

(Word count: 4,998)

# Experimental Philosophy Meets Formal Epistemology<sup>1</sup>

Jonah N. Schupbach, *Philosophy, University of Utah*

**Abstract.** *Formal epistemology* is just what it sounds like: epistemology done with formal tools. Coinciding with the general rise in popularity of experimental philosophy, formal epistemologists have begun to apply experimental methods in their own work. In this entry, I survey some of the work at the intersection of formal and experimental epistemology. I show that experimental methods have unique roles to play when epistemology is done formally, and I highlight some ways in which results from formal epistemology have been used fruitfully to advance epistemically-relevant experimental work. The upshot of this brief, incomplete survey is that formal and experimental methods often constitute mutually informative means to epistemological ends.

**Keywords.** Bayesian epistemology, coherence, confirmation, conjunction fallacy, explanatory power, explication, formal epistemology, judgment aggregation, probabilistic reasoning, simplicity

---

<sup>1</sup> Thanks to Vincenzo Crupi, Igor Douven, and two anonymous reviewers for helpful feedback on earlier drafts of this chapter. Research for this entry was supported by an Aldrich Fellowship from the University of Utah's Tanner Humanities Center.

*Formal epistemology* is a flourishing subfield of philosophy that investigates classic epistemic topics – such as knowledge, belief, rationality, and justification – using logicomathematical methods. Formal epistemology is far from a new movement, tracing its roots back at least to Aristotle’s formal theory of the syllogism and his application of this theory in the characterization of *epistêmê*. Indeed, given the centrality of the notion of *rationality* to epistemology, and noting that formal logic traditionally developed as the science of *rational* inference, one may plausibly claim that the history of formal logic itself is a narrative belonging to the history of formal epistemology.

To note that formal epistemology has a venerable pedigree, however, is not to suggest that its most important contributions are ancient history. On the contrary, there has arguably never been a more exciting, productive time to be a formal epistemologist. In part, this is because of the impressive development of formal methods. As such methods evolve (and as we appreciate more and more useful ways to apply them), formal epistemologists not only gain tools that could potentially lend new insight to ongoing projects, but they are also able to approach problems using formal methods that were previously formally intractable.

This entry focuses on another reason to be excited about formal epistemology today: the budding interest in connecting it to experimental research. Often, it is only once a theory is stated precisely that its empirical implications (possibly along with the fact that it even *has* empirical implications) become clear. At the intersection of formal and experimental epistemology, philosophers use logicomathematical tools to introduce precise philosophical theory, which then enables them to apply and test that theory experimentally. How this goes in detail depends on one’s philosophical ends.

Much of the work pursued by formal epistemologists either aims at *concept engineering* or *concept clarification*. In either case, one *explicates* concepts, where explication involves “transforming a given more or less inexact concept [the *explicandum*] into an exact one [the *explicatum*]” (Carnap 1950, 3). In concept engineering, this is done so that we might improve our conceptual frameworks; we *replace* explicanda with explicata that better assist us in achieving epistemic ends. Alternatively, in concept clarification, philosophers explicate concepts to shed new light on the nature and implications of those explicanda; the goal is to illuminate rather than replace explicanda.

This distinction is especially important to keep in mind when one is deciding whether an explication is satisfactory. Carnap (1950) helpfully proposes four desiderata for explication evaluation; for our purposes, we may focus on two. First, an explicatum should retain a *similarity* to the explicandum so that it applies in most (if not all) of the cases in which the latter applies. Second, the explicatum should, in some sense, be epistemically *fruitful*, e.g., by allowing us to learn things more efficiently.<sup>2</sup> But these two desiderata can conflict; greater similarity may only be attainable at the expense of fruitfulness and vice versa. When this happens, the concept engineer prizes fruitfulness above similarity in the pursuit of improved conceptual schema. The concept clarifier instead prizes similarity, as any explication only has a chance of illuminating the explicandum insofar as its explicatum (fruitful or not) maintains the requisite resemblance.

