European Journal for Philosophy of Science (2025) 15:3
https://doi.org/10.1007/s13194-024-00630-4

PAPER IN GENERAL PHILOSOPHY OF SCIENCE

Stopping rule and Bayesian confirmation theory

Yunbing Li[1] · Yongfeng Yuan[1]

Received: 5 April 2024 / Accepted: 17 December 2024

Abstract
This article mainly investigates whether common Bayesian confirmation measures are affected by stopping rules. The results indicate that difference measure d, log-ratio measure r, and log-likelihood measure l are not affected by non-informative stopping rules, but affected by informative stopping rules. In contrast, Carnap measure τ, normalized difference measure n, and Mortimer measure m are affected by (non-)informative stopping rules sometimes but sometimes aren't. Besides, we use two examples to further illustrate that confirmation measures d, r, and l are better than τ, n, and m.

Keywords Degree of confirmation · Confirmation measure · Informative stopping rule · Non-informative stopping rule

1 Introduction

New evidence may confirm, disconfirm, or be irrelevant to a hypothesis. Confirmation theories study the evaluation of a hypothesis based on pieces of evidence. Bayesian confirmation theorists have proposed various measures that rely on Bayes' theorem (Bayes, 1763) to measure the degree of confirmation of a hypothesis H provided by a piece of evidence E (Fitelson, 2001; Christensen, 1999; Carnap, 1950). These functions are called *confirmation measures*, and denoted by $c(H, E)$. According to most Bayesian confirmation theories, a piece of evidence *confirms* a hypothesis, iff, the probability of the hypothesis given the evidence is greater than it would be without the evidence; iff, the posterior probability of the hypothesis is greater than its prior probability. Formally speaking, confirmation measures should typically satisfy the following requirements (Glass, 2013; Brössel, 2013):

Requirement 1 *E confirms H, i.e.,* $c(H, E) > 0$*, iff* $Pr(H|E) > Pr(H)$.

✉ Yongfeng Yuan
ganlanshux@126.com

Yunbing Li
yunbing94@gmail.com

1 Department of Philosophy (Zhuhai), Sun Yat-Sen University, Tangjiawan, Xiangzhou District, Zhuhai 519082, Guangdong, China

Published online: 28 December 2024

Requirement 2 *E is irrelevant to H, i.e.,* $c(H, E) = 0$*, iff* $Pr(H|E) = Pr(H)$.

Requirement 3 *E disconfirms H, i.e.,* $c(H, E) < 0$*, iff* $Pr(H|E) < Pr(H)$.

In scientific research, scientists usually evaluate (confirm or disconfirm) a hypothesis based on experimental data/evidence. Nevertheless, sometimes scientists will consider the way for an experiment to stop as a piece of evidence, and evaluate its significant impact on the confirmation of the hypothesis. We call such a way for an experiment to stop as *stopping rules*. There are two kinds of stopping rules, namely informative stopping rules and non-informative stopping rules. A stopping rule is *non-informative* if the decision to stop or continue the experiment is only based on the experimental data and is independent of the hypothesis (Gandenberger, 2017); otherwise, it is *informative*.

Whether stopping rules, mainly those non-informative ones, influence the evaluation of hypotheses has been debated for several years between Frequentists and Bayesians. According to Frequentists, stopping rules should be regarded as evidence to evaluate hypotheses (Dienes, 2016). This is because stopping rules are not mysterious, can be reported after experiments, and are conclusively verifiable (Fletcher, 2019). Moreover, Frequentists' statistical methods, such as hypothesis testing based on p-values, are indeed influenced by stopping rules (Berger & Berry, 1988; Goodman & Royall, 1988; Urbach & Howson, 1993).

In contrast, Bayesians argue that the evaluation of a hypothesis through statistical methods should depend only on the experimental data, regardless of the non-informative stopping rules. Thus, under different non-informative stopping rules, if the experimental data is the same, then the evaluation of the hypothesis will also be the same. Besides, regarding non-informative stopping rules as evidence or not should not affect the final evaluation of the hypothesis. Hence, *hypothesis evaluation based on Bayesian theory should not be affected by non-informative stopping rules* (C1). For instance, Bayes factor hypothesis testing is unaffected by non-informative stopping rules (Goodman, 2005), and this is regarded as one of the reasons for its advantage over Frequentist hypothesis evaluation methods (Berger & Berry, 1988). As theories for Bayesians, we desire that Bayesian confirmation measures should be unaffected by non-informative stopping rules as well. Besides non-informative stopping rules, the informative ones should also be considered. Since informative stopping rules depend on hypotheses directly, *Bayesian confirmation measures are desired to be affected by informative stopping rules* (C2).

Note that, for Bayesians, if a Bayesian confirmation measure does not depend only upon the experimental data, but also be affected by a non-informative stopping rule which is independent of or irrelevant to the hypothesis, then this measure is not objective; because it is irrational for any measure to be influenced by an independent or irrelevant factor. For example, let the non-informative stopping rule be that: if we reach a pre-specified large number of trials, then we stop the experiment. Because this stopping rule is irrelevant to the hypothesis, it should not affect the evaluation of the hypothesis. For informative stopping rules, how the experiment stops is relevant to the hypothesis. Thus, if a Bayesian confirmation measure is not affected by informative stopping rules, then this measure is also not objective; because it is irrational for any measure to be unaffected by a dependent or relevant factor. For example, suppose that an experiment stops after obtaining a piece of data E, and this stopping event is more likely to occur when the hypothesis H is true compared to when $\neg H$ is true. In such a case, the stopping rule is informative, and the stopping event itself should support H. Hence, for convenience, we call the requirements (C1) and (C2) on the two kinds of stopping rules as the *objective requirements* of Bayesian confirmation theory. Then we have an interesting question: do all Bayesian confirmation measures satisfy such objective requirements? In this article, we will argue that: some confirmation measures are objective while some others are not. Besides, we will use some examples to illustrate why objective confirmation measures outperform the non-objective ones.

The structure of the paper is as follows. Section 2 will further illustrate how non-informative stopping rules affect Frequentists' statistical methods, while Bayesian statistical methods are unaffected. It is desired that Bayesian confirmation measures should not be affected by non-informative stopping rules. Section 3 will firstly introduce some relevant concepts and common confirmation measures, and then argue that difference measure d, log-ratio measure r, and log-likelihood measure l are unaffected by non-informative stopping rules but affected by informative stopping rules. After that, we will argue that Carnap measure τ, normalized difference measure n and Mortimer measure m are affected by (non-)informative stopping rules sometimes but sometimes are not. Thus, according to the concept of objectivity, d, r and l are objective, while τ, n and m are not objective; and the former three are considered better than the latter three. Besides, we also study the relation between the concept of objectivity and Final Probability. Section 4 will use two examples to further illustrate that confirmation measures with objectivity are better than those without objectivity.

2 Classical statistical tests, Bayesian statistical tests and non-informative stopping rules

In this section, we will further illustrate how non-informative stopping rules affect Frequentists' statistical methods, while Bayesian statistical methods are unaffected. Now let's consider the following example first.

Example 1 Let's test whether a coin is fair. The hypothesis H_0 is that the chance of heads is 0.5. We have tossed this coin 20 times, and then the outcome space is the set $X = \{H, T\}^{20}$. The experimental result is 6 heads and 14 tails.

We use Fisherian statistical tests on behalf of Frequentists. Typically, Fisherian statistical tests are based on p-values. The p-value of an experiment depends on the possible outcomes, that is to say, depends on the outcome space. For any experimental data E of an experiment x and *null* hypothesis H_0, the p-value of this experiment can be calculated as follows:

$$p\text{-}value_{x,E} = \sum_{\{E' \in X | Pr(E'|H_0) \leq Pr(E|H_0)\}} Pr(E'|H_0)$$

It means that: the p-value of an experiment is the probability of observing the observed data or other possible data that is at least as extreme as that observed, given that the null hypothesis H_0 is true. If the p-value is lower than the predefined threshold (known as the significance level, e.g., 0.05), then we reject the null hypothesis; otherwise, we will conclude that there is no sufficient reason to reject the null hypothesis. Now let's consider two non-informative stopping rules, S_1 and S_2, for Example 1 to provide a further illustration. The stopping rule S_1 stops the experiment after 20 tosses, while the stopping rule S_2 stops the experiment after obtaining 6 heads. Let H_0 be the null hypothesis, and let's take the most common level 0.05 to be the significance level. If the stopping rule in Example 1 is S_1, then

$$\begin{aligned} p\text{-}value_{x,E} &= \sum_{\{E' \in X | Pr(E'|H_0) \leq Pr(E|H_0)\}} Pr(E'|H_0) \\ &= \sum_{0 \leq i \leq 6} C_{20}^{i} (\frac{1}{2})^{i} \times (\frac{1}{2})^{20-i} + \sum_{14 \leq i \leq 20} C_{20}^{i} (\frac{1}{2})^{i} \times (\frac{1}{2})^{20-i} \\ &= 0.116. \end{aligned}$$

Since $0.116 > 0.05$, we do not have sufficient reason to reject H_0. If the stopping rule in Example 1 is S_2, then

$$p\text{-}value_{x,E} = \sum_{n \geq 20} C_{n-1}^{5} (\frac{1}{2})^{5} \times (\frac{1}{2})^{n-6} \times \frac{1}{2} = 0.0319.$$

Since $0.0319 < 0.05$, we should reject H_0. Hence, with the same evidence, different stopping rules will lead to different evaluations of hypotheses for classical statistical tests based on p-value. Since non-informative stopping rules have an impact on the evaluation of hypotheses for classic statistical tests, they are supposed to be applied to evaluate hypotheses.

