$$

if $\mathrm{Pr}(H) > 0$ and $1 > \mathrm{Pr}(E) > 0$.

The measure of explanatory power suggested in Crupi and Tentori (2012) is the following, ep^3:

Definition 5 (*explanatory power 3*) If hypothesis H explains evidence E, then the explanatory power of H regarding E with respect to probability function Pr is:

$$ep^3_{\mathrm{Pr}}(H, E) = \begin{cases} \frac{\mathrm{Pr}(E|H)-\mathrm{Pr}(E)}{1-\mathrm{Pr}(E)} & \text{if } \mathrm{Pr}(E|H) \geq \mathrm{Pr}(E) > 0 \\ \frac{\mathrm{Pr}(E|H)-\mathrm{Pr}(E)}{\mathrm{Pr}(E)} & \text{if } \mathrm{Pr}(E|H) < \mathrm{Pr}(E) \end{cases}$$

For all three measures of explanatory power one can show that they not only capture the spirit of requirements 1–3, but that they satisfy them.

Theorem 1 (ep^1, ep^2, ep^3 and the Requirements 1–3) *For all probability functions* Pr, *the three measures of explanatory power* ep^1_{Pr}, ep^2_{Pr}, *and* ep^3_{Pr} *satisfy the requirements* 1–3.
(The proof for this theorem can be found in the appendix.)

On the basis of each of these measures of explanatory power, one can define under which conditions a hypothesis can be considered the best explanation, for some evidence, of all the available hypotheses. Three possible definitions come to mind.

Definition 6 (*best explanation 1*) For all probability functions Pr, all measures of explanatory power $\mathfrak{ep}_{\mathrm{Pr}}(\cdot, \cdot)$ and all sets of hypotheses $\{H_1, \ldots, H_n\}$ and all bodies of evidence E with $H_i \in \mathcal{A}$ and $E \in \mathcal{A}$:
H_i is the best available explanation for E with respect to the set of available hypotheses $\{H_1, \ldots, H_n\}$ iff H_i explains E and for all hypotheses H_j that explain E (with $i \neq j$): $\mathfrak{ep}_{\mathrm{Pr}}(H_i, E) > \mathfrak{ep}_{\mathrm{Pr}}(H_j, E)$.

According to this definition, the best explanation for the evidence is simply the hypothesis that displays the highest explanatory power with respect to that evidence of all the hypotheses that explain the evidence. However, since we do not know whether explanatory power is truth conducive, one might argue that this conception of 'best explanation' is not adequate. A hypothesis might display considerable explanatory power with respect to the evidence even though it is very implausible. In that case, one might argue that it is not a good explanation and, therefore, that it is possibly not the best explanation for the evidence. An alternative definition, which takes these worries into account, is the following:

Definition 7 (*best explanation 2*) For all probability functions Pr, all measures of explanatory power $\mathfrak{ep}_{\Pr}(\cdot,\cdot)$ and all sets of hypotheses $\{H_1,\ldots,H_n\}$ and all bodies of evidence E with $H_i \in \mathcal{A}$ and $E \in \mathcal{A}$:
H_i is the best available explanation for E with respect to the set of available hypotheses $\{H_1,\ldots,H_n\}$ iff H_i explains E and for all hypotheses H_j that explain E (with $i \neq j$): $\mathbf{E}\mathfrak{ep}_{\Pr}(H_i,E) > \mathbf{E}\mathfrak{ep}_{\Pr}(H_j,E)$, where $\mathbf{E}\mathfrak{ep}_{\Pr}(H_i,E) =_{def} \Pr(H_i|E) \times \mathfrak{ep}_{\Pr}(H_i,E)$.
(*Note that for calculating and comparing expected explanatory power, measures of explanatory power that allow for negative values of explanatory power have to be rescaled in order to exclude these negative values and to ensure that the minimal value is 0. This applies in particular for the measures* $ep^2_{\Pr}$ *and* $ep^3_{\Pr}$.)

The following alternative definition also takes the plausibility of the hypotheses into account. In addition, it is considerably simpler than the previous one and closer to the one suggested by Harman (1967):

Definition 8 (*best explanation 3*) For all probability functions Pr, all measures of explanatory power $\mathfrak{ep}_{\Pr}(\cdot,\cdot)$ and all sets of hypotheses $\{H_1,\ldots,H_n\}$ and all bodies of evidence E with $H_i \in \mathcal{A}$ and $E \in \mathcal{A}$:
H_i is the best available explanation for E with respect to the set of available hypotheses $\{H_1,\ldots,H_n\}$ iff (i) H_i explains E, (ii) $\Pr(H_i|E) > .5$, and (iii) $\mathfrak{ep}_{\Pr}(H_i,E) > \mathfrak{ep}_{\Pr}(H_j,E)$, for all hypotheses H_j such that $\Pr(H_j|E) > .5$ and H_j explains E (with $i \neq j$).

It is important to stress that these measures presuppose that we apply them to explanatory hypotheses and that according to the above definitions the best available explanation among the available hypotheses $H_1,\ldots,H_n$ is indeed an explanation of the evidence E. Accordingly, the presented theories of explanatory power must presuppose that we have an adequate theory of explanation at our disposal. However, even if an adequate theory of the qualitative notion of explanation is available, there is another more troublesome problem. In the context of rational inference from evidence to hypothesis, it is usually the case that a hypothesis does not explain all the evidence available to the agent. Consider the following toy example: suppose on a flat and level piece of ground there are two flagpoles p_1 and p_2. Let the background knowledge B say this and that p_1 is 10 m tall (the length of p_2 is not specified in the background knowledge) and that at t_2 the length of the shadow of p_2 is 15 m. Then we receive three pieces of evidence: e_1 states that at t_1 the length of the shadow of p_1 is 8391 m, e_2 states that the length of flagpole p_2 is 21.422 m, and e_3 states that at t_1 the shadow of flagpole p_2 was 17.975. From this evidence we can conclude that the following hypothesis must be true. H: between t_1 and t_2 the sun's elevation changed from 50 to 55. Nevertheless, in the light of B the following holds: (i) H explains e_1, (ii) H predicts, but does not explain e_2 (H predicts e_2 in the sense that one can derive the latter from the former. H does not explain e_2 because the length of the shadow of a flagpole mentioned in B does not explain the pole's length, and since e_2 is about the length of the flagpole p_2, which presumably is the same at t_1 and at t_2, H and B can be considered to provide a retrodiction of the length of the flagpole p_2 at t_1 and a prediction of its length at t_2.), and finally (iii) H retrodicts, but neither explains nor predicts e_3 (H does not predict e_3 because e_3 takes place before the relevant conditions mentioned in B take place. H does not explain e_3 because the length of the shadow

of a flagpole at t_2 mentioned in B and H do not explain the pole shadow's length at t_1 as mentioned in e_3, H retrodict e_3 in the sense that one can derive the latter from the former given the background B and e_3 happens before the relevant conditions stated in B). Thus, even though our hypothesis H does not explain all the available evidence, it is nevertheless able to "systematize" the available evidence in the light of the specified background knowledge and we do not hesitate to infer its truth.

More generally, it is possible that the evidence E available to the scientists consists of three different pieces of evidence e_1, e_2, and e_3, but that the hypothesis H under consideration explains e_1, predicts but does not explain e_2, and retrodicts e_3. By assumption, H is not an explanation of E in this case and thus none of the measures of explanatory power introduced above can be employed to measure the explanatory power of H with respect to E. Even more importantly, this shows that the above measures of explanatory power typically can only be applied if the evidence is restricted; comprehensive bodies of evidence typically cannot be completely explained by any hypothesis and, thus, IBE is typically not applicable to large bodies of evidence. Already Harman (1967), who suggested the best developed formal account of IBE so far, sees this problem. He writes:

> I have not presented a set of sufficient conditions [for the acceptance of hypotheses], indeed a version of the lottery paradox is not met by the set of conditions I have mentioned. For it is possible that there are N different explanations, each accounting for a different fraction of the total evidence, each satisfying the conditions so far mentioned, even though the evidence also ensures that one of these explanations cannot be correct. (Harman 1967, p. 410)

And this diagnosis is also in line with Rescher (2005), who argues that

> [t]he inference from explanatory optimality of truth is impeded by the consideration that explanatory optimality is generally a local phenomenon whereas truth is by nature global and context-independent. In terms of the practical politics of the matter, optimal systematization is the best we can do. (Rescher 2005, p. 100)

It is important to note that Rescher does not have a specific measure of explanatory power or a certain specified inference schema in mind when he criticizes IBE; nor does he have in mind a specific measure of systematic power or a certain inference schema when he defends IBS. For agreeing with his conclusion, the only important consideration is that typically hypotheses cannot explain the total evidence available to the agent, but only parts of this evidence. However, for scientific inference the agent should take into account the total evidence available to her and not just the part of the evidence which her hypotheses can explain. Thus, the conclusion is acceptable if and only if the respective notion of explanation is not very permissive. In his discussion of systematization and IBE, however, Weisberg (2009) notes that

> [a]ccording to the best-systems view, something is a law of nature just in case it is a theorem in all of the true deductive systems that best balance simplicity and informative strength (Lewis 1973). […] One way to explain, says the best-systems advocate, is to locate the explanandum in a simple and orderly overall picture. So when the Humean explains the color of this emerald by appeal to

> the general law, she explains by locating this particular, local matter of fact in a simple, unifying, and informative picture of all the particular, local matters of fact. (Weisberg 2009, p. 139)

If one follows the Humeans, as portrayed by Weisberg, in this regard, then one could argue that prediction and retrodictions also locate "a particular, local matter of fact in a simple, unifying, and informative picture of all the particular, local matters of fact" and that therefore they are in a wide sense explanatory as well. One also might insist that we keep on calling the relevant inference schema 'Inference to the Best Explanation' even though not all matters of facts can be explained in the sense of the more restrictive causal theories of explanation. Summing up, if we do not have a very permissive notion of explanation, we can follow Rescher who concludes:

> Systematization is a resource of cognitive validation that is significantly different from explanation. Explanation is a retail commodity: one generally explains facts one at a time. But systematization is a wholesale commodity. […] We cannot appropriately "infer" the best explanation E_1 of a fact f_1 precisely because there may be some other fact f_2 whose best explanation E_2 is incompatible with that aforementioned E_1. But with systematization the matter stands differently. By its very nature as such, systematization must be coherent overall. (Rescher 2005, p. 103)

2.2 Systematic power

Discussing Rescher's conceptions of IBE and IBS and all his arguments against or respectively for them would take us beyond the scope of the present paper.[6] Instead, the goal is to offer a precise notion of systematic power and IBS that supersedes Rescher's vague notions. To start with, we suggest that the above measures of explanatory power can quantify the *systematic power* of H with respect to E, if we drop the assumption that the hypothesis explains the evidence. Indeed, this is the natural proposal in the light of Hempel and Oppenheim's famous symmetry thesis concerning explanations and predictions, which was first introduced in their seminal 1948 paper, "Studies in the Logic of Explanation".

> [I]t seems sometimes possible to compare different theories, at least in an intuitive manner, in regard to their explanatory, or predictive powers: Some theories seem powerful in the sense of permitting the derivation of many data from a small amount of initial information, others seem less powerful, demanding comparatively more initial data, or yielding fewer results. Is it possible to give a precise interpretation to comparisons of this kind by defining, in a completely general manner, a numerical measure for the explanatory or predictive power of a theory? […] Since explanation and prediction have the same logical structure, namely

[6] Especially since such an undertaking would not only require a great deal of interpretational work, but we would also have to elaborate the basis upon which we disagree with Rescher's arguments.