Experiments connect to formal epistemology in different ways, depending on whether one is pursuing concept clarification or engineering. The concept clarifier may experimentally test the similarity of explicata to explicanda; such experiments potentially

---

<sup>2</sup> Carnap more stringently requires the explicatum to be *fruitful* in the sense of being “useful for the formulation of many universal statements (empirical laws [or] logical theorems [...]).”

inform – by enabling us to evaluate – work in formal epistemology. The concept engineer may experimentally explore the fruitfulness of an explicatum; and, if an explicatum proves fruitful, formal epistemological results can inform experimental research.<sup>3</sup>

This entry surveys such formal-experimental research. Section 2 discusses work in which experiments test whether formal epistemological theory maintains a substantive similarity to cognitive factors. Section 3 explores work in which formal epistemological results fruitfully advance experimental research. Specific topics covered relate to an array of epistemological sub-disciplines – including mainstream epistemology, the epistemology of science, social epistemology, the general study of human reasoning, and of course formal epistemology.

## 1. Experiments Informing Formal Epistemology

Formal epistemologists often use explication to shed new light on epistemic concepts. As noted above, it is crucial to such work that explicata maintain a similarity to explicanda. This section offers examples of experimental tests for similarity. I focus on three topics: confirmation, explanatory reasoning, and coherence.

In Bayesian ***confirmation theory***, formal epistemologists offer the following explication: evidence E confirms hypothesis H (relative to a set of background beliefs K) exactly when E raises H's probability (conditional on K):  $\Pr(H|E\&K) > \Pr(H|K)$ .<sup>4</sup> While Bayesians agree on that much, they disagree on the best explication of *the extent to which* E

---

<sup>3</sup> (Schupbach, Unpublished) provides an in-depth discussion of explication, concept engineering, concept clarification, and the bearing of experimental work on these.

<sup>4</sup> Pr is a probability function defined over a field of propositions containing H, E, and K (where K may be thought of as the conjunction of propositions collectively describing relevant background beliefs). The solidus "|" may be read as "conditional on" or "given that"; accordingly, one may read the above inequality as "the probability of H conditional on E and K is greater than the probability of H conditional on K alone."

confirms H. The question is how to measure the strength of the above inequality. Obvious candidates include a difference measure  $C_d(H,E|K)=Pr(H|E\&K)-Pr(H|K)$  and ratio measure  $C_r(H,E|K)=Pr(H|E\&K)/Pr(H|K)$ , but many others have been defended (**Table 1**).

Importantly, which of these explicata one adopts can make a crucial difference to philosophical conclusions about confirmation we draw from them; e.g., Fitelson (1999) shows that various Bayesian “resolutions” of well-known paradoxes hinge sensitively on one’s choice of measure.

<b>Table 1.</b> Some alternative Bayesian explicata for degree of confirmation	
$C_d(H,E K)=Pr(H E\&K)-Pr(H K)$	$C_s(H,E K)=Pr(H E\&K)-Pr(H \sim E\&K)$
$C_m(H,E K)=Pr(E H\&K)-Pr(E K)$	$C_n(H,E K)=Pr(E H\&K)-Pr(E \sim H\&K)$
$C_c(H,E K)=Pr(H\&E K)-Pr(H K)Pr(E K)$	$C_r(H,E K)=Pr(H E\&K)/Pr(H K)$
$C_G(H,E K)=1-[Pr(\sim H E\&K)/Pr(\sim H K)]$	$C_z(H,E K)=\begin{cases} \frac{Pr(H E\&K)-Pr(H K)}{1-Pr(H K)}, & \text{if } Pr(H E\&K)\geq Pr(H K) \\ \frac{Pr(H E\&K)-Pr(H K)}{Pr(H K)}, & \text{if } Pr(H E\&K)< Pr(H K) \end{cases}$
$C_k(H,E K)=\frac{Pr(E H\&K)-Pr(E \sim H\&K)}{Pr(E H\&K)+Pr(E \sim H\&K)}$	

Considerations of *similarity* are crucial here. Can we verify that any one of these measures is similar to the concept of confirmation applied in human cognition? Our explication will only illuminate this concept to the extent that we can.<sup>5</sup> Crupi, Tentori, and collaborators (Tentori et al. 2007; Crupi, Tentori, and Gonzalez 2007) use experiments to investigate this question. In personal interviews, participants were informed that they had before them one of two urns, urn A containing 30 black and 10 white balls, urn B containing 15 black and 25 white balls. The experimenter then began sampling balls from the urn. After each of ten draws without replacement, participants assessed the extent to which the

<sup>5</sup> This is a serious issue as some philosophers (e.g., Strawson 1963) have challenged the very idea of using formal explication for concept clarification by arguing that formal explicata cannot maintain a close fit to informal explicanda. See (Schupbach, Unpublished) for a discussion of this challenge and experimentation’s role in potentially meeting it.

samples confirm the respective hypotheses that urn A/B was the chosen urn. These data allowed Crupi and Tentori to compare theoretical degrees of confirmation derived from candidate measures to human assessments of confirmation.