Now we turn to Bayesian statistical tests and show that they are not affected by non-informative stopping rules. Unlike Fisherian statistical tests, Bayesian tests depend only on the data we obtained. We use the Bayes factor to help us evaluate hypotheses. For any experimental data E of an experiment x and hypotheses H_1 and H_2, the Bayes factor of this experiment can be calculated as follows:

$$BF_{x,E} = \frac{Pr(E|H_1)}{Pr(E|H_2)}$$

Bayesians may establish a set of thresholds to decide whether a piece of evidence provides more evidential support to a hypothesis than to the other. For instance, it is suggested that: when $1 \leq BF_{x,E} < 3.2$, the support that E provides for H_1 compared to H_2 is negligible; when $3.2 \leq BF_{x,E} < 10$, E provides substantial support for H_1; when $10 \leq BF_{x,E} < 100$, E provides strong support for H_1; finally, when $100 \leq BF_{x,E}$, E decisively supports H_1 (Jeffreys, 1961, p. 396). Clearly, none of $Pr(E|H_1)$, $Pr(E|H_2)$, or the thresholds will be affected by non-informative stopping rules. This means that, no matter what the non-informative stopping rule is, the evaluation of the evidence E in favor of H_1 against H_2 remains the same. So the evaluation of hypotheses by such a Bayesian statistical test is not affected by non-informative stopping rules.

It is natural to expect that Bayesian confirmation measures also have the property which Bayesian statistical testing methods have. That is to say, with the same evidence, the choice of non-informative stopping rules does not affect the evaluation results for the hypotheses. The evaluation of hypotheses should be independent of the non-informative stopping rules. For example, in an experiment, the data E alone supports the hypothesis H. Then after considering the non-informative stopping rule, the strength of evidence E in favor of hypothesis H should not change, and certainly should not shift from a supporting relationship to a weakening or unaffected one. However, as we will see in Section 3, not all confirmation measures satisfy such a requirement.

3 Common confirmation measures and stopping rules

In this section, we will firstly introduce some relevant concepts and common confirmation measures, and then argue that some confirmation measures are unaffected by non-informative stopping rules but affected by informative stopping rules, and thus they are objective for Bayesians. After that, we will argue that some other confirmation measures are affected by (non-)informative stopping rules sometimes but sometimes are not, and thus they are not objective for Bayesians.

For a more rigorous analysis and discussion, we need to clarify firstly some relevant concepts. For example, the concepts of informative and non-informative stopping rules, what do we mean by "the stopping rule affects the degree of confirmation", and the concept of objectivity of Bayesian confirmation measures. Now let's denote the hypothesis by H, and the experimental data by E. Let $0 < Pr(H|E) < 1$ and $0 < Pr(\neg H|E) < 1$, which means that E is non-conclusive evidence for the hypothesis H. Note that, the evidence E in the following are all non-conclusive. Since $Pr(H|E) \neq 0$, $Pr(H|E) = \frac{Pr(E|H)\cdot Pr(H)}{Pr(E|H)\cdot Pr(H)+Pr(E|\neg H)\cdot Pr(\neg H)} \neq 0$, so we have $Pr(E|H) > 0$; similarly, we also have $Pr(E|\neg H) > 0$. And let's denote the stopping rule by R, and use R_E to denote the stopping event where the experiment stops due to the hypothesis H, stopping rule R and experimental data E. Note that, what can confirm or disconfirm a hypothesis must be a particular fact or evidence. Thus, when we say that the degree of confirmation of the hypothesis is affected (unaffected) by the *normative* stopping rule R, it is actually the stopping event R_E which provides confirmation or disconfirmation for the hypothesis. Nevertheless, to retain the terminology in the literature, we will still use expressions such as "a stopping rule confirms or disconfirm a hypothesis" in our discussions.

Since, for a non-informative stopping rule, the decision to stop/continue the experiment is only based on the experimental data and is independent of the hypothesis, the probability of the stopping event R_E should be independent of the hypothesis, namely should not be affected by H and $\neg H$, no matter what the evidence E is. In contrast, for an informative stopping rule, the decision to stop/continue the experiment depends upon the hypothesis. So the probability of the stopping event R_E should depend upon the hypothesis, namely should be affected by H and $\neg H$, no matter what the evidence E is. Thus, we have the following definition.

Definition 1 Roberts (1967)

(1) A stopping rule R is non-informative iff $\forall E(Pr(R_E|H\&E) = Pr(R_E|\neg H\&E))$.
(2) A stopping rule R is informative iff $\forall E(Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E))$.[1]

Note that, $Pr(R_E|H\&E)$ and $Pr(R_E|\neg H\&E)$ are chances rather than the probability of a particular event's occurrence. Nevertheless, the stopping event R_E's probability may or may not be different with respect to different hypotheses H and $\neg H$. When the decision to stop/continue the experiment is *independent of* or *irrelevant to* the hypothesis (e.g., stop the experiment when the equipment is broken), the stopping event R_E's probability should not be affected by H or $\neg H$, i.e., $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$. It is similar to the situation that: given two independent events A and B, we have $Pr(A|B) = Pr(A|\neg B)$. In contrast, when the decision to stop/continue the experiment depends upon, or is *not independent* of, the hypothesis, the stopping event R_E's probability should be affected by H and $\neg H$, i.e., $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. It is similar to the situation that: given two dependent events A and B, we usually have $Pr(A|B) \neq Pr(A|\neg B)$.

In intuition, if the degree of confirmation of a hypothesis H is unaffected by a stopping rule R, then the degree of confirmation provided by the evidence E should be equal to that provided by the evidence $E\&R_E$; if a stopping rule R supports a hypothesis H, then the degree of confirmation provided by the evidence E should be less than that provided by the evidence $E\&R_E$; similarly, if a stopping rule R weakens a hypothesis H, then the degree of confirmation provided by the evidence E should be greater than that provided by the evidence $E\&R_E$. Thus, we have the following definition.

Definition 2 (1) The degree of confirmation of a hypothesis H is unaffected by a stopping rule R iff $c(H, E) = c(H, E\&R_E)$;
(2) A stopping rule R supports a hypothesis H iff $c(H, E) < c(H, E\&R_E)$;
(3) A stopping rule R weakens a hypothesis H iff $c(H, E) > c(H, E\&R_E)$.

As we introduced in Section 1, Bayesian confirmation measures should be unaffected by non-informative stopping rules (C1), but affected by informative stopping rules (C2). We call such requirements as the *objective requirements* of Bayesian confirmation theory. For convenience, here we present the following definition.

Definition 3 A confirmation measure $c(H, E)$ is objective iff it satisfies the following conditions:

(1) it is unaffected by any non-informative stopping rule R; and
(2) it is affected by any informative stopping rule R, i.e., R supports or weakens H.

Note that, the concept of objectivity newly defined here is different from others in the theories of probability (e.g., "objective Bayesian"), it merely considers whether the confirmation measure is affected by (non-)informative stopping rules.

[1] Someone may regard a stopping rule R as informative if $\exists E(Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E))$. But we think such restriction is too weak. If there is a piece of evidence E in a given experiment such that $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$, then it is hard to tell why the stopping rule here is informative. Even there might exist another possible evidence E' such that $Pr(R_E|H\&E') \neq Pr(R_E|\neg H\&E')$, the stopping rule does not provide any information in the previous situation.

In the following, we follow Steel (2003) to discuss whether the following well-known confirmation measures satisfy the objective requirements in Definition 3.