> that of a deductive systematization, we shall use the neutral term "systematic power" to refer to the intended concept. (Hempel and Oppenheim 1948, p. 164)

In his 1958 "The Theoretician's Dilemma—A Study in the Logic of Theory Construction", Hempel makes a more fine-grained distinction between explanations, predictions, and retrodictions—or as he calls them, following Reichenbach, "postdictions"— and still upholds the symmetry thesis:

> Scientific explanations, predictions, and postdictions all have the same logical character: they show that the fact under consideration can be inferred from certain other facts by means of specified general laws. In the simplest case, the type of argument may be schematized as a deductive inference of the following form:
>
> $$\frac{\begin{array}{c} C_1, C_2, \ldots, C_k \\ L_1, L_2, \ldots, L_r \end{array}}{E}$$
>
> Here $C_1, C_2, \ldots, C_k$ are statements of particular occurrences (e.g., of the position and momenta of certain celestial bodies at a specified time), and $L_1, L_2, \ldots, L_r$ are general laws (e.g., those of Newtonian mechanics); finally E is a sentence stating whatever is being explained, predicted, or postdicted. And the argument has its intended force only if its conclusion, E, follows deductively from the premises. (Hempel 1958, pp. 37–38)

More importantly, Hempel admits the possibility of probabilistic explanations, predictions and postdictions, though all the while upholding the symmetry thesis. Consequently, Hempel proposes to distinguish between deductive and inductive systematization. In particular, there are cases in which:

> the statement E describing the occurrence under explanation or prediction or postdiction (for example, Johnny's catching the measles) is not logically deducible from the explanatory statements adduced (for example, (C_1) Johnny was exposed to the measles; (C_2) Johnny had not previously had the measles; (L) For persons who have not previously had the measles and are exposed to it, the probability is .92 that they will contract the disease); rather, on the assumption that the explanatory statements adduced are true, it is very likely, though not certain, that E is true as well. This kind of argument, therefore, is inductive rather than strictly deductive in character […]. An argument of this kind—no matter whether it is used for explanation, prediction, or postdiction, or for yet another purpose—will be called an *inductive systematization*. (Hempel 1958, pp. 39–40)

Following philosophers such as Hempel and Salmon, we now reject Hempel and Oppenheim's original suggestion concerning the deductive structure of explanations, retrodictions, and predictions and we also allow for probabilistic or inductive explanations, predictions, and retrodictions. Following the above considerations we might rephrase the original quote by Hempel and Oppenheim and say that "since explanation[, retrodiction,] and prediction have the same logical [or probabilistic] structure, namely that of a deductive [or inductive] systematization, we shall use the neutral

term "systematic power" to refer to the intended concept." Finally, the idea that "the notions of explanatory and predictive power can be combined within the notion of systematic power" (Niiniluoto 2011, Sect. 3.3) is still sound and is generally accepted. The reason for this is that if $\Pr(E|H) > \Pr(E)$, but H is not an explanation for E, then H is a more or less good retrodiction or prediction of E, and how well H predicts or retrodicts E depends on how much H increases the probability of E (i.e., how much H increases our expectation of the data). If $\Pr(E|H) \leq \Pr(E)$, we might discuss whether we want to say that hypothesis H is an explanation (see footnote 2), retrodiction, or prediction of the evidence E in the qualitative sense of these words. However, it is natural to use the quantitative notions of degree of explanatory, retrodictive, or predictive power even in these cases. (At least it is as natural as in the case of confirmation, where it is standard to speak of degree of confirmation even though in the qualitative sense of the word E does not confirm H in case $\Pr(H|E) \leq \Pr(H)$.) Accordingly, given our assumption that the measures introduced above are indeed adequate measures of explanatory power, it is fair to presuppose that they can be used as measures of systematic power in those contexts in which the hypothesis does not explain the total evidence available to the agent. Indeed, from a Bayesian perspective the only consideration that would speak against such a generalization would be if one argued that though some measure of explanatory power is an adequate measure of explanatory power, given that H explains E, it is not a measure of H's predictive power with respect to E, given that H predicts, but does not explain E. However, given that measures of explanatory power are closely related to confirmation measures, it seems implausible that such argumentation is possible. For example, in the context of measures of explanatory power, Schupbach, one of the modern champions of theories of explanatory power, says: "these measures are structurally equivalent to the confirmation measures" (Schupbach 2011, p. 814). In particular, measures of confirmation quantify how much the evidence increases the probability of the hypothesis, and measures of explanatory power quantify how much the explanans (the hypothesis) increases the probability of the explanandum (the evidence). The measure of explanatory power endorsed by Schupbach and Sprenger (2011), for example, is structurally equivalent to the measure of factual support proposed by Kemeny and Oppenheim (1952), the measure of explanatory power suggested by Crupi and Tentori (2012) is structurally equivalent to the measure of confirmation suggested by Crupi (2007), and the measure of explanatory power suggested by Popper (1959) is structurally equivalent to measures of confirmation endorsed by Horwich (1982), Keynes (1921), and Milne (1996). Thus, the picture that these champions of explanatory power advocate is this: *if* some hypothesis H explains the evidence, then we can gauge the explanatory power of some hypothesis concerning the evidence by measuring how much the hypothesis confirms the evidence. In a second step, it is only natural to say that if the hypothesis does not explain the evidence, we can use the same measure of confirmation to determine how strongly the hypothesis predicts or retrodicts the evidence.

Suppose the notions of explanatory, retrodictive, and predictive power can be combined within the notion of systematic power. This raises the question whether considerations of systematic power should play a crucial role in scientific inference. In the context of rational inference from evidence to hypothesis, it is a natural idea that we select a hypothesis on the basis of its systematic power (which, of course, includes

its explanatory power) and not just its explanatory power alone. This is already indicated in the quote by van Fraassen, when he writes that "explanatory power is certainly *one* criterion of theory choice", thereby implying that there are other criteria as well. To further motivate this idea, consider the following example: Newton's theory of gravitation is typically considered to describe the effects of gravitation satisfactorily, but not to explain how or why gravity can be causally effective. As Newton himself puts it in his *General Scholium* from the *Mathematical Principles of Natural Philosophy*:

> I have not been able to discover the cause of those properties of gravity from phœnomena, and I frame no hypothesis. …to us it is enough, that gravity does really exist, and acts according to laws which we have explained, and abundantly serves to account for all the motions of the celestial bodies, and of our sea. (Newton 1729, pp. 506–507)

According to causal theories of explanation, Newton's theory does not explain the motions of the planets in our system, although with the help of his theory we can describe, predict and retrodict the motions of the planets. However, we cannot explain them satisfactorily, since Newton's theory does not state the cause of these motions. Nevertheless, we are strongly inclined to consider Newton's theory to have a high systematic power with respect to our evidence concerning the motion of planets independently of the question of whether Newton's theory can causally explain these motions or not. This systematic power of Newton's theory of gravitation concerning the motion of planets led us to consider it the best theory available until Einstein's theory of relativity. This illustrates that the inference schema that we use in the context of theory choice is best described as IBS. Therefore, let us drop the presupposition that the hypotheses in question actually explain the evidence and rename the measures ep^1_{Pr}–ep^3_{Pr} as sp^1_{Pr}–sp^3_{Pr}, for *systematic power*. Then, we can define which hypothesis is the best (available) systematization of the evidence as follows:

Definition 9 (*best systematization 1*) For all probability functions Pr, all measures of systematic power $\mathfrak{sp}_{\mathrm{Pr}}(\cdot,\cdot)$ and all sets of hypotheses $\{H_1,\ldots,H_n\}$ and all bodies of evidence E with $H_i \in \mathcal{A}$ and $E \in \mathcal{A}$:
H_i is the best available systematization for E with respect to the set of available *hypotheses* $\{H_1,\ldots,H_n\}$ *iff for all hypotheses* H_j (*with* $i \neq j$): $\mathfrak{sp}_{\mathrm{Pr}}(H_i,E) > \mathfrak{sp}_{\mathrm{Pr}}(H_j,E)$.

Note that the only difference between Definition 6 and Definition 9 is that for the former we presuppose that the hypotheses in question actually explain the evidence and in the latter definition this presupposition is dropped.

Again, one might want argue that this conception of 'best systematization' is not adequate. A hypothesis might display a high systematic power with respect to the evidence even though it is very implausible. In that case, one might argue that it is not a very good systematization and therefore possibly not the best systematization for the evidence. Alternative definitions, that take this worry into account, are the following:

Definition 10 (*best systematization 2*) For all probability functions Pr, all measures of systematic power $\mathfrak{sp}_{\mathrm{Pr}}(\cdot,\cdot)$ and all sets of hypotheses $\{H_1,\ldots,H_n\}$ and all bodies of evidence E with $H_i \in \mathcal{A}$ and $E \in \mathcal{A}$:

H_i is the best available systematization for E with respect to the set of available *hypotheses* $\{H_1, \ldots, H_n\}$ *iff for all hypotheses* H_j *(with* $i \neq j$*):* $\mathbf{E}\mathfrak{sp}_{\mathrm{Pr}}(H_i, E) > \mathbf{E}\mathfrak{sp}_{\mathrm{Pr}}(H_j, E)$*, where* $\mathbf{E}\mathfrak{sp}_{\mathrm{Pr}}(H_i, E) =_{def} \mathrm{Pr}(H_i|E) \times \mathfrak{sp}_{\mathrm{Pr}}(H_i, E)$.
(*Note that for calculating and comparing expected systematic power, measures of systematic power that allow for negative values of systematic power have to be rescaled in order to exclude these negative values and to ensure that the minimal value is 0. This applies in particular for the measures* sp^2_{Pr} *and* sp^3_{Pr}.)

Definition 11 (*best systematization 3*) For all probability functions Pr, all measures of systematic power $\mathfrak{sp}_{\mathrm{Pr}}(\cdot, \cdot)$ and all sets of hypotheses $\{H_1, \ldots, H_n\}$ and all bodies of evidence E with $H_i \in \mathcal{A}$ and $E \in \mathcal{A}$:
H_i is the best available systematization for E with respect to the set of available hypotheses $\{H_1, \ldots, H_n\}$ iff (i) $\mathrm{Pr}(H_i|E) > .5$ and (ii) for all hypotheses H_j: $\mathrm{Pr}(H_j|E) \leq .5$ or $\mathfrak{ep}_{\mathrm{Pr}}(H_i, E) > \mathfrak{ep}_{\mathrm{Pr}}(H_j, E)$.

On the basis of Definitions 9–11, we can also provide the definition of the corresponding inference schema *Inference to the Best Systematization*:

Inference to The Best Systematization (IBS)

$$\frac{\begin{array}{c} E \\ H \text{ is the best systematization for } E \text{ with respect to the set of available hypotheses } \{H_1, \ldots, H_n\} \end{array}}{\text{Therefore: } H_i}$$

Given these definitions, the question is now whether systematic power can serve as a criterion of theory choice and whether the inference schema IBS leads us to accept true hypotheses.

3 From systematic power to theory choice

As discussed above, in most cases the inference schema that we use in the context of theory choice is better described as *Inference to the Best Systematization* instead of *Inference to the Best Explanation*. The inference schema should be understood as *Inference to the Best Explanation* if and only if the hypothesis explains the agent's total evidence—because then the systematic power of the hypothesis is nothing other than its explanatory power. However, in general this presupposition is not satisfied and there are theoretical virtues besides explanatory power, i.e., the predictive and retrodictive power of the hypothesis. We can now turn to the questions of (i) whether systematic power can serve as a criterion of theory choice and (ii) whether we can justify the inference from the premise that a given hypothesis would provide the best systematization for the evidence to the conclusion that the given hypothesis is acceptable or true. To begin, let us concentrate on the first question.

3.1 Systematic power as a criterion for theory choice

Can systematic power serve as a criterion of theory choice? This question can indeed be answered positively, at least for those Bayesians who consider the convergence

theorems of, for example, Gaifman and Snir (1982) or Schervish and Seidenfeld (1990) to be a success story of Bayesian epistemology that confers at least partial vindication on the Bayesian norms of reasoning. In particular, we can utilize the convergence theorems to prove that all three measures of systematic power introduced above are truth-conducive. Since this result depends heavily on the convergence theorems, let us review one of it briefly.

First, we present the Gaifman–Snir Theorem [for a proof see Gaifman and Snir (1982)].