The results were as follows:  $C_Z$  outperformed all other candidates in a paired *t*-test comparing the average (Pearson) correlations between derived results and participant responses; i.e.,  $C_Z$  proved to be a significantly better (linear) predictor of confirmation judgments under the experimental conditions. This supports the *comparative* conclusion that  $C_Z$  comes closer to our notion of confirmation than the other considered Bayesian measures.<sup>6</sup> There remains the question, however, of how closely  $C_Z$  – or any other measure – fits *overall* with this notion. Tentori, Crupi, and Osherson (2007) present a separate study that sheds some light on this question. In two experiments, they compare purely probabilistic approaches to confirmation in general with human judgments of confirmation. Their results support the conclusion that such judgments “depend on more than probabilities over E and H.” The implication for Bayesian confirmation theory is that, while certain measures may maintain a close resemblance to our notion of confirmation, any such explicatum will not be able to provide a perfect fit. But what would constitute a sufficiently good fit, and can any probabilistic account provide that? These lingering questions call out for further (philosophical and empirical) investigation.

Formal epistemological research on *explanatory reasoning* focuses on explicating a variety of attributes people look for in good explanations. Several of these “explanatory virtues” relate an explanatory hypothesis to what is being explained (e.g., unification, power, and consilience), while others are properties of the hypothesis alone (e.g., “internal”

---

<sup>6</sup> Rusconi et al. (Forthcoming) present a series of experiments which lend further comparative support to  $C_Z$  over a variety of (Bayesian and non-Bayesian) explications of “the value of evidence.”

versions of plausibility, coherence, and simplicity). In any case, formal epistemologists attempt to illuminate, by explicating, these virtues.

The study of the virtue of *power* parallels Bayesian work on confirmation. A hypothesis *H* has power over an explanandum *E* to the extent that it makes *E* less surprising (or more expected). Probabilistically, power is the flipside of confirmation; confirmation is measured as the extent to which *H* is more probable in light of *E*, whereas power is measured as the extent to which *E* is more probable given *H*. Accordingly, as with confirmation, various Bayesian measures of power have been proposed (**Table 2**).

<b>Table 2.</b> Some alternative Bayesian explicata for power	
$E_d(E,H K) = \Pr(E H\&K) - \Pr(E K)$	$E_c(E,H K) = \Pr(E H\&K) - \Pr(E \sim H\&K)$
$E_p(E,H K) = \frac{\Pr(E H\&K) - \Pr(E K)}{\Pr(E H\&K) + \Pr(E K)}$	$E_m(E,H K) = \Pr(E H\&K) / \Pr(E K)$
$E_s(E,H K) = \frac{\Pr(H E\&K) - \Pr(H \sim E\&K)}{\Pr(H E\&K) + \Pr(H \sim E\&K)}$	$E_z(E,H K) = \begin{cases} \frac{\Pr(E H\&K) - \Pr(E K)}{1 - \Pr(E K)}, & \text{if } \Pr(E H\&K) \geq \Pr(E K) \\ \frac{\Pr(E H\&K) - \Pr(E K)}{\Pr(E K)}, & \text{if } \Pr(E H\&K) < \Pr(E K) \end{cases}$

To evaluate how similar these alternatives are to our intuitive notion of power, Schupbach (2011) asks participants for assessments of explanatoriness in experimental settings based closely on Tentori et al.’s ball-and-urn design (e.g., how well does the hypothesis that urn A was chosen explain the results sampled so far?). Two analyses support the conclusion that  $E_s$  “comes significantly closer to participant judgments than any other candidate measure.” Regarding the question of overall fit, the study reveals that  $E_s$ ’s results not only sit quite closely, on average, to participant judgments, but their mean is also centered on that of participant judgments (other measures’ results are biased, on average, to one side of these judgments).