(1) Difference measure (Carnap, 1950): $d(H, E) = Pr(H|E) - Pr(H)$
(2) Log-ratio measure (Milne, 1996):

$$r(H, E) = \begin{cases} log[\frac{Pr(H|E)}{Pr(H)}] & \text{if } Pr(H|E) > 0 \\ -\infty & \text{if } Pr(H|E) = 0 \end{cases}$$

(3) Log-likelihood measure (Good, 1960; Fitelson, 2001):

$$l(H, E) = \begin{cases} log[\frac{Pr(E|H)}{Pr(E|\neg H)}] & \text{if } Pr(E|H) > 0 \text{ and } Pr(E|\neg H) > 0 \\ \infty & \text{if } Pr(E) > 0 \text{ and } Pr(E|\neg H) = 0 \\ -\infty & Pr(E) = 0 \text{ or } Pr(E|H) = 0 \end{cases}$$

(4) Carnap measure (Carnap, 1950): $\tau(H, E) = Pr(H\&E) - Pr(H) \cdot Pr(E)$
(5) Normalized difference measure (Nozick, 1981): $n(H, E) = Pr(E|H) - Pr(E|\neg H)$
(6) Mortimer measure (Mortimer, 1988): $m(H, E) = Pr(E|H) - Pr(E)$

We are now ready to investigate whether the six common Bayesian confirmation measures above are affected by stopping rules. We begin with confirmation measures d, r, and l. The following is the main result related to the difference measure d.

Theorem 1 *The difference measure d is unaffected by non-informative stopping rules, but affected by informative stopping rules. So difference measure d is objective.*

Proof Let's first calculate the degree of confirmation of hypothesis H based on the compositive evidence $E\&R_E$ as follows. According to Bayes' theorem, we have:

$$\begin{aligned} d(H, E\&R_E) &= Pr(H|E\&R_E) - Pr(H) \\ &= \frac{Pr(E\&R_E|H) \cdot Pr(H)}{Pr(E\&R_E|H) \cdot Pr(H) + Pr(E\&R_E|\neg H) \cdot Pr(\neg H)} - Pr(H) \\ &= \frac{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H)}{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) + Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H)} - Pr(H) \end{aligned}$$

Now let's analyze the impact of (informative and non-informative) stopping rules on the difference measure d, respectively.

(1) Let R be a non-informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$. Since all scientific experiments will eventually stop, it is reasonable to assume that both of $Pr(R_E|H\&E)$ and

$Pr(R_E|\neg H\&E)$ are greater than 0. Then we have:

$$\begin{aligned} d(H, E\&R_E) &= \frac{Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)}{Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)+Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H)\\ &= \frac{Pr(E|H)\cdot Pr(H)}{Pr(E|H)\cdot Pr(H)+Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H)\\ &= Pr(H|E) - Pr(H)\\ &= d(H, E) \end{aligned}$$

Thus, according to Definition 2, non-informative stopping rules do not affect the difference measure d.

(2) Let R be an informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. Suppose that $Pr(R_E|\neg H\&E) = \alpha$, $Pr(R_E|H\&E) = \beta$, and $\alpha \neq \beta$. Then we have:

$$\begin{aligned} d(H, E\&R_E) &= \frac{Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)}{Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)+Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H)\\ &= \frac{\beta\cdot Pr(E|H)\cdot Pr(H)}{\beta\cdot Pr(E|H)\cdot Pr(H)+\alpha\cdot Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H). \end{aligned}$$

Since all scientific experiments will eventually stop, it is reasonable to assume that at least one of $Pr(R_E|H\&E)$ and $Pr(R_E|\neg H\&E)$ is greater than 0. First assume $\alpha > 0$ and $\beta > 0$. Then we have:

$$\begin{aligned} d(H, E\&R_E) &= \frac{\beta\cdot Pr(E|H)\cdot Pr(H)}{\beta\cdot Pr(E|H)\cdot Pr(H)+\alpha\cdot Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H)\\ &= \frac{Pr(E|H)\cdot Pr(H)}{Pr(E|H)\cdot Pr(H)+\frac{\alpha}{\beta}\cdot Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H). \end{aligned}$$

Since $\alpha \neq \beta$, we have the following two situations. When $\alpha > \beta$, we can derive $d(H, E\&R_E) < d(H, E)$. Then, according to Definition 2, the informative stopping rule R weakens the hypothesis H. When $\alpha < \beta$, we have $d(H, E\&R_E) > d(H, E)$. Then, according to Definition 2, the informative stopping rule R supports the hypothesis H. Next, suppose that $\alpha = 0$ and $\beta \neq 0$, we have:

$$\begin{aligned} d(H, E\&R_E) &= \frac{\beta\cdot Pr(E|H)\cdot Pr(H)}{\beta\cdot Pr(E|H)\cdot Pr(H)+\alpha\cdot Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H)\\ &= 1 - P(H) \end{aligned}$$

Since $d(H, E) = P(H|E) - P(H)$ and E is non-conclusive, we have $P(H|E) < 1$. Thus $d(H, E) < d(H, E\&R_E)$. Then, according to Definition 2, the stopping rule supports the hypothesis H. When $\alpha \neq 0$ and $\beta = 0$, we have:

$$\begin{aligned} d(H, E\&R_E) &= \frac{\beta\cdot Pr(E|H)\cdot Pr(H)}{\beta\cdot Pr(E|H)\cdot Pr(H)+\alpha\cdot Pr(E|\neg H)\cdot Pr(\neg H)} - Pr(H)\\ &= -P(H) \end{aligned}$$

Since $d(H, E) = P(H|E) - P(H)$ and E is non-conclusive, we have $P(H|E) > 0$. Thus $d(H, E) > d(H, E\&R_E)$. Then, according to Definition 2, the stopping rule weakens the hypothesis H. Thus according to Definition 2, difference measure d is affected by informative stopping rules.

Hence, according to Definition 3, the difference measure d does satisfy the objective requirements. □

For log-ratio measure r and log-likelihood measure l, it turns out that: they are unaffected by non-informative stopping rules, but affected by informative stopping rules just like the difference measure d.

Theorem 2 *The log-ratio measure r and log-likelihood measure l are unaffected by non-informative stopping rules, but affected by informative stopping rules. So they are objective.*

Proof Similar to the proof for Theorem 1, please see Appendix A. □

The results above show that confirmation measures d, r, and l are all objective. Now let us turn to Carnap measure τ, the normalized difference measure n, and Mortimer m to see whether they are affected by stopping rules. The following is the main result related to the Carnap measure τ.

Theorem 3 *The Carnap measure τ is affected by stopping rules sometimes and sometimes not. So Carnap measure τ is not objective.*

Proof Let's first calculate the degree of confirmation of hypothesis H based on compositive evidence $E\&R_E$. The calculation is as follows:

$$\begin{aligned}\tau(H, E\&R_E) &= Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H) - Pr(H)\cdot Pr(E\&R_E)\\ &= Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H) - Pr(H)\cdot (Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)\\ &\quad + Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H)\cdot Pr(\neg H))\end{aligned}$$

In the following, let's analyze whether (informative and non-informative) stopping rules will affect Carnap measure τ, respectively.

(1) Let R be a non-informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$. Since all scientific experiments will eventually stop, it is reasonable to assume that both of $Pr(R_E|H\&E)$ and $Pr(R_E|\neg H\&E)$ are greater than 0. Let's denote their values as i ($i > 0$). Then we have:

$$\begin{aligned}\tau(H, E\&R_E) &= Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H) - Pr(H)\cdot (Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)\\ &\quad + Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H)\cdot Pr(\neg H))\\ &= i(Pr(E|H)\cdot Pr(H) - Pr(H)\cdot (Pr(H)\cdot Pr(E|H) + Pr(E|\neg H)\cdot Pr(\neg H))\\ &= i(Pr(H\&E) - Pr(H)\cdot Pr(E))\\ &= i(\tau(H, E))\end{aligned}$$

When $i = 1$, we have $\tau(H, E\&R_E) = \tau(H, E)$. Then, according to Definition 2, Carnap measure τ is not affected by the non-informative stopping rule R. Nevertheless, when $i \neq 1$, $\tau(H, E\&R_E) \neq \tau(H, E)$, and then, according to Definition 2, Carnap measure τ is affected by the non-informative stopping rule R.

(2) Let R be an informative stopping rule. Then, according to Definition 1, $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. Let's assume that $Pr(R_E|\neg H\&E) = \alpha$, $Pr(R_E|H\&E) = \beta$ $(\alpha \neq \beta)$, $Pr(E|H) = \gamma$ $(0 < \gamma < 1)$, $Pr(E|\neg H) = \delta$ $(0 < \delta < 1)$ and $Pr(H) = \epsilon$ $(0 < \epsilon < 1)$, then we have:

$$\begin{aligned}\tau(H, E\&R_E) &= Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) - Pr(H) \cdot (Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) \\ &\quad + Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H)) \\ &= \epsilon \cdot (1 - \epsilon) \cdot (\beta \cdot \gamma - \alpha \cdot \delta) \\ \tau(H, E) &= Pr(E|H) \cdot Pr(H) - (Pr(E|H) \cdot Pr(H) + Pr(E|\neg H) \cdot Pr(\neg H)) \cdot Pr(H) \\ &= \epsilon \cdot (1 - \epsilon) \cdot (\gamma - \delta)\end{aligned}$$

It is easy to see that: if $\beta = \frac{\alpha \cdot \delta + \gamma - \delta}{\gamma}$ then $\tau(H, E\&R_E) = \tau(H, E)$. According to Definition 2, when $\beta = \frac{\alpha \cdot \delta + \gamma - \delta}{\gamma}$, Carnap measure τ is unaffected by informative stopping rule R. In contrast, when $\beta \neq \frac{\alpha \cdot \delta + \gamma - \delta}{\gamma}$, Carnap measure τ is affected by the informative stopping rule R.