The Gaifman–Snir Theorem *Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. Now let e_1,..., e_n, ...be a sequence of propositions of $\mathcal{A}$ which separates W, and for all $w \in W$ let $e_i^w = e_i$ if $w \vDash e_i$, and $\neg e_i$ otherwise. Let* Pr *be a probability function on $\mathcal{A}$. Let* Pr* *be the unique σ-additive probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying* $\Pr^*(A) = \Pr(A)$ *for all $A \in \mathcal{A}$.*
Then there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ *so that the following holds for every $w \in W'$ and all propositions A of $\mathcal{A}$:*

$$lim_n \Longrightarrow_\infty \Pr(A|E_n^w) = \mathcal{I}(A, w)$$

where $\mathcal{I}(A, w) = 1$, if $w \vDash A$ and 0 otherwise.

(*The elements of the algebra $\mathcal{A}$ are interpreted as propositions expressible in some language $\mathcal{L}$ suitable for arithmetic. In particular, let $\mathcal{L}$ be some first-order language containing the numerals '1', '2', '3', ...as names, respectively, individual constants, and symbols for addition, multiplication, identity etc. In addition, let $\mathcal{L}$ contain finitely many relations and functional symbols. Gaifman and Snir* (1982) *call them the 'empirical symbols'. Accordingly we can think of the possibilities in W as models for that language $\mathcal{L}$, which agree on the interpretation of the mathematical symbols but can disagree on the interpretation of the empirical symbols. In addition, the elements of W are thought of as models that satisfy the following property: if a statement of the form $\exists x\phi[x]$ is true in a model then there is an individual constant c such that the proposition $\phi[c]$ is true in that model. This ensures that we can think of the proposition expressed by the formula $\exists x\phi[x]$ as the infinite disjunction of the propositions expressed by formulae of the form $\phi[c]$. In addition, Gaifman and Snir also require that the probability function* Pr *satisfies the following property:* $\Pr(\exists x\phi[x]) = \lim_{n\to\infty} \Pr(\bigcup \phi[a_i]))$

It is part of the Bayesian folklore that these convergence theorems represent a vindication of Bayesian norms of reasoning. The standard interpretation of this theorem is that it shows that in the long run the subjective probability of a proposition converges to the proposition's truth value (almost surely). To ensure that a Bayesian agent's subjective probabilities converge to the truth value of the respective propositions in some possible world $w \in W$, three conditions must be satisfied.

First, there must be a sequence e_1, ..., e_n, ...of propositions of $\mathcal{A}$ which separates W (Of course the elements of $\mathcal{A}$ are still to be interpreted as propositions expressible in some language $\mathcal{L}$ as required by the Gaifman–Snir Theorem (see above).). A sequence of propositions e_1, ..., e_n, ...separates the set of possibilities W if and only if for every pair of worlds w_i and $w_j \in W$ (with $w_i \neq w_j$) there is one proposition in the sequence such that it is true in one of the possible worlds and false in the other. (Given our interpretation of the elements of W, we can say that a sequence of propositions e_1, ..., e_n, ...separates the models of some language $\mathcal{L}$—i.e., W—if and only if for

all models M_1 and M_2 such that $M_1 \neq M_2$ there is a statement e_i in the sequence that is true in one of the models but false in the other.) The assumption of the mere existence of such a sequence is unproblematic (at least from a purely mathematical point of view). Indeed, the sequence of all atomic sentences of the language $\mathcal{L}$ is such a separating sequence.

Second, the agent must determine directly (i.e., without additional Bayesian inferences), step by step for each of these propositions, whether it is true or false in a respective world w, and thus the agent's evidence e_i^w at stage i of the sequence in possible world w is e_i if $w \vDash e_i$ and $\neg e_i$ otherwise. This second condition is not unproblematic. If an agent only misses one observation, there might be a proposition such that our Bayesian agent's degree of belief in that proposition does not converge to its truth value. The assumption becomes obviously problematic if we consider that there might be propositions in the sequence such that our Bayesian agent is not in a position to directly determine its truth value. We recall, in line with the Gaifman–Snir Theorem, that the elements of $\mathcal{A}$ can be interpreted as propositions expressible in some language $\mathcal{L}$. If that language contains theoretical vocabulary, e.g., the predicate 'negatively charged particle', then we cannot determine the truth value of a proposition expressing that some object is negatively charged. Now suppose that there are two possible worlds whose main difference is that a is a negatively charged particle in one of the worlds but not in the other (and of course the differences that come with the latter, let us suppose for the sake of the argument these differences are not observable). Then, according to our separability requirement, there must be a proposition e_i that is true in one possible world and false in the other, where we can directly determine the truth value. However, this seems to be impossible in this case, since the respective predicate, 'negatively charged particle', is a theoretical predicate. One might argue that this criticism is not problematic per se, since an inductive method is of course only as good as its inductive basis: the evidence. Suppose you failed to observe the only pink elephant that will ever exist and you only see grey ones. Nobody would blame a given inductive method for not leading you to believe that there are pink elephants given the evidence that in the past and in the future all observed elephants are grey. Or suppose, for example, that the Bayesian agent is colorblind: no one then blames the Bayesian norms of reasoning if this agent cannot learn the truth about specific animal-color hypotheses, even though she follows the Bayesian norms. The point can be generalized to deductive inferences, as has already been noted by Schupbach (2014):

> [M]odus ponens itself provides us with no reason to believe that we will instantiate it with true premises. The same point holds for any inference form: by virtue of their formal character, they provide us with few constraints on the quality of the material that may be used to instantiate them on any occasion. But when working with bad material content, virtually any inference form will likely commend a false conclusion. (Schupbach 2014, p. 58)

Thus, we should consider a given set of norms of inductive inference (in this case systematic power as a criterion of theory choice) as justified if and only if following these norms would lead us to true and informative hypotheses, provided the premises (i.e., the evidence) of an inductive inference are true and sufficiently informative.

Nevertheless, Bayesians have to admit that even if humans were Bayesian agents, the Bayesian norms of reasoning would not guarantee epistemic success, since human agents are not ideal observers either.

Third, the possible world $w \in W$ is not in a set of possible worlds with measure zero with respect to the agent's probability Pr. Or, to put it differently, there are possible worlds in which the agent's degrees of belief are not guaranteed to converge to the truth value of the respective propositions, but these are possible worlds to which the agent assigns an a priori probability of zero.

Here is not the place to take up the discussion of in how far the convergence results, given the presuppositions that they make, provide a vindication for Bayesian norms of reasoning. For detailed discussion of the claim that these convergence results (at least partially) vindicate the Bayesian norms of reasoning, see Belot (2013), Earman (1992), Hawthorne (2014), and Huttegger (2015a, b). For applications of these results in Bayesian philosophy of science see for example Brössel (2008, 2014, 2015), Huber (2008), and Huttegger (2015a, b). The following theorem shows that systematic power as a criterion of theory choice is vindicated by the convergence theorems to the same extent as Bayesian norms of reasoning in general are vindicated by these theorems:

Theorem 2 (truth-conduciveness of systematic power) *Let W be a set of possible worlds and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ which separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let* Pr *be a strict (or regular) probability function on $\mathcal{A}$.*[7] *Let* $\Pr^*$ *be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$. Then there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all hypotheses $H_1, H_2 \in \mathcal{A}$ and for all $\mathfrak{sp}_{\Pr}$ satisfying Requirements* 1–3.

1. *if $w \vDash H_1$ and $w \vDash \neg H_2$, then:*
   $\exists n \forall m \geq n : [\mathfrak{sp}_{\Pr}(H_1, E_m^w) > \mathfrak{sp}_{\Pr}(H_2, E_m^w)]$.
2. *if $w \vDash H_1 \cap H_2$ and $H_1 \vDash H_2$ but $H_2 \nvDash H_1$, then:*
   $\exists n \forall m \geq n : [\mathfrak{sp}_{\Pr}(H_1, E_m^w) > \mathfrak{sp}_{\Pr}(H_2, E_m^w)]$.

where $E_m^w = \bigcap_{0 \leq i \leq m} e_i^w$.

[7] In this theorem and in Theorem 3 and Corollary 1 below we assume that the probability function is a strict or regular probability function. Thus, in Howson's words, the theorems again confirm

> Hume's argument that there is no sound inductive argument from experiential data that does not incorporate an inductive premise, and it also tells us what the inductive premise will look like: it will be a probability assignment that is not deducible from the probability axioms. (Howson 2003, p. 134)

Without the "inductive premise" that we are dealing with a strict or regular probability function, we would have to replace the '$>$' in Point 2 of Theorems 1–3 by '$\geq$'. In consequence, considerations of systematic power could not distinguish between a logically stronger true hypothesis H_1 and a logically weaker true hypothesis H_2, if both of them had the same prior probability despite the difference in their logical strength. Indeed no purely probabilistic inference rule could distinguish between them, since from a probabilistic perspective there would be no discernible difference between the hypotheses, because in this case $\Pr(H_2 \rightarrow H_1) = 1$ and, thus, Bayesians would treat $(H_2 \rightarrow H_1 \wedge H_2)$ as if it were a logical truth. For a longer and more comprehensive discussion of the connection between logical strength, prior probabilities (or informativity) and theory choice see Brössel (2014).

(The proof for this theorem can be found in the appendix.)

According to this theorem, all measures of systematic power that satisfy the minimal requirements 1–3 assign a higher degree of systematic power to true hypotheses than to false hypotheses if we confront them with the total evidence available to an agent: after receiving finitely many pieces of evidence (in a sequence of separating observational statements), and for every piece of evidence thereafter, true hypotheses provide a higher degree of systematic power than false hypotheses. In addition, we see that measures of systematic power that satisfy the minimal requirements 1–3 allow us to further distinguish between true hypotheses. In particular, if one compares two hypotheses, both of which are true but where one is logically stronger, then after receiving finitely many pieces of evidence (in a sequence of separating pieces of evidence) and for every piece of evidence thereafter, the logically stronger hypothesis displays a higher degree of systematic power with respect to the evidence than the logically weaker hypothesis. This demonstrates that systematic power is a very powerful criterion of theory choice. After finitely many pieces of evidence, and for every piece of evidence thereafter, comparisons of systematic power reflect our preferences in theory choice. In particular, systematic power reflects the two requirements on theory choice that Hempel (1960), Huber (2008), and Levi (1967) agree upon[8]:

Requirement 4 (preference in theory choice) For all hypothesis H_1 and H_2, H_1 is to be preferred to H_2 if: (i) H_1 is true and H_2 is false, or (ii) H_1 and H_2 are true, but H_1 is logically stronger than H_2.

Accordingly, we can conclude that systematic power is indeed a very promising criterion of theory choice. More importantly, even though the debate concerning which probabilistic measure is best suited for measuring explanatory and systematic power is still ongoing, it is a simple corollary of the preceding considerations (Theorems 1 and 2) that the previously suggested measures of systematic power sp^1–sp^3 are useful for guiding theory choice. After receiving finitely many pieces of evidence (in a sequence of separating observational statements), and for every piece of evidence thereafter, true hypotheses display a higher degree of systematic power than false hypotheses; and if one compares two hypotheses, both of which are true but where one is logically stronger, then after receiving finitely many pieces of evidence (in a sequence of separating pieces of evidence) and for every piece of evidence thereafter, the logically stronger hypothesis displays a higher degree of systematic power with respect to the evidence than the logically weaker hypothesis.

Corollary 1 (truth-conduciveness of systematic power) *Let W be a set of possible worlds and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ which separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let* Pr *be a strict (or regular) probability function on $\mathcal{A}$. Let* Pr* *be the unique probability function on the smallest*

[8] Hempel (1960), and Levi (1967) and Huber (2008) disagree in their evaluation of false hypotheses. Where Hempel prefers logically weaker false hypotheses to logically stronger false hypotheses, Levi and Huber prefer logically stronger false hypotheses to logically weaker false hypotheses. For discussion, see Brössel (2014).

σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ *for all $A \in \mathcal{A}$. Then there is a $W' \subseteq W$ with* $\Pr^*(W') = 1$ *such that the following holds for every $w \in W'$ and all hypotheses $H_1, H_2 \in \mathcal{A}$.*

1. *if $w \vDash H_1$ and $w \vDash \neg H_2$, then:*
   $\exists n \forall m \geq n : [\mathfrak{sp}_{\Pr}(H_1, E_m^w) > \mathfrak{sp}_{\Pr}(H_2, E_m^w)]$, *if* $\mathfrak{sp}_{\Pr} \in \{sp^1, sp^2, sp^3\}$.
2. *if $w \vDash H_1 \cap H_2$ and $H_1 \vDash H_2$* but *$H_2 \nvDash H_1$, then:*
   $\exists n \forall m \geq n : [\mathfrak{sp}_{\Pr}(H_1, E_m^w) > \mathfrak{sp}_{\Pr}(H_2, E_m^w)]$, *if* $\mathfrak{sp}_{\Pr} \in \{sp^1, sp^2, sp^3\}$.

where $E_m^w = \bigcap_{0 \leq i \leq m} e_i^w$.
(This corollary is a simple consequence of the Theorems 1 and 2. The proof is omitted here.)

3.2 Inference to the best systematization

Can we provide a similar vindication for the inference from the premise that a given hypothesis is the best systematization for the evidence to the conclusion that the given hypothesis is acceptable or true? More specifically, can we derive a vindication of the inference schema IBS from the above vindication of systematic power as a criterion of theory choice?

Inference to The Best Systematization (IBS)

$$\frac{\begin{array}{c} E \\ H \text{ is the best systematization for } E \text{ with respect to the set of available hypotheses } \{H_1, \ldots, H_n\} \end{array}}{\text{Therefore: } H_i}$$

In order to apply this inference schema at some arbitrary time point t_0, it is required that the hypothesis H_i is the best systematization of the evidence E_{t_0} available at time point t_0 (in the sense of Definition 9, 10, or 11) if compared with the hypotheses in the set $\{H_1, \ldots, H_n\}$ under consideration. If the probabilities underlying the definition of systematic power are interpreted as subjective probabilities, then it is trivial for the agent to determine which hypothesis is the best systematization for the evidence E_{t_0} available at time point t_0. Thus, in that case we do not need an additional epistemology for how agents learn which hypothesis is the best systematization for the evidence E_{t_0} available at time point t_0. However, if the probabilities underlying the definition of systematic power are interpreted as objective chances, then the we need additionally an epistemology for how agents learn about these objective chances in order to determine which hypothesis is the best systematization for the evidence E_{t_0} available at time point t_0. In the following I presuppose that the agent is in the position to determine the systematic power of a hypothesis with respect to the available evidence (thus I ignore the epistemology of objective chances, as I ignore the topic of rule-following concerning inference rules in general). One cannot apply the inference schema at time point t_0, if there is no hypothesis H_i that is the best systematization of the evidence E_{t_0} with respect to the set of hypotheses $\{H_1, \ldots, H_n\}$ under consideration (e.g., if two or more hypotheses fare equally well in systematizing the evidence). In order to apply IBS, it is not required that at time point t_0 the agent already knows that H_i provides

the best systematization for all future evidence E_{t_1} where t_1 is later than t_0. It is only required that the agent "knows" or is justified in believing that the hypothesis H_i is the best systematization of the evidence E_{t_0} presently available to the agent at time point t_0. Thus, it might be the case that after obtaining more evidence between t_0 and t_1 the agent infers at t_1 another hypothesis with the help of the inference schema IBS because the second hypothesis displays higher systematic power with respect to the evidence E_{t_1} available at t_1.

In the light of this understanding of IBS it follows from the considerations in the preceding subsection that we can provide at least a partial vindication for IBS, given certain assumptions. (i) If there is at least one true hypothesis in the set of available hypotheses $\{H_1, \ldots, H_n\}$, then after finitely many pieces of evidence and for every piece of evidence thereafter, IBS will allow us to infer a true hypothesis. (ii) If there is a logically strongest true hypothesis in the set of available hypotheses $\{H_1, \ldots, H_n\}$, then after finitely many pieces of evidence, and for every piece of evidence thereafter, IBS will allow us to infer the logically strongest hypothesis. Of course, these results are again subject to the same conditions as the convergence results. More formally:

Theorem 3 (A (partial) vindication of IBS) *Let W be a set of possible worlds and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ which separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let* Pr *be a strict (or regular) probability function on $\mathcal{A}$. Let* Pr* *be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying* $\mathrm{Pr}^*(A) = \mathrm{Pr}(A)$ *for all $A \in \mathcal{A}$. Then there is a $W' \subseteq W$ with* $\mathrm{Pr}^*(W') = 1$ *so that the following holds for every $w \in W'$ and all hypotheses $H_1, \ldots, H_n \in \mathcal{A}$ and for all $\mathfrak{sp}_{\mathrm{Pr}}$ satisfying Requirements* 1–3.

1. *If there is a $H_j \in \{H_1, \ldots, H_n\}$ such that $w \vDash H_j$, then:*
   $\exists n \forall m \geq n$: *if H_i is the best systematization for E_m^w with respect to the set of hypotheses $\{H_1, \ldots, H_n\}$, then $w \vDash H_i$.*
2. *If there is a $H_j \in \{H_1, \ldots, H_n\}$ such that $w \vDash H_j$ and $H_j \vDash H_i$ for all H_i with $w \vDash H_i$, then:*
   $\exists n \forall m \geq n$: *H_j is the best systematization for E_m^w with respect to the set of hypotheses $\{H_1, \ldots, H_n\}$.*

where $E_m^w = \bigcap_{0 \leq i \leq m} e_i^w$.
(The proof for this theorem can be found in the appendix.)

From a formal perspective, both assumptions—(i) that there is at least one true hypothesis among the hypotheses under consideration and (ii) that there is a logically strongest true hypothesis among the hypotheses under consideration—are weak. If the set of available hypotheses contains a "catchall" hypothesis, i.e., the negation of the disjunction of all other hypotheses in the set, it is guaranteed that there is a true hypothesis among the hypotheses under consideration. If the set of available hypotheses is a finite sub-algebra of $\mathcal{A}$, it is also guaranteed that there is a logically strongest true hypothesis among the hypotheses under consideration. It should be noted that it is no problem to close a finite set of hypotheses under negation and conjunction to form a finite sub-algebra of $\mathcal{A}$. That both assumptions are weak from a formal perspective

indicates that it does not take much effort to ensure that these assumptions are actually true. Indeed, given a very permissive sense of 'under consideration' it is extremely plausible that both assumptions are actually satisfied, after all the language of science allows for an unrestricted use of conjunctions and negations. In particular, one could argue that especially the negation of a hypothesis is taken into consideration whenever the hypothesis is under consideration. To not consider $\neg H$ if one considers H is like not considering the possibility that one's hypothesis is false. Scientists usually consider the hypothesis that their initial hypothesis is false. One could also argue that whenever two hypothesis H_1 and H_2 are under consideration, scientists also consider the possibility that both of them are true, at least if they are logically consistent with each other. Nevertheless, even if both assumptions are satisfied, one has to admit that the inference schema IBS as portrayed above would be subject to van Fraassen's (1989) famous *Best of a Bad Lot Objection* if the latter objection had not been rebutted by Schupbach (2014). As Schupbach convincingly argues with respect to IBE:

> [T]he bad lot objection is powerless against the inference form of IBE (it is powerless against other inference forms too); but in that case, the objection provides no motivation for revamping the form of explanatory inference. On the other hand, the bad lot objection is more compelling when framed as a problem for particular inferences to the best explanation (e.g., those used by realists); but in this case, the bad lot objection is not an objection to IBE, but rather an objection to the material content involved in particular instances of IBE. In neither case does the bad lot objection call for us to discard IBE and replace it with a more modest formulation of explanatory inference. (Schupbach 2014, p. 63)

Following Schupbach, the *Best of a Bad Lot Objection* against IBS might be compelling against a particular inference to the best systematization, but in that case it is rather an objection against the set of hypotheses considered, not against the inference schema IBS itself. For example, one might argue that since in a particular instance of IBS the uninformative catch-all hypothesis came out as the best systematization with respect to the evidence, we should consider additional hypotheses. Or one might argue that the finite sub-algebra of $\mathcal{A}$ considered in a particular instance of IBS is too coarse-grained, and thus the achieved conclusion too uninformative. Instead, one should consider a more fine-grained sub-algebra of $\mathcal{A}$. However, such objections would only meet the set of hypotheses under consideration and not the inference schema IBS itself.

Interestingly, given the underlying notion of systematic power, scientists are also in a position to judge whether the set of hypotheses considered contains only bad hypotheses, respectively whether the choice of considered hypotheses is adequate or appropriate for the scientist's purposes at hand. If for an extended number of observations none of the hypotheses under consideration has a high systematic power with respect to the evidence, then it is perhaps time to consider additional hypotheses or more fine-grained algebras. Such judgments, should not be understood as some sort of meta-inductive rule. In particular, we are not proposing to infer something about the truth or falsity of a hypothesis from the fact that none of the hypotheses under consideration displays a high systematic power. On the contrary, for a given set of

hypotheses one should employ IBS and accept the hypothesis that displays the highest systematic power. However, when it comes to the particular choice of the algebra or the set of hypotheses under consideration it does not make sense to say that such a decision is true or false. Instead the choice of the algebra or the set of hypotheses is useful or not, and if none of the hypotheses displays a high systematic power with respect to the evidence then scientists should ask whether their particular choice serves their scientific purposes. In that sense the choice of the algebra or the set of hypotheses under consideration is similar to Carnap's problem of the choice of appropriate linguistic frameworks. With respect to the question of the rationality of such changes Carnap (1950) says the following:

> After the new forms are introduced into the language, it is possible to formulate with their help internal questions and possible answers to them. A question of this kind may be either empirical or logical; accordingly a true answer is either factually true or analytic. From the internal questions we must clearly distinguish external questions, i.e., philosophical questions concerning the existence or reality of the total system of the new entities.[…] To be sure, we have to face at this point an important question; but it is a practical, not a theoretical question; it is the question of whether or not to accept the new linguistic forms. The acceptance cannot be judged as being either true or false because it is not an assertion. It can only be judged as being more or less expedient, fruitful, conducive to the aim for which the language is intended. (Carnap 1950, Sect. 3)

Friedman was one of the first to clearly recognize the importance of these observations for the testing of scientific theories in actual science.

> Just as, for Carnap, the logical rules of a linguistic framework are constitutive of the notion of "correctness" or "validity" relative to this framework, so a particular paradigm governing a given episode of normal science, for Kuhn, yields generally-agreed-upon (although perhaps only tacit) rules constitutive of what counts as a "valid" or "correct" solution to a problem within this episode of normal science. Just as, for Carnap, external questions concerning which linguistic framework to adopt are not similarly governed by logical rules, but rather require a much less definite appeal to conventional and/or pragmatic considerations, so changes of paradigm in revolutionary science, for Kuhn, do not proceed in accordance with generally-agreed-upon rules as in normal science, but rather require something more akin to a conversion experience. (Friedman 2002, p. 181)

We suggest the same holds true for considerations of the systematic power of the hypotheses in a chosen algebra. The decision for a certain algebra or for a specific set of hypotheses includes certain "conventional and/or pragmatical considerations" since they depend crucially on the linguistic framework adopted by the scientist and additional pragmatic constraints. Thus, following Carnap and Friedman the acceptance of a certain algebra or a specific set of hypotheses cannot be taken to be either true or false but can "be judged as being more or less expedient, fruitful, conducive to the aim for which the language is intended" (Carnap 1950, Sect. 3) and within the

boundaries of the scientist's pragmatic constraints. Based on the past development of the systematic power of the available hypotheses, scientists can come to a decision on whether the set of available hypotheses contains only unfruitful or inexpedient hypothesis for the scientist's purposes at hand. More specifically, if for an extended number of observations none of the hypotheses under consideration displays a high systematic power with respect to the evidence (even if there is, for example, a catchall hypothesis that displays a slightly higher systematic power than the other hypotheses), then it is perhaps time to rethink one's linguistic framework or one's pragmatically motivated limitations and consider additional hypotheses or more fine-grained sub-algebras. Thus, measures of systematic power can be considered a valuable tool that informs but does not determine our conventional and/or pragmatical considerations. However, this should not be understood as meta-inductive rule which we intend to apply to infer something about the truth or falsity of the set of hypotheses under consideration, but only about the usefulness of the given convention concerning what hypotheses are considered to be relevant. Thus, considerations of systematic power can be an important aid in coming to a decision concerning questions that resemble Carnap's external questions.