While the virtue of power relates hypotheses to explananda, *simplicity* is predicated monadically to hypotheses alone. Experiments by Lombrozo (2007) focus on a notion of simplicity formalized as number of causes cited by a hypothesis. Her research broaches two questions: First, is this formal idea recognized as an explanatory virtue in actual human reasoning (the similarity question)? Second, how do people's judgments of simplicity relate to their judgments of probability?

Participants are asked to diagnose an alien patient's two symptoms by appealing either to a single disease X or to a combination of two other diseases, Y&Z. Participants are told that X "always causes both" symptoms, while Y and Z jointly (but not individually) account fully for both symptoms. With this information, an overwhelming majority of participants opt for X over Y&Z as the best diagnosis, suggesting that the number of causes cited does indeed track a perceived explanatory virtue.

Variations on this experiment explore the relation between explanatory and probability judgments. For example, when given the above information along with base rates revealing that Y&Z are together more prevalent in the alien population than X, participants still often favor the simpler diagnosis. Lombrozo notes that these findings are consistent with the idea that people grant simple hypotheses higher prior probabilities, which they subsequently update appropriately in light of base rate information. Bonawitz and Lombrozo (2012) replicate these same findings with pre-school aged children.<sup>7</sup>

Experiments like the above prove useful in testing formal accounts of explanatory virtues. And by revealing contexts in which explanatory reasoning tracks formal explicata, this work potentially justifies the drawing of normative conclusions about human

---

<sup>7</sup> For another recent study in this vein, see Pacer et al.'s (2013) investigation into whether four computational models of overall explanatoriness "correspond to general human intuitions."



reasoning from analyses of formal correlates. Bayesian explications of power (**Table 2**), for example, imply that the more power a hypothesis has over a given explanandum, the better confirmed it is by that explanandum (using any Bayesian measure of confirmation). Insofar as experiments reveal contexts in which explanatory reasoning fits closely with one of these explicata then, we gain a neat account of why explanatory reasoning may be good in those contexts: because it reliably tracks probabilistic confirmation relations.<sup>8</sup>

Epistemologists speak of *coherence* vaguely as the property things have to the extent that they fit together, dovetail, or agree. Coherence can plausibly be thought of as another explanatory virtue; the more coherent an explanation, the better. But people also assess beliefs, norms, choices, and values to be more or less coherent; so coherence may for example underlie practical reasoning (Millgram and Thagard 1996). Bayesians attempt probabilistic explications of coherence, and recent experimental work investigates whether these accounts connect to human cognition (Harris and Hahn 2009). However, there has been far more experimental work testing Thagard's computational account of coherence.

This explication associates coherence-based reasoning with constraint satisfaction (Thagard 2000). In any "coherence problem," we hunt for an "accepted" subset of a set of elements  $E = \{e_1, e_2, \dots, e_n\}$  which best satisfies a given set of constraints. These constraints may link elements in a positive way (e.g., if  $e_j$  implies  $e_k$ , a positive constraint may require that we accept both), or in a negative way (e.g., if  $e_j$  is inconsistent with  $e_k$ , a negative constraint may require that we accept only one). Each particular constraint is assigned a weight. Thagard explicates coherence as  $W$ , the sum of the weights corresponding to the

---

<sup>8</sup> The reader may wish to consult (Lombrozo, this volume) for more information regarding formal-experimental work on explanation.

constraints that get satisfied by a particular partition of  $E$ ; thus, maximizing coherence amounts to partitioning  $E$  into two sets (accepted and rejected) in a way that maximizes  $W$ .

By specifying the nature of the elements and constraints in question, one may distinguish varieties of coherence. *Explanatory coherence*, which has received the most experimental attention, is that type of coherence in which elements are propositions and constraints implement various explanatory conditions – e.g., “A hypothesis coheres with what it explains” and “the more hypotheses it takes to explain something, the lower the degree of coherence.” Thagard (1989) develops a program, “ECHO”, that computes approximate solutions to explanatory coherence problems using a connectionist (neural network) algorithm.

Through a series of experiments, Schank and Ranney (1991) investigate “the relationship between human explanatory reasoning and explanatory evaluation as implemented in ECHO.” For example, in one of their experiments, participants evaluated the “believability” of two hypotheses, H1 and H2, in two scenarios differing only in how many pieces of evidence one of the hypotheses (H2) explained. ECHO predicts that which hypothesis is part of the most explanatorily coherent solution depends on which scenario one is in. The question, for the sake of testing the similarity of this explication of coherence, is whether participants’ judgments agree with ECHO’s predictions.