Hence, according to Definition 3, Carnap measure τ does not satisfy the objective requirements for some situations. □

For normalized measure n and Mortimer measure m, it turns out that they are affected by stopping rules sometimes and sometimes are not, just like the Carnap measure τ.

Theorem 4 *The normalized measure n and the Mortimer measure m are affected by stopping rules sometimes and sometimes are not. So they are not objective.*

Proof Similar to the proof for Theorem 3, please see Appendix A. □

In summary, the confirmation measures τ, n, and m are unaffected by non-informative stopping rules in some situations, but affected in other situations. Similarly, they are unaffected by informative stopping rules in some situations, but affected in other situations. Thus, all of them don't satisfy the objective requirements. According to the above discussions, we know that: the confirmation measures d, r and l are objective, but the confirmation measures τ, n, and m lack the property of objectivity. Their properties are listed in Table 1, where ✗ means the confirmation measure is unaffected by (non-)informative stopping rules and ✓ means the confirmation measure is affected by (non-)informative stopping rules.

Interestingly, in the literature there is a requirement labeled *Final Probability* (Crupi, 2015, p.642), which is a popular adequacy condition for Bayesian confirmation

Table 1 Confirmation Measures and Stopping Rules

Confirmation Measures	Non-informative Stopping Rule	Informative Stopping Rule
Difference measure d	✗	✓
Log-ratio measure r	✗	✓
Log-likelihood measure l	✗	✓
Carnap measure τ	Sometimes ✓	Sometimes ✓
	Sometimes ✗	Sometimes ✗
Normalized difference measure n	Sometimes ✓	Sometimes ✓
	Sometimes ✗	Sometimes ✗
Mortimer measure m	Sometimes ✓	Sometimes ✓
	Sometimes ✗	Sometimes ✗

measures, and all objective confirmation measures we mentioned above satisfy this requirement, whereas those non-objective ones do not (Brössel, 2013, Observation 3). For convenience, we reformulate the requirement of Final Probability as follow.

Requirement 4 (Final Probability, FP for short)
For any hypothesis H and pieces of evidence E_1 and E_2, $c(H, E_1) \gtreqless c(H, E_2)$ iff $Pr(H|E_1) \gtreqless Pr(H|E_2)$.

This observation fosters a key question about the logical relationship between FP and the above objectivity requirement: is FP a sufficient condition or a necessary condition for objectivity? We can prove the following theorem.

Theorem 5 *If a confirmation measure c satisfies FP, then it is objective.*

Proof Let H, E be a hypothesis and a piece of evidence respectively, and suppose c satisfies FP (Requirement 4). We need to show that c satisfies the two conditions (1) and (2) in Definition 3.

(1) We show c satisfies Definition 3(1). Let R be an arbitrary non-informative stopping rule and R_E be a stopping event. According to Definition 2, we only need to show $c(H, E) = c(H, E\&R_E)$. Since R is non-informative, by Definition 1 we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$. Suppose that $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E) = 0$. Then we have $Pr(R_E\&H\&E) = Pr(R_E\&\neg H\&E) = 0$, and then $Pr(R_E\&E) = 0$. However, E is the evidence, and R_E is the stopping event, so $Pr(R_E\&E) \neq 0$. Contradiction. So $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E) \neq 0$. As c satisfies Requirement 4, we have $c(H, E) = c(H, E\&R_E)$ iff $Pr(H|E) = Pr(H|E\&R_E)$. So we only need to

prove that $Pr(H|E) = Pr(H|E\&R_E)$:

$$\begin{aligned} Pr(H|E\&R_E) &= \frac{Pr(R_E|H\&E) \times Pr(H\&E)}{Pr(R_E|H\&E) \times Pr(H\&E) + Pr(R_E|\neg H\&E) \times Pr(\neg H\&E)} \\ &= \frac{Pr(R_E|H\&E) \times Pr(H\&E)}{Pr(R_E|H\&E) \times Pr(H\&E) + Pr(R_E|H\&E) \times Pr(\neg H\&E)} \\ &= \frac{Pr(R_E|H\&E) \times Pr(H\&E)}{Pr(R_E|H\&E) \times (Pr(H\&E) + Pr(\neg H\&E))} \\ &= \frac{Pr(H\&E)}{Pr(E)} \\ &= Pr(H|E) \end{aligned}$$

Since R and E are arbitrary, c satisfies Definition 3(1).

(2) We show c satisfies Definition 3(2). Let R be an arbitrary informative stopping rule and R_E be a stopping event. According to Definition 2, we only need to show $c(H, E) \neq c(H, E\&R_E)$. Since R is informative, by Definition 1 we have $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. As c satisfies Requirement 4, we have $c(H, E) = c(H, E\&R_E)$ iff $Pr(H|E) = Pr(H|E\&R_E)$. Thus, we only need to prove that $Pr(H|E) \neq Pr(H|E\&R_E)$. There are two possibilities as follows.

(i) Suppose $Pr(R_E|H\&E) \neq 0$. Then we have

$$\begin{aligned} Pr(H|E\&R_E) &= \frac{Pr(R_E|H\&E) \times Pr(H\&E)}{Pr(R_E|H\&E) \times Pr(H\&E) + Pr(R_E|\neg H\&E) \times Pr(\neg H\&E)} \\ &= \frac{Pr(H\&E)}{Pr(H\&E) + \frac{Pr(R_E|\neg H\&E)}{Pr(R_E|H\&E)} \times Pr(\neg H\&E)}. \end{aligned}$$

Since E is the evidence, we have $Pr(E) \neq 0$. As E is non-conclusive evidence for the hypothesis H, we have $Pr(\neg H|E) \neq 0$ (please see the beginning of this section). Thus, $Pr(\neg H\&E) = Pr(\neg H|E) \times Pr(E) \neq 0$. Since $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$, we have $\frac{Pr(R_E|\neg H\&E)}{Pr(R_E|H\&E)} \neq 1$. Hence,

$$\begin{aligned} Pr(H|E\&R_E) &= \frac{Pr(H\&E)}{Pr(H\&E) + \frac{Pr(R_E|\neg H\&E)}{Pr(R_E|H\&E)} \times Pr(\neg H\&E)} \\ &\neq \frac{Pr(H\&E)}{Pr(H\&E) + Pr(\neg H\&E)} \\ &= Pr(H|E). \end{aligned}$$

(ii) Suppose $Pr(R_E|H\&E) = 0$. Then we have

$$\begin{aligned} Pr(H|E\&R_E) &= \frac{Pr(R_E|H\&E) \times Pr(H\&E)}{Pr(R_E|H\&E) \times Pr(H\&E) + Pr(R_E|\neg H\&E) \times Pr(\neg H\&E)} \\ &= 0. \end{aligned}$$

However, since E is non-conclusive evidence for the hypothesis H, we have $Pr(H|E) \neq 0$ (please see the beginning of this section). Thus, we have $Pr(H|E\&R_E) \neq Pr(H|E)$.

Hence, we have $Pr(H|E) \neq Pr(H|E\&R_E)$ in both of the two possibilities. According to Requirement 4, we have $c(H, E) \neq c(H, E\&R_E)$. Since R and E are arbitrary, c satisfies Definition 3(2).

Therefore, according to Definition 3, the confirmation measure c is objective. □

The result that all confirmation measures satisfying FP are objective is important. We can use this result to derive directly that other confirmation measures are objective, so long as they satisfy FP. For example, the confirmation measures in Crupi et al. (2013); Brössel (2013).

Now consider the question: is FP a necessary condition for objectivity? Let the confirmation measure c be objective. We need to show that it satisfies Requirement 4, namely $c(H, E_1) \gtreqless c(H, E_2)$ iff $Pr(H|E_1) \gtreqless Pr(H|E_2)$ (E_1 and E_2 are arbitrary). Firstly, according to Definitions 1, 2 and 3, the concept of objectivity involves a stopping event R_E, whereas E_1, E_2 in Requirement 4 are arbitrary and may not involve any stopping event. Secondly, the two pieces of evidence E_1, E_2 in Requirement 4 are arbitrary and may be not related to each other, whereas the two pieces of evidence E, $E\&R_E$ involved in the concept of objectivity (Definitions 1, 2 and 3) are closely related to each other. Thus, we cannot derive FP from objectivity, and objectivity should be regarded as a candidate adequacy condition of its own. Someone may turn to ask: why can we prove Theorem 5? The reason is that: we can replace the two pieces of evidence E_1, E_2 in Requirement 4 with E, $E\&R_E$ respectively to obtain $c(H, E) = c(H, E\&R_E)$ iff $Pr(H|E) = Pr(H|E\&R_E)$, and then we can use such an equivalence to prove the objectivity.