More importantly, from the preceding considerations we can conclude that IBS is indeed a very fruitful inference schema in science. In particular, *within* the given algebra, or respectively for the given set of hypotheses under consideration, IBS leads us after finitely many steps of observation and for every observation thereafter to accept the logically strongest true hypothesis. Thus, IBS is the appropriate tool for answering what Carnap calls the internal questions of science (at least for those Bayesians who consider the convergence theorems of Bayesian epistemology to confer at least partial vindication on the Bayesian norms of reasoning).

3.3 Inferences to the best explanation and inferences to the best systematization in philosophy

Interestingly, the inference schema IBS and the related inference schema IBE are not only used in science, but also in philosophy. As mentioned earlier, a classic example from the philosophy of science is the debate surrounding scientific realism. Proponents of scientific realism argue that we are justified in believing that the theoretical terms of our best scientific hypotheses refer to objects in the world, and defend the view that the hypothesis that these theoretical terms refer to objects in the world is the best explanation for the apparent success of science. The existence of successful theories whose theoretical terms did not refer to objects in the world would be a miracle. A classic example from epistemology is the debate on external world skepticism. Here, the opponents of external world skepticism argue that this hypothesis must be false, since the existence of external world objects is the best explanation for our perceptions. Given the structure and predictability of many of our perceptions, it would be a miracle if they were not caused by objects in the external world. Given these so-called inferences to the best explanation, the question is whether the Theorems 1–3 show that the inference schemas IBE and IBS are also fruitful within philosophy.

Before discussing whether IBS or IBE are also useful within philosophy, it is important to note that the results presented above are not decisive, as regards the debate about scientific realism. In particular, in order to prove the above results, one assumption is required that makes the results achieved irrelevant for the debate about scientific realism. This assumption is that the sequence of pieces of evidence $e_0, \ldots, e_n, \ldots$ must separate the set W, where a sequence of pieces of evidence separates the set of possibilities W if and only if for every pair of worlds w_i and $w_j \in W$ (with $w_i \neq w_j$) there is one piece of evidence in the sequence such that it is true in one of the possibilities and false in the other. The assumption implies that the Theorems 1–3 do not speak about scientific hypotheses that contain theoretical vocabulary, i.e., vocabulary which includes non-observational terms. More specifically, if we allow for hypotheses that contain theoretical vocabulary, then we allow for two possible worlds w_i and $w_j \in W$ (with $w_i \neq w_j$), such that there is no piece of evidence in the sequence (that we can obtain via observation) that is true in one of the worlds and false in the other. Thus, these theorems do not prove that successful hypotheses—i.e., hypotheses that display a strong systematic power after finitely many pieces of evidence and for every piece of evidence thereafter—that contain theoretical vocabulary are true, and therefore they are not decisive with respect to the debate about scientific realism. For discussion, see Brössel (2014).

In a second step, one can also recognize that Theorems 1–3 do not settle the general question whether IBS and IBE are fruitful inference schemas within philosophy. In particular, the assumption that there is a sequence of pieces of evidence $e_0, \ldots, e_n, \ldots$ that separates the set W becomes implausible if we assume that subsets of W are philosophical hypotheses. Even if we assume that intuitions can serve as philosophical evidence, and even if we additionally assume that these intuitions are fully reliable, it is very implausible that such philosophical evidence can separate the set of possibilities W. This would require that for every pair of philosophically conceivable possible worlds, the following holds: there is some proposition e that is true in only one of these two worlds and we intuit e in the world in which e is true and we intuit $\neg e$ in the world in which e is false. Furthermore, we cannot think of any other conception of philosophical evidence that would make the satisfaction of this requirement plausible. Thus, Theorems 1–3 do not settle the question whether IBS and IBE are fruitful inference schemas within philosophy.

4 From systematic power to Bayes' rule

Sections 3.1 and 3.2 demonstrated that IBS is a fruitful inference schema in science (but not in philosophy, as discussed in Sect. 3.3). It leads us to accept the logically strongest, true hypothesis within a set of available hypotheses. Here, it is important to note that the neutral term 'acceptance' is used instead of the term 'full belief'. Clearly, for most philosophers, accepting a scientific hypothesis means fully believing that the hypothesis is true. The above result can then be taken to show that fully believing a hypothesis which is the best systematization of the evidence is justified: after finitely many pieces of evidence, and for every piece of evidence thereafter, the epistemic agent believes the logically strongest true hypothesis available. From an

epistemological point of view, there is nothing wrong with believing such a hypothesis. Some philosophers reject the idea that accepting a hypothesis requires fully believing that it is true. For example, according to scientific anti-realists like van Fraassen (1980), accepting a hypothesis only requires fully believing that the hypothesis is empirically adequate. Thus, using the neutral term 'acceptance' will allow scientific anti-realists of this kind to employ IBS for scientific reasoning. For further discussion see also Brössel (2014).

However, there are also philosophers who completely reject the idea that we are allowed to believe or accept a given hypothesis. According to this view, all inductive inference rules that make you believe or accept a hypothesis are unacceptable. In particular, radical Bayesians like Jeffrey (1956, 1992) reject the notions of full belief or acceptance. They would argue that even though IBS is based on some probabilistic measure of systematic power, IBS is strictly speaking not Bayesian at heart. IBS makes agents believe or accept a hypothesis or reject it for another, where "[t]he framework of probabilism replaces the two Cartesian options, affirmation and denial, by a continuum of judgmental probabilities in the interval from 0 to 1, endpoints included" (Jeffrey 1992, p. 194). For these reasons in particular, radical Bayesians argue that, like the notions of accepting or fully believing a hypothesis, IBS itself has to be rejected.

Suppose we follow radical Bayesians in rejecting the notions of full belief and acceptance. Then we have to reject inference schemas such as IBE and IBS. More generally, it seems that considerations of explanatory and systematic power are "in conflict with other forms of change of opinion, that we accept as rational" as noted by van Fraassen (1989). In particular, such considerations seem to be in conflict with *Bayes' Rule*.

Definition 12 (*Bayes' rule*) If $\Pr_{t_0}$ is the agent's probability function at time point t_0, E is the logically strongest proposition that the agent became absolutely certain of between time points t_0 and t_1, and $\Pr_{t_0}(E) > 0$, the agent's probability function should change to $\Pr_{t_1}$, which is defined as follows:

$$\Pr_{t_1}(H) = \Pr_{t_0}(H|E)$$

for all $H \in \mathcal{A}$.

On the surface, it seems that neither the explanatory power nor the systematic power of the hypothesis H with respect to the evidence E has any influence on the posterior degree of belief in H at time point t_1. However, closer considerations reveal that the systematic power of the hypothesis with respect to the evidence E has a tremendous effect on the posterior degree of belief in H at time point t_1. The so-called Bayes' Theorem is the following:

Theorem 4 (Bayes' Theorem)

$$\Pr_{t_0}(H|E) = \frac{\Pr_{t_0}(E|H) \times \Pr_{t_0}(H)}{\Pr_{t_0}(E)}$$

(The proof is omitted here.)

A trivial reformulation of Bayes' Theorem shows what we want to prove, i.e., that the systematic power of the hypothesis with respect to the evidence has a tremendous effect on the new degree of belief in H at time point t_1. The trivial reformulation of Bayes' Theorem is this:

Theorem 5

$$\Pr_{t_1}(H) = \Pr_{t_0}(H|E) = sp^1_{\Pr_{t_0}}(H, E) \times \Pr_{t_0}(H)$$

(The proof is omitted here.)

Similar results can be obtained for the other two measures of systematic power. The results are the following:

Theorem 6

$$\begin{aligned}\Pr_{t_1}(H) &= \Pr_{t_0}(H|E)\\ &= \frac{\dfrac{e^{\tanh^{-1}[sp^2_{\Pr_{t_0}}(H,E)]}\Pr_{t_0}(E)}{e^{\tanh^{-1}[sp^2_{\Pr_{t_0}}(H,E)]}\Pr_{t_0}(E) + e^{\tanh^{-1} sp^2_{\Pr_{t_0}}(H,\neg E)}\Pr_{t_0}(\neg E)}}{\Pr_{t_0}(E)} \times \Pr_{t_0}(H)\end{aligned}$$

(The proof can be found in the appendix.)

Theorem 7

$$\begin{aligned}\Pr_{t_1}(H) &= \Pr_{t_0}(H|E)\\ &= \begin{cases} \dfrac{[\Pr_{t_0}(\neg E) \times sp^3_{\Pr_{t_0}}(H, E) + \Pr_{t_0}(E)] \times \Pr_{t_0}(H)}{\Pr_{t_0}(E)} & if\ \Pr_{t_0}(E|H) \geq \Pr_{t_0}(E)\\ \dfrac{[\Pr_{t_0}(E) \times sp^3_{\Pr_{t_0}}(H, E) + \Pr_{t_0}(E)] \times \Pr_{t_0}(H)}{\Pr_{t_0}(E)} & if\ \Pr_{t_0}(E|H) < \Pr_{t_0}(E)\end{cases}\end{aligned}$$

(The proof is omitted here.)

Theorems 5–7 show that an agent's degree of belief in some hypothesis depends on her a priori degree of belief in that hypothesis and the systematic power of that hypothesis with respect to the evidence. In the case of the measures of systematic power sp^2 and sp^3 this depends additionally on the agent's a priori degree of belief in the evidence. These theorems therefore demonstrate that considerations of systematic power are not in conflict with Bayes' Rule. Rather, such considerations can be taken to inform the agent's posterior degree of belief in the hypothesis in the light of the evidence E. Thus, even radical Bayesians, who reject all inference rules that allow one to accept or fully believe a hypothesis, must admit that systematic power is an integral component of Bayesian reasoning. The agent's posterior degree of belief in the hypothesis depends crucially on the hypothesis's systematic power with respect to the available evidence and, *ceteris paribus*, the higher the systematic power of the hypothesis, the higher the agent's posterior degree of belief in it.

5 Putting the results in perspective: van Fraassen and the role of explanatory and systematic power in scientific reasoning

It should be obvious that it is not the aim of this paper to give a comprehensive, detailed, and historical account of van Frassen's various arguments against IBE. For this purpose other papers are more pertinent: Okasha (2000) and Psillos (1996) must be mentioned here. Naturally, this is also not the place to discuss comprehensively and in detail in how far the results achieved here can be considered a rebuttal of van Fraassen's "attack" against IBE or its close cousin IBS. The latter task presupposes the former. In addition, it would require intensive speculation about how van Fraassen would formulate his criticism against IBS and systematic power. (Furthermore, it might be even more interesting to investigate in more detail the relations and differences between the philosophical conclusions of Okasha (2000) and Weisberg (2009) and the results presented here.) Nevertheless, because of van Fraassen's importance for the debate a few comments are in order (and these comments presuppose that van Fraassen would raise any objection raised already against IBE and explanatory power, against IBS and systematic power as well).

According to Okasha (2000, p. 692), "[t]o understand [van Fraassen's] attack, it is necessary to look briefly at van Fraassen's views on induction." So let's do this! We begin with van Fraassen's general picture of induction and rational reasoning.