Results were favorable to Thagard’s explication. An ANOVA test revealed that participants’ assessments of H1 did not differ significantly across the two scenarios, but those pertaining to H2 did. Believability-ratings of H2 were significantly higher than those of H1 only in the second scenario. All of this was as predicted by ECHO. Other experimental tests were similarly favorable to this explicatum, leading Schank and Ranney

to conclude, “ECHO’s ability to model the subjects’ propositional evaluations well [...] suggests that ECHO shows continued promise for predicting human explanatory reasoning.” Schank and Ranney’s experiments test ECHO in cases where only the breadth of explanations is varied; Read and Marcus-Newhall (1993) extend this work, confirming the predictive accuracy of ECHO under other sorts of variations (e.g., simplicity).

Interestingly, the majority of experiments on the constraint-satisfaction notion of coherence have *not* focused on testing for similarity. Rather, these *use* ECHO to advance experimental psychological research on various topics – e.g., how people construct mental models of others (Read and Miller 1993), or how humans transform complex, ambiguous information into coherent decisions (Simon and Holyoak 2002). Here, instead of being tested by experimental work, formal epistemology is informing experimental work. This way of combining formal epistemology and experimentation is the primary focus of the next section.

## 2. Formal Epistemology Informing Experiments

In this section, we take a brief look at recent work in which the products of formal epistemology advance experimental research. Here, formal epistemology informs experimental work; however, the direction of information runs in the other direction too, as such work effectively tests the fruitfulness of an explicatum. I focus on two topics: the conjunction fallacy and judgment aggregation.

Since Tversky and Kahneman (1983) first presented their seminal research on the topic, psychologists have puzzled over the fact that, placed under certain conditions, a majority of people make the probabilistically erroneous judgment that a conjunction is

more likely than one of its conjuncts. In the most famous example of this putative **conjunction fallacy**, subjects are given the following information: “Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.” When faced with this information, a majority of participants (85% in the original study) judge it more likely that “Linda is a bank teller and is active in the feminist movement” than that “Linda is a bank-teller” (Tversky and Kahneman 1983).

A number of thinkers recently use the formal epistemological tools of confirmation theory to develop a precise account of these problematic judgments (e.g., Sides et al. 2002, Levi 2004, and Crupi, Fitelson, and Tentori 2008). Informally, the idea is simple: when people judge that a conjunction is more likely than is one of its conjuncts, their fallacious judgments accurately track confirmation relations. The information subjects are given about Linda, for example, plausibly confirms the hypothesis that she is a feminist bank-teller far more so than it does the alternative hypothesis that she is a bank-teller.

Crupi, Fitelson, and Tentori (2008) put forward a detailed account along these lines. Where the fallacious judgment takes the form  $\Pr(H1\&H2|E) > \Pr(H1|E)$ , they point out that cases in which participants are susceptible to the fallacy share a set of formal traits (letting  $c$  stand for any of the measures in **Table 1**):

- (i)  $E$  is negatively (if at all) correlated with  $H1$ :  $c(H1,E) \leq 0$ ;
- (ii)  $E$  is positively correlated with  $H2$ , conditional on  $H1$ :  $c(H2,E|H1) > 0$ ;
- (iii)  $H1$  and  $H2$  are mildly (if at all) negatively correlated.

As Crupi, Fitelson, and Tentori prove, (i) and (ii) imply that the conjunction  $H1\&H2$  is better confirmed by  $E$  than  $H1$  alone:  $c(H1\&H2,E) > c(H1,E)$ . Inspired by this result, they hypothesize that “on conditions (i)-(iii), most participants may depart from the relevant

probabilistic relationship between  $\Pr(H1\&H2|E)$  and  $\Pr(H1|E)$  because of the perception that  $c(H1\&H2,E) > c(H1,E)$ .”