Before we leave this section, we need to clarify some doubts. Someone may think that for confirmation measures τ, n, and m, non-informative stopping rules just result in linear scaling of their values, in proportion to the probability of stopping, it is unclear why we should not just re-normalize the confirmation measure in proportion to this probability of stopping. For example, for Carnap measure τ, when the probability of stopping according to the non-informative stopping rule R based on E is i, $\tau(H, E\&R_E) = i(\tau(H, E))$. So we just need a re-normalized confirmation measure $\tau^*(H, E) = \frac{1}{i} \cdot \tau(H, E)$ to represent the degree of confirmation based on the compositive evidence $E\&R_E$. The confirmation measure τ can be considered unaffected by the non-informative stopping rule R if $\tau^*(H, E\&R_E) = \tau(H, E)$. Here we have $\tau^*(H, E\&R_E) = \frac{1}{i} \cdot \tau(H, E\&R_E) = \tau(H, E)$. So Carnap measure τ is unaffected by the non-informative stopping rule if we admit such a re-normalization. However, we find that this approach is untenable. For different non-informative stopping rules, the probability of stopping based on the same evidence is highly likely different. That means we cannot find a unique re-normalized confirmation measure τ^* to satisfy the requirement that for *any* non-informative stopping rule R, $\tau(H, E) = \tau^*(H, E\&R_E)$. Someone may say that we can have different re-normalized confirmation measures for different non-informative stopping rules. For each stopping rule R, if the probability for the experiment to stop after obtaining E is x, we can let $\tau^*(H, E) = \frac{1}{x} \cdot \tau(H, E)$ to make τ to be unaffected by this stopping rule. Then here comes the second worry. Note that the re-normalization of a confirmation measure based on evidence E requires the exact value of the probability of the experiment stopping at E. But in many cases, such a probability is vague or even unknowable. For example, in chemistry experiments, it is common for several people to share a bottle of reagents when the requirement for precision is not very high. When you have completed several sets of experiments and got some data, it is quite possible that you may have to stop your experiment due to the depletion of reagents. But in such a case, it is hard, or even impossible to know the exact probability for you to stop after getting the data you have obtained. Therefore, we have reasons to believe that the approach of re-normalization is not feasible.

4 Effects of stopping rules on confirmation measures

In this section, we use two examples to illustrate the distinctions between confirmation measures with and without objectivity, and explain why confirmation measures with objectivity are better ones. Let's firstly consider the following example with a non-informative stopping rule.

Example 2 A pharmaceutical company is conducting a sequential experiment to verify the significant therapeutic effect of Drug A in treating a disease. Researcher Robert oversees the experiment to assess the effectiveness of Drug A. Let 'H' denote the hypothesis "Drug A has a significant therapeutic effect". After verifying 500 samples of patients, the experiment abruptly stops because the expensive equipment is overheating and broken. Let 'R' denote the stopping rule "the experiment stops when the equipment is broken".[2] Among the 500 patients, 400 of them have recovered (denoted by 'E'). Thus, Researcher Robert concludes that the results confirm the significant therapeutic effect of Drug A.

According to the stopping rule R, the experiment stops once the equipment is broken, no matter what the hypothesis H is and what the evidence E is. So the stopping rule R is independent of or irrelevant to the hypothesis H, and thus the stopping rule R is non-informative. According to the stopping rule R, the stopping event R_E is an event that: due to the stopping rule, the experiment stopped with the data E. Since the stopping rule R is independent of or irrelevant to the hypothesis H, the stopping rule R and the stopping event R_E, intuitively speaking, should not affect the degree of confirmation of the hypothesis H.

In the following, we use the *objective* difference measure d and the *non-objective* Carnap measure τ as representatives to verify whether they reflect such an intuition. Suppose that the prior probability $Pr(H) = Pr(\neg H) = 0.5$, $Pr(E|H) = 0.8$ and $Pr(E|\neg H) = 0.2$. And assume that the probability of equipment overheating and breaking in each experiment is $\frac{1}{1000}$, denoted as $Pr(i)$. Since the stopping rule R is non-informative, we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$. Since the stopping rule is that the experiment stops when the equipment is broken, their values depend

[2] Note that, this kind of stopping rules are common, ordinary or usual, researchers usually have to stop the experiments when the expensive equipment are broken, or the remaining samples are mistakenly destroyed, etc.

upon $Pr(i)$, i.e., $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E) = Pr(i)$. Now let's first use the difference measure d to calculate the degree of confirmation of the hypothesis H before and after integrating the stopping event R_E into the evidence. Before integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H can be calculated as follows.

$$\begin{aligned} d(H, E) &= Pr(H|E) - Pr(H) \\ &= \frac{Pr(E|H) \cdot Pr(H)}{Pr(E|H) \cdot Pr(H) + Pr(E|\neg H) \cdot Pr(\neg H)} - Pr(H) \\ &= \frac{0.8 \times 0.5}{0.8 \times 0.5 + 0.2 \times 0.5} - 0.5 \\ &= 0.3 \end{aligned}$$

The result means that the data E provides some support to H. After integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H is calculated as follows:

$$\begin{aligned} d(H, E\&R_E) &= Pr(H|E\&R_E) - Pr(H) \\ &= \frac{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H)}{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) + Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H)} - Pr(H) \\ &= \frac{Pr(E|H) \cdot Pr(H)}{Pr(E|H) \cdot Pr(H) + Pr(E|\neg H) \cdot Pr(\neg H)} - Pr(H) \\ &= Pr(H|E) - Pr(H) \\ &= 0.3 \end{aligned}$$

As we can see, before and after integrating the stopping event R_E into the evidence, the degree of confirmation calculated by the difference measure d remains the same. This means that the support provided by the evidence $E\&R_E$ is the same as that provided by the evidence E alone. Hence, it reflects our intuition that the non-informative stopping rule should not affect the degree of confirmation of the hypothesis, and aligns with Theorem 1 in Section 3.

Secondly, let's use Carnap measure τ to calculate the degree of confirmation of the hypothesis H before and after integrating the stopping event R_E into the evidence. Before integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H is:

$$\begin{aligned} \tau(H, E) &= Pr(H\&E) - Pr(H) \cdot Pr(E) \\ &= Pr(E|H) \cdot Pr(H) - Pr(H) \cdot (Pr(E|\neg H) \cdot Pr(\neg H) + Pr(E|H) \cdot Pr(H)) \\ &= 0.8 \times 0.5 - 0.5 \times (0.2 \times 0.5 + 0.8 \times 0.5) \\ &= 0.15 \end{aligned}$$

This result means that the data E provides some support to H. After integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H is:

$$
\begin{aligned}
\tau(H, E\&R_E) &= Pr(H\&E\&R_E) - Pr(H) \cdot Pr(E\&R_E) \\
&= Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) - Pr(H) \cdot (Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) \\
&\quad + Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H)) \\
&= Pr(i) \cdot (Pr(E|H) \cdot Pr(H) - Pr(H) \cdot Pr(E|H) \cdot Pr(H) - Pr(H) \cdot Pr(E|\neg H) \cdot Pr(\neg H)) \\
&= Pr(i) \cdot (Pr(H\&E) - Pr(H) \cdot Pr(E)) \\
&= Pr(i) \cdot (\tau(H, E)) \\
&= \frac{1}{1000} \times 0.15 \\
&= 1.5 \times 10^{-4}
\end{aligned}
$$

It is clear that, under the non-objective Carnap measure τ, the degree of confirmation for the hypothesis H provided by $E\&R_E$ is close to 0. This is significantly lower than the confirmation degree for the hypothesis H provided merely by E (0.15). It means that, although E does support H, $E\&R_E$ hardly provides any support to H. Hence, the non-informative stopping rule R substantially weakens the hypothesis H based on the evidence E. This result contradicts our intuition and expectation that the non-informative stopping rule should not affect the degree of confirmation of the hypothesis. In contrast, the objective confirmation measure d does not face such a problem. Hence, according to the above discussions, the objective difference measure d can be considered better than the non-objective Carnap measure τ.

Now let's consider the following example with an informative stopping rule.

Example 3 Researcher Hank is studying the quake phenomenon on a planet X. After some observations, he makes a hypothesis H: the maximum number of those aftershocks, which exceed magnitude 5, of magnitude 7 quakes $m \leq 10$.[3] Now a magnitude 7 quake occurs in a region of the planet X. Hank begins to monitor the number of aftershocks exceeding magnitude 5. The stopping rule R is that: if the number of aftershocks exceeding magnitude 5 (denoted by 'n') doesn't increase for two weeks, then stop the monitoring.[4] Finally, the number of aftershocks observed in the monitor $n = 8$ (denoted by 'E'), and the experiment stops. Thus, Hank believes that this result supports the hypothesis H.