> [V]an Fraassen is no inductive sceptic; he grants the rationality of our beliefs about the unobserved. What enables van Fraassen to reject the traditional ideal of induction without falling into inductive scepticism is a particular thesis about rationality. Rationality is a concept of permission, not obligation, he maintains: it concerns what you may believe, not what you must. Therefore, rational belief change need not be governed by rules which tell you how to respond to evidence; two agents can respond very differently to the same evidence, without one of them being irrational. (Okasha 2000, p. 693)

Thus, the radical Bayesian's criticism against IBE or IBS (see Sect. 4) is not backed by van Fraassen. Nothing in van Fraassen's general picture of induction forbids us to believe or accept a hypothesis. In the light of this very general view of induction and rational reasoning the crucial point for van Fraassen seems to be this: there are no arguments that support the view that one *ought* to apply IBE, its close cousin IBS, or any other inference rule. Since the paper neither proves nor claims that IBE or IBS are the only inference rules that lead one to accept the logically strongest, true hypothesis from among the available hypotheses, nothing in the present paper stands in contradiction with this very general picture of inductive inference. In fact, there are many inference rules that lead one to accept the logically strongest, true hypothesis of all available hypothesis (see Huber 2008; Brössel 2014), and thus we must conclude that these inference rules are just as justified as IBS. Thus, the paper establishes that one is permitted to apply IBS and IBE, but certainly not that one ought to apply them. However, van Fraassen (1989, p. 142) clearly thinks that his arguments against IBE and IBS are stronger, when he writes that there are many charges to be laid against IBE and IBS: "One is that it pretends to be something other than it is. Another is that

it is supported by bad arguments. A third is that it conflicts with other forms of change of opinion, that we accept as rational." Thus, it is fair to interpret van Fraassen as saying that IBE and, presumably, its close cousin IBS, are inference rules that you are not even permitted to apply. In contrast to this, the paper argues that IBS is an inference rule that one is permitted to apply in scientific inquiry. The basic premises in support of this conclusion are: (i) Schupbach's successful mitigation of the force of the *Best of a Bad Lot Objection* against IBE, IBS, or other variants of inductive inference rules (together with the fact the measures of systematic power are good indicators of whether the available hypotheses are a *Bad Lot*) and (ii) the proof that after finitely many steps of observation and for every observation thereafter IBS leads you to accept the logically strongest, true hypothesis available. The only important premise that one has to accept to obtain the proof is this: explanatory and systematic power and hence IBE and IBS can be defined in terms of Bayesian probabilities.

Van Fraassen nevertheless thinks that in the light of his counterarguments advocates of IBE and IBS at least have to retrench.

> This retrenchment can take two forms. The first form is that the special features which make for explanation among empirically unrefuted theories, make them (more) likely to be true. The second form is that the notion of rationality itself requires these features to function as relevant factors in the rules for rational response to the evidence. […] Let us note beforehand that the first must lean on *intrinsic explanatoriness*, which can be discerned prior to empirical observations, and the second specifically on *explanatory success* after the observational results come in. What the criteria are for either, we shall leave up to the retrencher. (1989, p. 146)

Thus, according to van Fraassen (1989, p. 146), it is not even possible to design a retrenched IBE rule, which specifies how we should "allocate our personal probabilities with due respect to explanation." This indicates that van Fraassen thinks that there is no room in Bayesian norms of reasoning for considerations of explanatory power. Before going on we should take a look at what van Fraassen thinks about Bayesian norms of reasoning. According to Okasha:

> Van Fraassen […] is nonetheless a Bayesian of sorts. He accepts the Bayesian representation of opinion in terms of degrees-of-belief, and he agrees that synchronic probabilistic coherence is a necessary condition of rationality. However, he does not accept the Bayesian thesis that conditionalization is the only rational way to respond to new evidence; though he allows that it is *a* rational way. (Okasha 2000, p. 693)

Thus, according to van Fraassen one might not be obliged to use strict conditionalization for changing one's degrees of belief in the light of evidence, but one is certainly permitted to do so. Theorems 5–7 show that if one is permitted to use strict conditionalization, then one is also permitted to update one's degrees of belief by a rule which allocates the agent's degrees of belief with due respect to considerations of explanatory, predictive, and retrodictive power, i.e., the hypothesis's systematic power. In particular, these theorems show that one can design update rules with the following two

important properties: (i) according to these update rules, an agent's updated degrees of belief in a hypothesis are a strictly increasing function (solely) from the agent's old degrees of belief in that hypothesis and its systematic power with respect to the evidence (combining its explanatory, predictive, and retrodictive power), and (ii) these update rules are logically equivalent to strict conditionalization. Thus, such update rules can be plausibly understood as retrenched inferences to the best systematization and one is permitted to apply them since they are logically equivalent to an update rule that van Fraassen explicitly permits in his (1989) *Laws of Symmetry.* Thus, as long as the relevant probabilistic measures can be understood as measures of a hypothesis's systematic power with regard to the evidence, van Fraassen is not only wrong in rejecting the inference rule IBS, he is also wrong in rejecting the retrenched versions of IBS that specify how systematic power influences posterior beliefs in hypotheses. Again the only important premise that one has to accept is this: explanatory and systematic power and hence IBE and IBS as well can be defined in terms of Bayesian probabilities.

It is presumably this last premise that van Fraassen either rejects or does not consider. Van Fraassen considers only two possibilities regarding how to assign a role to considerations of explanatory (and systematic) power within Bayesian epistemology. According to the first, hypotheses that display a high explanatory (or systematic) power with regard to the evidence are assigned a higher prior probability than hypotheses that display less explanatory (or systematic) power. Since this would require that we know a priori which hypotheses are more explanatory with regard to our future a posteriori evidence, this approach is not tenable. Here we agree with van Fraassen. For a more detailed and critical discussion of van Fraassen's argument against this first variant see Okasha (2000, Sect. 4) and Douven (2011). According to the second, we assign a higher a posteriori probability to explanatory hypotheses when they display *explanatory success* after the observational results come in. "Combining the ideas of personal probability and living by rules, the new rule of IBE would be a recipe for adjusting our personal probabilities while respecting the explanatory (as well as predictive) success of hypotheses" (van Fraassen 1989, p. 149). Van Fraassen claims that it is impossible to construct such inference rules and maintain that they are rational. Van Fraassen's argument presupposes that such rules require one to adjust one's probabilities in such a way that one assigns extra probability to hypotheses that not only show a certain amount of predictive success but that are also explanatory. Van Fraassen's invites us to imagine a conversation between a Preacher or Explanationism and a perfect Bayesian, he calls him Peter:

> [T]he Preacher goes on to say: in view of this explanatory success, you should raise your credence in the more explanatory hypotheses.
>
> 'What?' exclaims Peter. 'More than I would anyway?' 'Yes', says the Preacher. 'Our forefathers all inferred to the best explanation, and in daily commerce, our humbler brothers still do. We who have seen the light of probability should not disdain their insight, but give due respect to explanatory success.' (van Fraassen 1989, p. 166)

Thus, according to van Fraassen, the retrenched rules of IBE should require the following: (i) an agent's updated degrees of belief in a hypothesis are a strictly increasing

function from the agent's old degrees of belief in that hypothesis, its predictive or retrodictive power with respect to the evidence, and, in addition, some additional bonus for explanatory success, in case the hypothesis explains the evidence; and (ii) these update rules are logically equivalent to strict conditionalization if none of the hypotheses explains the evidence, and they assign a higher degree of belief to a hypothesis than strict conditionalization does, if the hypothesis does explain the evidence (and thus a lower degree of belief if the hypothesis does not explain the evidence whereas one of the alternative hypotheses explains it).

In contrast to van Fraassen, but in the spirit of Hempel and Oppenheim (1948) and Hempel (1958) we presuppose a symmetry principle for measures of predictive, retrodictive and explanatory power. In particular, we argued that if we can use a probabilistic measure as a measure of the predictive power of the hypothesis with regard to the evidence, then we can use the same measure to gauge the explanatory power of a hypothesis with regard to the evidence if the hypothesis explains the evidence, and we can use it to gauge the retrodictive power of a hypothesis with regard to the evidence if the hypothesis retrodicts the evidence. Thus, when Bayesian Peter asks whether he should raise his degree of belief in the hypothesis displaying explanatory success more that he would anyway, the preacher of IBE and IBS should have said: 'No, our forefathers all inferred to the best systematization, discounting any difference between prediction, retrodiction, and explanation, and in daily commerce, our humbler brothers still do. We who have seen the light of Hempel's and Oppenheim's symmetry principle should not disdain their insight and give equal respect to explanatory, predictive and retrodictive success.'

6 Summary

For decades, the role of the inference schema *Inference to the Best Explanation* within scientific reasoning has been hotly debated, and for decades philosophers have suggested probabilistic measures of explanatory power. The present paper has shown that both inference schemas—*Inference to the Best Explanation* and *Inference to the Best Systematization*—can play a central role within scientific reasoning. Accepting the hypothesis with the highest systematic power leads one to accept the logically strongest true hypothesis among the available hypotheses, after finitely many pieces of evidence and for every piece of evidence thereafter. Furthermore, even if we adopt radical Bayesianism, and reject all qualitative notions of belief or acceptance and admit only of degrees of belief, we can show that explanatory and systematic power play an important role in scientific reasoning. In particular, the agent's posterior degree of belief in the hypothesis depends crucially on the hypothesis's systematic power with respect to the available evidence and, *ceteris paribus*, the higher the systematic power of the hypothesis, the higher the agent's posterior degree of belief in it. The crucial assumptions for this are: (i) explanatory and systematic power, and IBE and IBS, can be defined in terms of Bayesian probabilities; and (ii) a symmetry principle for measures of predictive, retrodictive and explanatory power so that we can combine these notions in one measure of systematic power. Certainly not all defenders of IBE and related inference rules will be happy with this presupposition. For example,

adopting the picture advocated here and with it the vindication of the application of these measures and inference rules in induction might feel like a Pyrrhic victory for Okasha (2000). Van Fraassen, on the other hand, a "Bayesian of Sorts", might actually be happy with the results achieved here. As long as explanatory and systematic power, and IBE and IBS, are constrained by Bayesianism, he might rule that we are permitted to use these inference rules.

Acknowledgments I am indebted to Ralf Busse, Vincenzo Crupi, Markus Eronen, Branden Fitelson (for making me aware of the Harman (1967) paper), Albert Newen, Gerhard Schurz (and the members of his research colloquium), and especially Matteo Colombo, Anna-Maria A. Eder, Jan Sprenger, and Ben Young. Finally, I am also grateful to two (very challenging) referees of this journal.

Appendix

Proof of Theorem 1

We have to show that ep^1–ep^3 satisfy requirements 1–3.