According to the stopping rule R, the stopping event R_E is an event that: due to the stopping rule, the monitoring stopped with the data E. Since both of the hypothesis H and the stopping rule R are about the number of aftershocks, they seem to be related/relevant to each other. Moreover, intuitively speaking, $E\&R_E$ provides more

[3] In seismology, some researchers distinguish different types of quakes based on the number and magnitude of aftershocks, please see Scholz (2002) (p. 224).

[4] Suppose that, there will be no more aftershocks exceeding magnitude 5 when n doesn't increase for two weeks, and it is meaningless to continue the monitoring.

evidential support to H ($m \leq 10$) than just E; because E means just that the number of aftershocks $n = 8$, whereas $E\&R_E$ means that *the data* $n = 8$ *and the monitoring stopped*. It means that R_E affects the evaluation of the hypothesis H. So R_E and R is relevant to or not independent of the hypothesis H. Thus, in intuition, the stopping rule R is an informative stopping rule. Since $E\&R_E$ provides more evidential support to H than just E, the stopping rule R and the stopping event R_E, intuitively speaking, affect the degree of confirmation of the hypothesis H.

In the following, we use the *objective* difference measure d and the *non-objective* Carnap measure τ as representatives to verify whether they reflect such an intuition. Before that, let's suppose that: the prior probability $Pr(H) = Pr(\neg H) = 0.5$; if H is true (i.e., $m \leq 10$), then it is likely to obtain the result E (i.e., $n = 8$), and thus let $Pr(E|H) = 0.7$; and if $\neg H$ is true (i.e., $m > 10$), then it is not likely to obtain the result E (i.e., $n = 8$), and thus let $Pr(E|\neg H) = 0.2$. Since given that H is true it is likely to obtain the result E, the monitoring is likely to stop with E, so let $Pr(R_E|H\&E) = 0.8$. Since given that $\neg H$ is true it is not likely to obtain the result E, the monitoring is not likely to stop with E, so let $Pr(R_E|\neg H\&E) = 0.3$.

Now let's firstly use the difference measure d to calculate the degree of confirmation of the hypothesis H before and after integrating the stopping event R_E into the evidence. Before integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H can be calculated as follows.

$$
\begin{aligned}
d(H, E) &= Pr(H|E) - Pr(H) \\
&= \frac{Pr(E|H) \cdot Pr(H)}{Pr(E|H) \cdot Pr(H) + Pr(E|\neg H) \cdot Pr(\neg H)} - Pr(H) \\
&= \frac{0.7 \times 0.5}{0.7 \times 0.5 + 0.2 \times 0.5} - 0.5 \\
&\approx 0.2778
\end{aligned}
$$

This result means that the data E provides some support to the hypothesis H. After integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H is calculated as follows:

$$
\begin{aligned}
d(H, E\&R_E) &= Pr(H|E\&R_E) - Pr(H) \\
&= \frac{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H)}{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) + Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H)} - Pr(H) \\
&= \frac{0.8 \times 0.7 \times 0.5}{0.8 \times 0.7 \times 0.5 + 0.3 \times 0.2 \times 0.5} - 0.5 \\
&\approx 0.4032
\end{aligned}
$$

As we can see, after integrating the stopping event R_E into the evidence, the degree of confirmation calculated by the difference measure d increases. Hence, it reflects our intuitions that: $E\&R_E$ provides more evidential support to H than just E, and the informative stopping rule R should affect the degree of confirmation of the hypothesis H; and it also aligns with Theorem 1 in Section 3.

Secondly, let's use Carnap measure τ to calculate the degree of confirmation of the hypothesis H before and after integrating the stopping event R_E into the evidence. Before integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H is:

$$\begin{aligned}\tau(H,E) &= Pr(H\&E) - Pr(H)\cdot Pr(E)\\ &= Pr(E|H)\cdot Pr(H) - Pr(H)\cdot(Pr(E|\neg H)\cdot Pr(\neg H) + Pr(E|H)\cdot Pr(H))\\ &= 0.7\times 0.5 - 0.5\times(0.2\times 0.5 + 0.7\times 0.5)\\ &= 0.125\end{aligned}$$

This result means that the data E provides some support to the hypothesis H. After integrating the stopping event R_E into the evidence, the degree of confirmation for the hypothesis H is:

$$\begin{aligned}\tau(H,E\&R_E) &= Pr(H\&E\&R_E) - Pr(H)\cdot Pr(E\&R_E)\\ &= Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H) - Pr(H)\cdot(Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)\\ &\quad + Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H)\cdot Pr(\neg H))\\ &= 0.8\times 0.7\times 0.5 - 0.5\times(0.8\times 0.7\times 0.5 + 0.3\times 0.2\times 0.5)\\ &= 0.125\end{aligned}$$

It is clear that, under the non-objective Carnap measure τ, the degree of confirmation for the hypothesis H provided by $E\&R_E$ is the same as that provided merely by E. This result contradicts our intuitions that: $E\&R_E$ provides more evidential support to H than just E, and the informative stopping rule R should affect the degree of confirmation of the hypothesis H. In contrast, the objective confirmation measure d does not face such a problem. Hence, the objective difference measure d can be considered better than the non-objective Carnap measure τ. To sum up, according to the above discussions, it is reasonable to believe that confirmation measures with objectivity are better than those without objectivity.

5 Conclusion

This article mainly investigates whether six common Bayesian confirmation measures d, r, l, τ, m, and n are affected by stopping rules, and whether they are objective for Bayesians. We have proved that, difference measure d, log-ratio measure r, and log-likelihood measure l are unaffected by non-informative stopping rules, but affected by informative stopping rules. In contrast, Carnap measure τ, normalized difference measure n, and Mortimer measure m are affected by (non-)informative stopping rules sometimes and sometimes are not. Thus, according to the concept of objectivity, the confirmation measures d, r and l are objective, while τ, n and m are not objective. Based on the dispute between Frequentists and Bayesians about non-informative stopping rules, as well as the requirement related to informative stopping rules, the objective measures are considered better than the non-objective ones. Besides, we also provide two examples to further illustrate that the confirmation measures with objectivity are better than those without objectivity.

Appendix A: Omitted proofs

Proof for Theorem 2

Proof We first show that the log-ratio measure r is unaffected by non-informative stopping rule but affected by informative stopping rule. We first calculate the degree of confirmation of hypothesis H based on the compositive evidence $E\&R_E$. According to Bayes' theorem, we have:

$$
\begin{aligned}
r(H, E\&R_E) &= log[\frac{Pr(H|E\&R_E)}{Pr(H)}] \\
&= log[\frac{Pr(E\&R_E|H) \cdot Pr(H)}{Pr(E\&R_E) \cdot Pr(H)}] \\
&= log[\frac{Pr(E\&R_E|H)}{Pr(E\&R_E)}] \\
&= log[\frac{Pr(E\&R_E|H)}{Pr(E\&R_E|H) \cdot Pr(H) + Pr(E\&R_E|\neg H) \cdot Pr(\neg H)}] \\
&= log[\frac{Pr(R_E|H\&E) \cdot Pr(E|H)}{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) + Pr(R_E|E\&\neg H) \cdot Pr(E|\neg H) \cdot Pr(\neg H)}]
\end{aligned}
$$

In the following, let's analyze the impact of stopping rules (informative and non-informative) on the log-ratio measure r, respectively.

(1) Let R be a non-informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E) > 0$. Hence, we have:

$$
\begin{aligned}
r(H, E\&R_E) &= log[\frac{Pr(R_E|H\&E) \cdot Pr(E|H)}{Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) + Pr(R_E|E\&\neg H) \cdot Pr(E|\neg H) \cdot Pr(\neg H)}] \\
&= log[\frac{Pr(E|H)}{Pr(E|H) \cdot Pr(H) + Pr(E|\neg H) \cdot Pr(\neg H)}] \\
&= log[\frac{Pr(E|H) \cdot \frac{Pr(H)}{Pr(E)}}{Pr(E|H) \cdot Pr(H) \cdot \frac{Pr(H)}{Pr(E)} + Pr(E|\neg H) \cdot Pr(\neg H) \cdot \frac{Pr(H)}{Pr(E)}}] \\
&= log[\frac{Pr(H|E)}{Pr(H) \cdot (Pr(H|E) + Pr(\neg H|E))}] \\
&= log[\frac{Pr(H|E)}{Pr(H)}]
\end{aligned}
$$

Since E is non-conclusive, we have $Pr(H|E) > 0$. Hence $r(H, E) = log[\frac{Pr(H|E)}{Pr(H)}] = r(H, E\&R_E)$. Thus, according to Definition 2, non-informative stopping rule does not affect log-ratio measure r.

(2) Let R be an informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. Suppose that $Pr(R_E|\neg H\&E) = \alpha$, $Pr(R_E|H\&E) = \beta$, and $\alpha \neq \beta$. Since all scientific experiments will eventually stop, it is reasonable to assume that at least one of $Pr(R_E|H\&E)$ and

$Pr(R_E|\neg H\&E)$ is greater than 0. First assume that $\alpha > 0$ and $\beta > 0$. Then we have:

$$
\begin{aligned}
r(H, E\&R_E) &= log[\frac{Pr(R_E|H\&E)\cdot Pr(E|H)}{Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H) + Pr(R_E|E\&\neg H)\cdot Pr(E|\neg H)\cdot Pr(\neg H)}] \\
&= log[\frac{\beta\cdot Pr(E|H)}{\beta\cdot Pr(E|H)\cdot Pr(H) + \alpha\cdot Pr(E|\neg H)\cdot Pr(\neg H)}] \\
&= log[\frac{Pr(E|H)}{Pr(E|H)\cdot Pr(H) + \frac{\alpha}{\beta}\cdot Pr(E|\neg H)\cdot Pr(\neg H)}] \\
&= log[\frac{Pr(E|H)\cdot\frac{Pr(H)}{Pr(E)}}{Pr(E|H)\cdot Pr(H)\cdot\frac{Pr(H)}{Pr(E)} + \frac{\alpha}{\beta}\cdot Pr(E|\neg H)\cdot Pr(\neg H)\cdot\frac{Pr(H)}{Pr(E)}}] \\
&= log[\frac{Pr(H|E)}{Pr(H)\cdot(Pr(H|E) + \frac{\alpha}{\beta}\cdot Pr(\neg H|E))}]
\end{aligned}
$$

Since $\alpha \neq \beta$, we have the following two situations. When $\alpha > \beta$, we can derive:

$$
\begin{aligned}
r(H, E\&R_E) &= log[\frac{Pr(H|E)}{Pr(H)\cdot(Pr(H|E) + \frac{\alpha}{\beta}\cdot Pr(\neg H|E))}] \\
&< log[\frac{Pr(H|E)}{Pr(H)\cdot(Pr(H|E) + Pr(\neg H|E))}] \\
&= r(H, E).
\end{aligned}
$$

Then, according to Definition 2, the informative stopping rule R weakens the hypothesis H. When $\alpha < \beta$, we have:

$$
\begin{aligned}
r(H, E\&R_E) &= log[\frac{Pr(H|E)}{Pr(H)\cdot(Pr(H|E) + \frac{\alpha}{\beta}\cdot Pr(\neg H|E))}] \\
&> log[\frac{Pr(H|E)}{Pr(H)\cdot(Pr(H|E) + Pr(\neg H|E))}] \\
&= r(H, E).
\end{aligned}
$$

Then, according to Definition 2, the informative stopping rule R supports the hypothesis H. Now assume that $\alpha = 0$ and $\beta \neq 0$, we have:

$$
\begin{aligned}
r(H, E\&R_E) &= log[\frac{\beta\cdot Pr(E|H)}{\beta\cdot Pr(E|H)\cdot Pr(H) + \alpha\cdot Pr(E|\neg H)\cdot Pr(\neg H)}] \\
&= -log[Pr(H)]
\end{aligned}
$$

Since $r(H, E) = log[\frac{Pr(H|E)}{Pr(H)}] = log[Pr(H|E)] - Pr(H)$ and E is non-conclusive, we have $log[Pr(H|E)] < 0$. Thus $r(H, E) < r(H, E\&R_E)$. Then, according to Definition 2, the informative stopping rule supports the hypothesis H. When $\alpha \neq 0$ and $\beta = 0$, we have: $Pr(H|E\&R_E) = \frac{Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H)}{Pr(R_E|H\&E)\cdot Pr(E|H)\cdot Pr(H) + Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H)\cdot Pr(\neg H)} = 0$. Thus $r(H, E\&$

$R_E) = -\infty < r(H, E)$. According to Definition 2, the informative stopping rule weakens the hypothesis H. Then, according to Definition 2, the log-ratio measure r is affected by informative stopping rules.

Hence, according to Definition 3, the log-ratio measure r does satisfy the objective requirements.

In the following, let's consider the log-likelihood measure l to see whether it satisfies the objective requirements. What is the degree of confirmation of the hypothesis H provided by the evidence $E\&R_E$? According to Bayes' theorem, we have:

$$l(H, E\&R_E) = log[\frac{Pr(E\&R_E|H)}{Pr(E\&R_E|\neg H)}] = log[\frac{Pr(R_E|H\&E) \cdot Pr(E|H)}{Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H)}]$$

Now let's analyze the impact of stopping rules (informative and non-informative) on the log-ratio measure l, respectively.

(1) Let R be a non-informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E) > 0$. Since E is non-conclusive, we have $Pr(E|H) > 0$. Since $Pr(E\&R_E|H) = Pr(E|H) \cdot Pr(R_E|H\&E)$, we have $Pr(E\&R_E|H) > 0$. Similarly, $Pr(E\&R_E|\neg H) > 0$. Therefore, we have:

$$l(H, E\&R_E) = log[\frac{Pr(R_E|H\&E) \cdot Pr(E|H)}{Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H)}] = log[\frac{Pr(E|H)}{Pr(E|\neg H)}]$$

Since E is non-conclusive, we have $Pr(E|H) > 0$ and $Pr(E|\neg H) > 0$. Thus $l(H, E) = log[\frac{Pr(E|H)}{Pr(E|\neg H)}] = l(H, E\&R_E)$. Then, according to Definition 2, non-informative stopping rules do not affect the log-likelihood measure l.

(2) Let R be an informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. Suppose that $Pr(R_E|\neg H\&E) = \alpha$, $Pr(R_E|H\&E) = \beta$, and $\alpha \neq \beta$. Since all scientific experiments will eventually stop with some experimental data E, it is reasonable to assume that at least one of $Pr(R_E|H\&E)$ and $Pr(R_E|\neg H\&E)$ is greater than 0. First assume that $\alpha > 0$ and $\beta > 0$. Since E is non-conclusive, we have $Pr(E|H) > 0$ and $Pr(E|\neg H) > 0$. Then we have $Pr(E\&R_E|H) = Pr(E|H) \cdot \beta > 0$ and $Pr(E\&R_E|\neg H) = Pr(E|\neg H) \cdot \alpha > 0$. Therefore, we have:

$$\begin{aligned} l(H, E\&R_E) &= log[\frac{Pr(R_E|H\&E) \cdot Pr(E|H)}{Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H)}] = log[\frac{\beta \cdot Pr(E|H)}{\alpha \cdot Pr(E|\neg H)}] \\ &= log[\frac{Pr(E|H)}{\frac{\alpha}{\beta} \cdot Pr(E|\neg H)}] \end{aligned}$$

Since $\alpha \neq \beta$, we have the following two situations. When $\alpha > \beta$, we can derive:

$$l(H, E) = log[\frac{Pr(E|H)}{Pr(E|\neg H)}] > log[\frac{Pr(E|H)}{\frac{\alpha}{\beta} \cdot Pr(E|\neg H)}] = l(H, E\&R_E).$$

Then, according to Definition 2, the informative stopping rule R weakens the hypothesis H. When $\alpha < \beta$, we have:

$$l(H, E) = log[\frac{Pr(E|H)}{Pr(E|\neg H)}] < log[\frac{Pr(E|H)}{\frac{\alpha}{\beta} \cdot Pr(E|\neg H)}] = l(H, E\&R_E).$$

Then, according to Definition 2, the informative stopping rule R supports the hypothesis H. Then assume that $\alpha = 0$ or $\beta = 0$. Since $Pr(E\&R_E) = Pr(E|H)\cdot \beta + Pr(E|\neg H)\cdot\alpha$ and E is non-conclusive, $Pr(E\&R_E) > 0$. When $\alpha = 0$ and $\beta \neq 0$, we have $Pr(E\&R_E|\neg H) = \alpha \cdot Pr(E|\neg H) = 0$. Thus $l(H, E\&R_E) = \infty > l(H, E)$. According to Definition 2, the stopping rule supports the hypothesis H. When $\alpha \neq 0$ and $\beta = 0$, $Pr(E\&R_E|\neg H) = \alpha \cdot Pr(E|\neg H) > 0$. And $Pr(E\&R_E|H) = \beta \cdot Pr(E|H) = 0$. Thus $l(H, E\&R_E) = -\infty < l(H, E)$. According to Definition 2, the stopping rule weakens the hypothesis H. Therefore, according to Definition 2, the log-likelihood measure l is affected by informative stopping rules.

Hence, according to Definition 3, the log-likelihood measure l does satisfy the objective requirements. □

Proof for Theorem 4

Proof We first show that the normalized difference measure n is affected by stopping rules sometimes and sometimes is not. We first calculate the degree of confirmation based on the compositive evidence $E\&R_E$? According to Bayes' theorem, we have:

$$\begin{aligned} n(H, E\&R_E) &= Pr(E\&R_E|H) - Pr(E\&R_E|\neg H) \\ &= Pr(R_E|H\&E)\cdot Pr(E|H) - Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H) \end{aligned}$$

In the following, let's analyze whether (informative and non-informative) stopping rules will affect the normalized difference measure n, respectively.