Proof for ep^1

1. ep^1 satisfies Requirement 1 trivially. It is defined in terms of probabilities.
2. ep^1 satisfies Requirement 2 with marker 1:

$$ep^1_{\mathrm{Pr}}(H, E) = \frac{\Pr(E|H)}{\Pr(E)} = \begin{cases} > 1, & \Pr(E|H) > \Pr(E) \\ = 1, & \Pr(E|H) = \Pr(E) \\ < 1, & \Pr(E|H) < \Pr(E) \end{cases}$$

3. ep^1 satisfies Requirement 3: if $\Pr(E|H_1) > \Pr(E|H_2)$, then

$$ep^1_{\mathrm{Pr}}(H_1, E) = \frac{\Pr(E|H_1)}{\Pr(E)} > \frac{\Pr(E|H_2)}{\Pr(E)} = ep^1_{\mathrm{Pr}}(H_2, E).$$

Proof for ep^2

1. ep^2 satisfies Requirement 1 trivially. It is defined in terms of probabilities.
2. ep^2 satisfies Requirement 2 with marker 0: First note that

$$ep^2_{\mathrm{Pr}}(H, E) = \frac{\Pr(H|E) - \Pr(H|\neg E)}{\Pr(H|E) + \Pr(H|\neg E)} = \begin{cases} > 0, & \Pr(H|E) > \Pr(H|\neg E) \\ = 0, & \Pr(H|E) = \Pr(H|\neg E) \\ < 0, & \Pr(H|E) < \Pr(H|\neg E) \end{cases}$$

Now we only have to see that

$$\Pr(H|E) \overset{>}{\underset{<}{=}} \Pr(H|\neg E) \Leftrightarrow \Pr(H|E) \overset{>}{\underset{<}{=}} \Pr(H) \Leftrightarrow \Pr(E|H) \overset{>}{\underset{<}{=}} \Pr(E)$$

3. ep^2 satisfies Requirement 3: if $\Pr(E|H_1) > \Pr(E|H_2)$, then
   (a) $\frac{\Pr(E|H_1)}{\Pr(E)} > \frac{\Pr(E|H_2)}{\Pr(E)}$
   (b) $\Pr(\neg E|H_1) < \Pr(\neg E|H_2)$ and, thus, also: $\frac{\Pr(\neg E|H_1)}{\Pr(\neg E)} < \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}$

(a) and (b) imply that:

$$\left[\frac{\Pr(E|H_1)}{\Pr(E)} \times \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right] - \left[\frac{\Pr(\neg E|H_1)}{\Pr(\neg E)} \times \frac{\Pr(E|H_2)}{\Pr(E)}\right] > \left[\frac{\Pr(E|H_2)}{\Pr(E)} \times \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)}\right] - \left[\frac{\Pr(\neg E|H_2)}{\Pr(\neg E)} \times \frac{\Pr(E|H_1)}{\Pr(E)}\right]$$

and that therefore:

$$\left[\frac{\Pr(E|H_1)}{\Pr(E)} \times \frac{\Pr(E|H_2)}{\Pr(E)} + \frac{\Pr(E|H_1)}{\Pr(E)} \times \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right] - \left[\frac{\Pr(\neg E|H_1)}{\Pr(\neg E)} \times \frac{\Pr(E|H_2)}{\Pr(E)} + \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)} \times \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right] > \left[\frac{\Pr(E|H_1)}{\Pr(E)} \times \frac{\Pr(E|H_2)}{\Pr(E)} + \frac{\Pr(E|H_2)}{\Pr(E)} \times \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)}\right] - \left[\frac{\Pr(\neg E|H_2)}{\Pr(\neg E)} \times \frac{\Pr(E|H_1)}{\Pr(E)} + \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)} \times \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right]$$

and

$$\left[\frac{\Pr(E|H_1)}{\Pr(E)} - \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)}\right] \times \left[\frac{\Pr(E|H_2)}{\Pr(E)} + \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right] > \left[\frac{\Pr(E|H_2)}{\Pr(E)} - \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right] \times \left[\frac{\Pr(E|H_1)}{\Pr(E)} + \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)}\right]$$

which implies:

$$\frac{\left[\frac{\Pr(E|H_1)}{\Pr(E)} - \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)}\right]}{\left[\frac{\Pr(E|H_1)}{\Pr(E)} + \frac{\Pr(\neg E|H_1)}{\Pr(\neg E)}\right]} > \frac{\left[\frac{\Pr(E|H_2)}{\Pr(E)} - \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right]}{\left[\frac{\Pr(E|H_2)}{\Pr(E)} + \frac{\Pr(\neg E|H_2)}{\Pr(\neg E)}\right]}$$

We can reformulate this as follows:

$$\frac{\left[\frac{\Pr(H_1|E)}{\Pr(H_1)} - \frac{\Pr(H_1|\neg E)}{\Pr(H_1)}\right]}{\left[\frac{\Pr(H_1|E)}{\Pr(H_1)} + \frac{\Pr(H_1|\neg E)}{\Pr(H_1)}\right]} > \frac{\left[\frac{\Pr(H_2|E)}{\Pr(H_2)} - \frac{\Pr(H_2|\neg E)}{\Pr(H_2)}\right]}{\left[\frac{\Pr(H_2|E)}{\Pr(H_2)} + \frac{\Pr(H_2|\neg E)}{\Pr(H_2)}\right]}$$

Finally, by cancelling $\Pr(H_1)$, respectively $\Pr(H_2)$ out of these formulae we get the desired result:

$$\frac{\left[\Pr(H_1|E) - \Pr(H_1|\neg E)\right]}{\left[\Pr(H_1|E) + \Pr(H_1|\neg E)\right]} > \frac{\left[\Pr(H_2|E) - \Pr(H_2|\neg E)\right]}{\left[\Pr(H_2|E) + \Pr(H_2|\neg E)\right]}$$

Proof for ep^3

1. ep^3 satisfies Requirement 1 trivially. It is defined in terms of probabilities.
2. ep^3 satisfies Requirement 2 with marker 0:

   First note that

$$ep^3_{\mathrm{Pr}}(H, E) = \begin{cases} \frac{\mathrm{Pr}(E|H)-\mathrm{Pr}(E)}{1-\mathrm{Pr}(E)} & \text{if } \mathrm{Pr}(E|H) \geq \mathrm{Pr}(E) > 0 \\ \frac{\mathrm{Pr}(E|H)-\mathrm{Pr}(E)}{\mathrm{Pr}(E)} & \text{if } \mathrm{Pr}(E|H) < \mathrm{Pr}(E) \end{cases}$$

   Thus,

$$ep^3_{\mathrm{Pr}}(H, E) = \begin{cases} \frac{\mathrm{Pr}(E|H)-\mathrm{Pr}(E)}{1-\mathrm{Pr}(E)} = \begin{cases} > 0, & \mathrm{Pr}(E|H) > \mathrm{Pr}(E) \\ = 0, & \mathrm{Pr}(E|H) = \mathrm{Pr}(E) \end{cases} & \text{if } \mathrm{Pr}(E|H) \geq \mathrm{Pr}(E) > 0 \\ \frac{\mathrm{Pr}(E|H)-\mathrm{Pr}(E)}{\mathrm{Pr}(E)} = \left\{ < 0, \quad \mathrm{Pr}(E|H) < \mathrm{Pr}(E) \right. & \text{if } \mathrm{Pr}(E|H) < \mathrm{Pr}(E) \end{cases}$$

3. ep^2 satisfies Requirement 3: if $\mathrm{Pr}(E|H_1) > \mathrm{Pr}(E|H_2)$, then
   (a) $ep^3_{\mathrm{Pr}}(H_1, E) = \frac{\mathrm{Pr}(E|H_1)-\mathrm{Pr}(E)}{1-\mathrm{Pr}(E)} > ep^3_{\mathrm{Pr}}(H_2, E) = \frac{\mathrm{Pr}(E|H_2)-\mathrm{Pr}(E)}{1-\mathrm{Pr}(E)}$, if $\mathrm{Pr}(E|H_1) \geq \mathrm{Pr}(E)$ and $\mathrm{Pr}(E|H_2) \geq \mathrm{Pr}(E)$.
   (b) $ep^3_{\mathrm{Pr}}(H_1, E) = \frac{\mathrm{Pr}(E|H_1)-\mathrm{Pr}(E)}{\mathrm{Pr}(E)} > ep^3_{\mathrm{Pr}}(H_2, E) = \frac{\mathrm{Pr}(E|H_2)-\mathrm{Pr}(E)}{\mathrm{Pr}(E)}$, if $\mathrm{Pr}(E|H_1) < \mathrm{Pr}(E)$ and $\mathrm{Pr}(E|H_2) < \mathrm{Pr}(E)$.
   (c) $ep^3_{\mathrm{Pr}}(H_1, E) = \frac{\mathrm{Pr}(E|H_1)-\mathrm{Pr}(E)}{1-\mathrm{Pr}(E)} > ep^3_{\mathrm{Pr}}(H_2, E) = \frac{\mathrm{Pr}(E|H_2)-\mathrm{Pr}(E)}{\mathrm{Pr}(E)}$, if $\mathrm{Pr}(E|H_1) \geq \mathrm{Pr}(E)$ and $\mathrm{Pr}(E|H_2) < \mathrm{Pr}(E)$.

Proof of Theorem 2

Let W be a set of possible worlds and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ which separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let Pr be a strict (or regular) probability function on $\mathcal{A}$. Let Pr* be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\mathrm{Pr}^*(A) = \mathrm{Pr}(A)$ for all $A \in \mathcal{A}$. Then there is a $W' \subseteq W$ with $\mathrm{Pr}^*(W') = 1$ so that the following holds for every $w \in W'$ and all hypotheses $H \in \mathcal{A}$ and for all $\mathfrak{sp}_{\mathrm{Pr}}$ satisfying Requirements 1–3.

Then, according to the Gaifman–Snir Theorem (1982), there is a $W' \subseteq W$ with $\mathrm{Pr}^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories H of $\mathcal{A}$:

$$\lim_{n \Longrightarrow \infty} \mathrm{Pr}(H|E_n^w) = \mathcal{I}(H, w)$$

where $\mathcal{I}(H, w) = 1$, if $w \vDash H$ and 0 otherwise.

1. Suppose $w \vDash H_1$ *and* $w \vDash \neg H_2$. Then $\lim_{n \Longrightarrow \infty} \Pr(H_1|E_n^w) = 1$ and $\lim_{n \Longrightarrow \infty} \Pr(H_2|E_n^w) = 0$ which implies that $\exists n \forall m \geq n : \Pr(H_1|E_n^w) > \Pr(H_1) \& \Pr(H_2|E_n^w) < \Pr(H_2)$. The latter entails by symmetry of probabilistic relevance that $\exists n \forall m \geq n : \Pr(E_n^w|H_1) > \Pr(E_n^w) \& \Pr(E_n^w|H_2) < \Pr(E_n^w)$ and that therefore $\exists n \forall m \geq n : \Pr(E_n^w|H_1) > \Pr(E_n^w|H_2)$. Thus, with Requirement 3 on measures of explanatory and systematic power we can conclude that:

$$\exists n \forall m \geq n : [\mathfrak{sp}_{\Pr}(H_1, E_m^w) > \mathfrak{sp}_{\Pr}(H_2, E_m^w)].$$

2. Suppose $w \vDash H_1 \cap H_2$ and $H_1 \vDash H_2$, but $H_2 \nvDash H_1$. We already know that

$$\lim_{n \to \infty} \left[\frac{\Pr(H|E_n^w)}{\Pr(H)} \right] = \frac{1}{\Pr(H)}, \text{ if } \lim_{n \Longrightarrow \infty} \Pr(H|E_n^w) = 1.$$

and thus that

$$\lim_{n \to \infty} \left[\frac{\Pr(E_n^w|H)}{\Pr(E_n^w)} \right] = \frac{1}{\Pr(H)}, \text{ if } \lim_{n \Longrightarrow \infty} \Pr(H|E_n^w) = 1.$$

The latter implies that

$$\lim_{n \to \infty} \left[\frac{\Pr(E_n^w|H_1)}{\Pr(E_n^w)} \right] = \frac{1}{\Pr(H_1)} > \lim_{n \to \infty} \left[\frac{\Pr(E_n^w|H_2)}{\Pr(E_n^w)} \right] = \frac{1}{\Pr(H_2)}$$

since $H_1 \vDash H_2$ implies that $H_2 \nvDash H_1$, $\Pr(H_1) < \Pr(H_2)$. This means that

$$\lim_{n \to \infty} \left[\Pr(E_n^w|H_1) \right] > \lim_{n \to \infty} \left[\Pr(E_n^w|H_2) \right]$$

and that therefore $\exists n \forall m \geq n : \Pr(E_n^w|H_1) > \Pr(E_n^w|H_2)$. Thus, with Requirement 3 on measures of explanatory and systematic power we can conclude that:

$$\exists n \forall m \geq n : [\mathfrak{sp}_{\Pr}(H_1, E_m^w) > \mathfrak{sp}_{\Pr}(H_2, E_m^w)].$$

where $E_m^w = \bigcap_{0 \leq i \leq m} e_i^w$.

Proof of Theorem 3

Let W be a set of possible worlds and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ which separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let Pr be a strict (or regular) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$. Then there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all hypotheses $H \in \mathcal{A}$ and for all $\mathfrak{sp}_{\Pr}$ satisfying Requirements 1–3.