(1) Let R be a non-informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$. Let's denote their values as $i(i > 0)$. Then we have:

$$\begin{aligned} n(H, E\&R_E) &= Pr(R_E|H\&E)\cdot Pr(E|H) - Pr(R_E|\neg H\&E)\cdot Pr(E|\neg H) \\ &= i(Pr(E|H) - Pr(E|\neg H)) \end{aligned}$$

When $i = 1$, we have $n(H, E\&R_E) = n(H, E)$. Then, according to Definition 2, the normalized difference measure n is not affected by the non-informative stopping rule R. Nevertheless, when $i \neq 1$, $n(H, E\&R_E) \neq n(H, E)$, and then, according to Definition 2, the normalized difference measure n is affected by the non-informative stopping rule R.

(2) Let R be an informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. Let's assume that $Pr(R_E|\neg H\&E) = \alpha$, $Pr(R_E|H\&E) = \beta$ $(\alpha \neq \beta)$, $Pr(E|H) = \gamma$ $(0 < \gamma < 1)$, $Pr(E|\neg H) = \delta$ $(0 < \delta < 1)$ and $Pr(H) = \epsilon$ $(0 < \epsilon < 1)$, then we have:

$$
\begin{aligned}
n(H, E\&R) &= Pr(R_E|H\&E) \cdot Pr(E|H) - Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \\
&= \beta \cdot \gamma - \alpha \cdot \delta \\
n(H, E) &= Pr(E|H) - Pr(E|\neg H) \\
&= \gamma - \delta
\end{aligned}
$$

Obviously, if $\beta = \frac{\alpha\cdot\delta+\gamma-\delta}{\gamma}$, then $n(H, E\&R_E) = n(H, E)$. According to Definition 2, when $\beta = \frac{\alpha\cdot\delta+\gamma-\delta}{\gamma}$, the normalized difference measure n is unaffected by the informative stopping rule R. In contrast, when $\beta \neq \frac{\alpha\cdot\delta+\gamma-\delta}{\gamma}$, the normalized difference measure n is affected by the informative stopping rule R.

Hence, according to Definition 3, the normalized difference measure n does not satisfy the objective requirements for some situations.

Finally, let's consider Mortimer measure m to see whether it satisfies the objective requirements. What is the degree of confirmation of the hypothesis H provided by the compositive evidence $E\&R_E$? According to Bayes' theorem, we have:

$$
\begin{aligned}
m(H, E\&R_E) =&Pr(E\&R_E|H) - Pr(E\&R_E) \\
=&Pr(R_E|H\&E) \cdot Pr(E|H) - Pr(E\&R_E|H) \cdot Pr(H) - Pr(E\&R_E|\neg H) \cdot Pr(\neg H) \\
=&Pr(R_E|H\&E) \cdot Pr(E|H) - Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) \\
&- Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H)
\end{aligned}
$$

In the following, let's analyze whether (informative and non-informative) stopping rules will affect the Mortimer measure m, respectively.

(1) Let R be a non-informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) = Pr(R_E|\neg H\&E)$. Let's denote their values as i $(i > 0)$. So we have:

$$
\begin{aligned}
m(H, E\&R_E) =&Pr(R_E|H\&E) \cdot Pr(E|H) - Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) \\
&- Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H) \\
=&i(Pr(E|H) - Pr(H\&E) - Pr(E\&\neg H)) \\
=&i(Pr(E|H) - Pr(E))
\end{aligned}
$$

When $i = 1$, we have $m(H, E\&R_E) = m(H, E)$. Then, according to Definition 2, Mortimer measure m is not affected by the non-informative stopping rule R. Nevertheless, when $i \neq 1$, $m(H, E\&R_E) \neq m(H, E)$, and then, according to Definition 2, Mortimer measure m is affected by the non-informative stopping rule R.

(2) Let R be an informative stopping rule. Then, according to Definition 1, we have $Pr(R_E|H\&E) \neq Pr(R_E|\neg H\&E)$. Let's assume that $Pr(R_E|\neg H\&E) = \alpha$, $Pr(R_E|H\&E) = \beta$ $(\alpha \neq \beta)$, $Pr(E|H) = \gamma$ $(0 < \gamma < 1)$, $Pr(E|\neg H) = \delta$ $(0 < \delta < 1)$ and $Pr(H) = \epsilon$ $(0 < \epsilon < 1)$, then we have:

$$\begin{aligned}
m(H, E\&R_E) =& Pr(R_E|H\&E) \cdot Pr(E|H) - Pr(R_E|H\&E) \cdot Pr(E|H) \cdot Pr(H) \\
&- Pr(R_E|\neg H\&E) \cdot Pr(E|\neg H) \cdot Pr(\neg H) \\
=& (\beta \cdot \gamma - \alpha \cdot \delta) \cdot (1 - \epsilon) \\
m(H, E) =& Pr(E|H) - Pr(H\&E) - Pr(E\&\neg H) \\
=& Pr(E|H) - Pr(E|H) \cdot Pr(H) - Pr(E|\neg H) \cdot Pr(\neg H) \\
=& (\gamma - \delta) \cdot (1 - \epsilon)
\end{aligned}$$

Clearly, if $\beta = \frac{\alpha \cdot \delta + \gamma - \delta}{\gamma}$, then $m(H, E\&R_E) = m(H, E)$. According to Definition 2, when $\beta = \frac{\alpha \cdot \delta + \gamma - \delta}{\gamma}$, Mortimer measure m is unaffected by the informative stopping rule R. In contrast, when $\beta \neq \frac{\alpha \cdot \delta + \gamma - \delta}{\gamma}$, Mortimer measure m is affected by the informative stopping rule R.

Hence, according to Definition 3, the Mortimer measure m does not satisfy the objective requirements for some situations. □

Acknowledgements Thanks to Zizhao Yuan for recommending an article about Earth science to support Example 3. Thanks to Charlie Siu for suggesting moving some proofs in Section 3 to the appendix. And thanks to the three anonymous reviewers of this journal for their valuable suggestions.

Funding This work was supported by the major project of the National Social Science Fund of China [23&ZD240].

Declarations

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London,* (53), 370–418.

Berger, J. O., & Berry, D. A. (1988). Statistical analysis and the illusion of objectivity. *American Scientist, 76*(2), 159–165.

Brössel, P. (2013). The problem of measure sensitivity redux. *Philosophy of Science, 80*(3), 378–397.

Carnap, R. (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Christensen, D. (1999). Measuring confirmation. *The Journal of Philosophy, 96*(9), 437–461.
Crupi, V. (2015). Inductive logic. *Journal of Philosophical Logic, 44*(6), 641–650.
Crupi, V., Chater, N., & Tentori, K. (2013). New axioms for probability and likelihood ratio measures. *British Journal for the Philosophy of Science, 64*(1), 189–204.
Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology, 72*, 78–89.
Fitelson, B. (2001). *Studies in Bayesian Confirmation Theory*. Ph. D. thesis, University of Wisconsin, Madison.
Fletcher, S. C. (2019). Stopping rules as experimental design. *European Journal for Philosophy of Science, 9*(2), 1–20.
Gandenberger, G. (2017). Differences among noninformative stopping rules are often relevant to Bayesian decisions. arXiv:1707.00214
Glass, D. H. (2013). Confirmation measures of association rule interestingness. *Knowledge-Based Systems, 44*, 65–77.
Good, I. J. (1960). Weight of evidence, corroboration, explanatory power, information and the utility of experiments. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 22*(2), 319–331.
Goodman, S. N. (2005). Introduction to Bayesian methods I: measuring the strength of evidence. *Clinical Trials, 2*(4), 282–290.
Goodman, S. N., & Royall, R. (1988). Evidence and scientific research. *American Journal of Public Health, 78*(12), 1568–1574.
Jeffreys, H. (1961). *Theory of Probability*. Oxford, U.K.: Oxford University Press.
Milne, P. (1996). log [p(h/eb)/p(h/b)] is the one true measure of confirmation. *Philosophy of Science, 63*(1), 21–26.
Mortimer, H. (1988). *The Logic of Induction*. New York: Halsted Press.
Nozick, R. (1981). *Philosophical Explanations*. Cambridge, Mass.: Harvard University Press.
Roberts, H. V. (1967). Informative stopping rules and inferences about population size. *Journal of the American Statistical Association, 62*(319), 763–775.
Scholz, C. H. (2002). *The Mechanics of Earthquakes and Faulting*. Cambridge: Cambridge University Press.
Steel, D. (2003). A Bayesian way to make stopping rules matter. *Erkenntnis, 58*(2), 213–227.
Urbach, P., & Howson, C. (1993). *Scientific Reasoning: The Bayesian Approach*. Chicago: Open Court.