1. Suppose there is a $H_j \in \{H_1, \ldots, H_n\}$ such that $w \vDash H_j$. Then according to Theorem 2, for every false hypothesis $H_i \in \{H_1, \ldots, H_n\}$ and all for all $\mathfrak{sp}_{\Pr}$ satisfying Requirements 1–3: $\exists n \forall m \geq n : [\mathfrak{sp}_{\Pr}(H_j, E_m^w) > \mathfrak{sp}_{\Pr}(H_i, E_m^w)]$ (note Requirement 3 is the crucial requirement here). Since there are only finitely many false hypotheses in $\{H_1, \ldots, H_n\}$ we can conclude that:

$$\exists n \forall m \geq n \text{ such that } \exists H_j \in \{H_1, \ldots, H_n\} \text{ with } w \vDash H_j \text{ and } \forall H_i \in \{H_1, \ldots, H_n\} \text{ with } w \vDash \neg H_i : [\mathfrak{sp}_{\Pr}(H_j, E_m^w) > \mathfrak{sp}_{\Pr}(H_i, E_m^w)]$$

   Thus, with Definition 9 we can conclude that: $\exists n \forall m \geq n$ such that if H_i is the best systematization for E_m^w with respect to the set of hypotheses $\{H_1, \ldots, H_n\}$, then $w \vDash H_i$. For Definitions 10 and 11 we can show the same since according to Theorem 2 the true hypothesis will be more probable than the false ones after finitely many steps of observation and for every observation thereafter.
2. The proof for the second part proceeds along the same lines.

where $E_m^w = \bigcap_{0 \leq i \leq m} e_i^w$.

Proof of Theorem 6

$$\Pr_{t_0}(H|E) = \frac{\dfrac{e^{\tanh^{-1}\left[sp^2{}_{t_0}(H,E)\right]} \Pr_{t_0}(E)}{e^{\tanh^{-1}\left[sp^2{}_{t_0}(H,E)\right]} \Pr_{t_0}(E) + e^{\tanh^{-1} sp^2{}_{t_0}(H,\neg E)} \Pr_{t_0}(\neg E)}}{\Pr_{t_0}(E)} \times \Pr_{t_0}(H)$$

$$= \frac{\dfrac{e^{\frac{1}{2}\left[\log\left[sp^2{}_{t_0}(H,E)+1\right]-\log\left[1-sp^2{}_{t_0}(H,E)\right]\right]} \Pr_{t_0}(E)}{e^{\frac{1}{2}\left[\log\left[sp^2{}_{t_0}(H,E)+1\right]-\log\left[1-sp^2{}_{t_0}(H,E)\right]\right]} \Pr_{t_0}(E) + e^{\frac{1}{2}\left[\log\left[sp^2{}_{t_0}(H,\neg E)+1\right]-\log\left[1-sp^2{}_{t_0}(H,\neg E)\right]\right]} \Pr_{t_0}(\neg E)}}{\Pr_{t_0}(E)} \times \Pr_{t_0}(H)$$

$$= \frac{\dfrac{e^{\frac{1}{2}\left[\log\left[\frac{sp^2{}_{t_0}(H,E)+1}{1-sp^2{}_{t_0}(H,E)}\right]\right]} \Pr_{t_0}(E)}{e^{\frac{1}{2}\left[\log\left[\frac{sp^2{}_{t_0}(H,E)+1}{1-sp^2{}_{t_0}(H,E)}\right]\right]} \Pr_{t_0}(E) + e^{\frac{1}{2}\left[\log\left[\frac{sp^2{}_{t_0}(H,\neg E)+1}{1-sp^2{}_{t_0}(H,\neg E)}\right]\right]} \Pr_{t_0}(\neg E)}}{\Pr_{t_0}(E)} \times \Pr_{t_0}(H)$$

Now we know that

$$\frac{\dfrac{\Pr_{t_0}(\neg E|H)}{\Pr_{t_0}(\neg E)}}{\dfrac{\Pr_{t_0}(E|H)}{\Pr_{t_0}(E)}} = \frac{e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|\neg E)}{\Pr_{t_0}(H|E)}\right]}}{e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}} \quad \text{(by the definition of } e \text{ and } \log)$$

$$\frac{\Pr_{t_0}(\neg E|H)}{\Pr_{t_0}(E|H)} = \frac{\Pr_{t_0}(\neg E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|\neg E)}{\Pr_{t_0}(H|E)}\right]}}{\Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}}$$

$$\frac{\Pr_{t_0}(\neg E|H)}{\Pr_{t_0}(E|H)} + \frac{\Pr_{t_0}(E|H)}{\Pr_{t_0}(E|H)} = \frac{\Pr_{t_0}(\neg E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|\neg E)}{\Pr_{t_0}(H|E)}\right]}}{\Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}} + \frac{\Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}}{\Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}}$$

$$\frac{1}{\Pr_{t_0}(E|H)} = \frac{\Pr_{t_0}(\neg E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|\neg E)}{\Pr_{t_0}(H|E)}\right]} + \Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}}{\Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}}$$

$$\Pr_{t_0}(E|H) = \frac{\Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}}{\Pr_{t_0}(\neg E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|\neg E)}{\Pr_{t_0}(H|E)}\right]} + \Pr_{t_0}(E) \times e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}}$$

$$\Pr_{t_0}(E|H) = \frac{e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}\Pr_{t_0}(E)}{e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|E)}{\Pr_{t_0}(H|\neg E)}\right]}\Pr_{t_0}(E) + e^{\frac{1}{2}\log\left[\frac{\Pr_{t_0}(H|\neg E)}{\Pr_{t_0}(H|E)}\right]}\Pr_{t_0}(\neg E)}$$

and since $\frac{sp^2{}_{t_0}(H,E)+1}{1-sp^2{}_{t_0}(H,E)} = \frac{\Pr(H|E)}{\Pr(H|\neg E)}$ we can conclude that

$$\Pr_{t_0}(H|E) = \frac{\Pr_{t_0}(E|H)}{\Pr_{t_0}(E)} \times \Pr_{t_0}(H)$$

References

Belot, G. (2013). Bayesian orgulity. *Philosophy of Science*, *80*, 483–503.

Brössel, P. (2012). *Rethinking Bayesian confirmation theory—Steps towards a a new Bayesian theory of confirmation*. PhD-thesis, University of Konstanz.

Brössel, P. (2008). Theory assessment and coherence. *Abstracta*, *4*, 57–71.

Brössel, P. (2014). Assessing theories: The coherentist approach. *Erkenntnis*, *79*, 593–623.

Brössel, P. (2015). Keynes's coefficient of dependence revisited. *Erkenntnis*, *80*, 521–553.

Brössel, P., & Eder, A.-M. A. (2014). How to resolve doxastic disagreement. *Synthese*, *191*, 2359–2381.

Carnap, R. (1950). Empiricism, semantics, and ontology. *Revue Inter-nationale de Philosophie, 4*, 20–40. Reprinted in the Supplement to Carnap, R. (1956). *Meaning and necessity: A study in semantics and modal logic*. Chicago: University of Chicago Press.

Crupi, V. (2013). Confirmation. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/archives/win2013/entries/confirmation/.

Crupi, V., et al. (2007). On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science*, *74*, 229–252.

Crupi, V., & Tentori, K. (2012). A second look at the logic of explanatory power (with two novel representation theorems). *Philosophy of Science*, *79*, 365–385.

Darwin, C. (1872). *The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life* (6th ed.). London: John Murray.

Douven, I. (2011). Abduction. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2011 Edition). http://plato.stanford.edu/archives/spr2011/entries/abduction/.

Earman, J. (1992). *Bayes or bust? A critical examination of Bayesian confirmation theory*. Cambridge: MIT Press.

Friedman, M. (2002). Kant, Kuhn, and the rationality of science. *Philosophy of Science*, *69*, 171–190.

Gaifman, H., & Snir, M. (1982). Probabilities over rich languages, testing, and randomness. *Journal of Symbolic Logic*, *47*, 495–548.

Good, I. (1960). Weight of evidence, corroboration, explanatory power, information and the utility of experiments. *Journal of The Royal Statistical Society. Series B (Methodological)*, *22*, 319–331.

Harman, G. (1965). The inference to the best explanation. *Philosophical Review*, *74*, 88–95.

Harman, G. (1967). Detachment, probability, and maximum likelihood. *Noûs*, *1*, 401–411.

Hawthorne, J. (2014). Inductive logic. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2014 Edition). http://plato.stanford.edu/archives/sum2014/entries/logic-inductive/.

Hempel, C. (1958). The theoretician's dilemma. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science* (Vol. 2). Minneapolis: University of Minnesota Press.

Hempel, C. (1960). Inductive inconsistencies. *Synthese*, *12*, 439–469.
Hempel, C., & Oppenheim, P. (1948). Studies in the logic of explanation. In C. Hempel (Ed.), *Aspects of scientific explanation and other essays in the philosophie of science* (pp. 245–291). New York: Free Press. (1965).
Horwich, P. (1982). *Probability and evidence*. Cambridge: Cambridge University Press.
Howson, C. (2003). *Hume's problem*. Oxford: Oxford University Press.
Huber, F. (2008). Assessing theories, Bayes' style. *Synthese*, *161*, 89–118.
Huttegger, S. (2015a). Merging of opinions and probability kinematics. *Review of Symbolic Logic* (Accepted).
Huttegger, S. (2015b). Bayesian convergence to the truth and the metaphysics of possible worlds. *Philosophy of Science* (Accepted).
Jeffrey, R. (1956). Valuation and acceptance of scientific hypotheses. *Philosophy of Science*, *23*, 237–246.
Jeffrey, R. (1992). Radical probabilism (prospectus for a user's manual). *Philosophical Issues*, *2*, 193–204.
Kemeny, J., & Oppenheim, P. (1952). Degree of factual support. *Philosophy of Science*, *19*, 307–324.
Keynes, J. (1921). *A treatise on probability*. London: Macmillan.
Kolmogorov, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Julius Springer.
Levi, I. (1967). *Gambling with truth*. New York: A. A. Knopf.
Lewis, D. (1973). *Counterfactuals*. Harvard: Harvard University of Press.
McMullin, E. (1992). *The inference that makes science*. Milwaukee: Marquette University Press.
Milne, P. (1996). $log[p(h/eb)/p(h/b)]$ is the one true measure of confirmation. *Philosophy of Science*, *63*, 21–26.
Newton, I. (1729). *The mathematical principles of natural philosophy* (A. Motte, Trans.). New York: Daniel Adler.
Niiniluoto, I. (2011). Scientific progress. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2011 Edition). http://plato.stanford.edu/archives/sum2011/entries/scientific-progress/.
Okasha, S. (2000). Van Fraassen's critique of inference to the best explanation. *Studies in History and Philosophy of Science*, *31*, 691–710.
Popper, K. (1959). *The logic of scientific discovery*. London: Hutchinson.
Psillos, S. (1996). On van Fraassen's critique of abductive reasoning. *The Philosophical Quarterly*, *46*, 31–47.
Rescher, N. (2005). *Studies in pragmatism*. Heusenstamm: Ontos.
Schervish, M., & Seidenfeld, T. (1990). An approach to consensus and certainty with increasing evidence. *Journal of Statistical Planning and Inference*, *25*, 401–414.
Schupbach, J. (2011). Comparing probabilistic measures of explanatory power. *Philosophy of Science*, *78*, 813–829.
Schupbach, J. (2014). Is the bad lot objection just misguided? *Erkenntnis*, *79*, 55–64.
Schupbach, J., & Sprenger, J. (2011). The logic of explanatory power. *Philosophy of Science*, *78*, 105–127.
van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
van Fraassen, B. (1989). *Laws and symmetry*. Oxford: Oxford University Press.
Vogel, J. (1990). Cartesian skepticism and inference to the best explanation. *Journal of Philosophy*, *87*, 658–666.
Weisberg, J. (2009). Locating IBE in the Bayesian framework. *Synthese*, *167*, 125–143.