Because this account uses the formal epistemological explication of confirmation, it has the great advantage of being unambiguous and testable; it is “sufficiently well defined to allow for critical examination” (Tentori, Crupi, and Russo 2013, 247). Schupbach (2012) performs and presents an initial experimental study of this account, which aims to establish the critical point that people are sometimes not susceptible to the fallacy in cases that satisfy (i)-(iii). However, in their response, Tentori and Crupi (2012) point out that the inequality  $c(H1\&H2,E) > c(H1,E)$  is plausibly quite weak in Schupbach’s experiments. This leads them to extend his study by varying the strength of the confirmation inequality across similar cases. Their results show that the percentage of participants who commit the fallacy varies (from 0% to nearly 100%) predictably with the confirmation relation, strongly supporting their original account.

In another recent study, Tentori, Crupi, and Russo (2013) compare their approach with other accounts of the fallacy. They note that these other accounts make a common prediction: that as the probability of the added conjunct increases, so too does people’s susceptibility to the fallacy. On their confirmation-theoretic account, this need not be the case. The results of four different experiments vindicate their account over competitors, concurring that “the conjunction fallacy predominantly depends on the perceived confirmation of the added conjunct rather than its perceived probability.”

Bayesian confirmation theory has also usefully been applied in recent accounts of the experimentally robust results of Wason’s selection task – see (Oaksford and Chater 2007, ch. 6) for a review. For example, Nickerson (1996) suggests that people’s seemingly

irrational preferences in these tasks may be shown to “have a robust rational basis” insofar as they track a confirmation-theoretic notion of “expected impact” of evidence (the expected, absolute value of  $C_d$  – **Table 1**). Oaksford and Chater (1994, 2007) offer a similar account using a related notion of expected decrease in uncertainty (measured using Shannon information). The implications and value of such accounts remain open to debate, however – see Fitelson and Hawthorne’s (2010) critical discussion.

These examples show that the tools of formal epistemology can help shed new light on perplexing psychological issues. In these cases, Bayesian confirmation theory is essential to the statement of new accounts of the phenomena. And the fact that these accounts are formally explicated guides us in experimentally testing their precise implications and predictions.

In *social epistemology*, philosophers study epistemic topics as they relate to groups as opposed to individuals. For example, social epistemologists analyze what it means when we predicate knowledge, rationality, and the like to groups of people. Formal work in social epistemology has focused on developing and refining concepts and methods, which are not meant to illuminate pretheoretical notions so much as to improve group performances. In other words, formal social epistemology typically aims to engineer precise concepts, designed to improve group judgments and decision-making.

A prime example can be found in work on *judgment aggregation*. To construct a collective, group judgment out of the judgments of group members, some method of aggregation is needed. There are a variety of aggregation methods available (averaging, deliberation procedures, voting methods, etc.), and each is potentially good for different purposes in different contexts – see (List 2012) and (Lyon and Pacuit 2014) for overviews.

In a recent study, Lyon, Fidler, and Burgman (2012) run an experiment to compare methods for aggregating individual interval judgments – e.g., the judgment that Salt Lake City’s average elevation is between 3,500 and 5,000 feet. Participants shared interval judgments in response to questions, along with corresponding confidence estimates (how likely they thought it was that the true value would be found in their given interval). Each subject was then asked for confidence estimates again, but this time regarding the interval judgments of another participant (the goal here was to reduce an overconfidence effect). Interval judgments were then aggregated using the following methods:

- *Maximum Confidence Slating*: choose the judgment with the highest associated confidence level. MJCS uses confidence levels assigned by the judge; MECS uses those assigned by the separate evaluator.
- *Maximum Information Slating*: choose the judgment with the highest associated value of confidence level divided by interval width. MJIS uses confidence levels assigned by the judge; MEIS uses those assigned by the separate evaluator.
- *Maximum Consensus Slating (MConS)*: choose the interval on which the two confidence estimates (the judge’s and evaluator’s) most closely agree.

This study produced several interesting results. First, when confidence estimates are given for others’ judgments instead of one’s own, overconfidence is greatly reduced; while both the confidence estimates of judges and evaluators displayed overconfidence, this was reduced 60.5% in the latter case. This important finding gives us a potential solution for cases in which the overconfidence of individuals is pernicious to group performance. The study also reveals that the various aggregation methods “have different strengths,” with different methods coming out on top on different dimensions (hit rate,

accuracy, calibration). For example, MECS has the highest hit rate of any considered method – meaning that the intervals it picks out contain the true value in a greater percentage of cases than do those picked out by other means – while MConS displayed consistently low overconfidence levels.

This experiment not only verifies that different aggregation methods are good for different purposes, but it also gives specific guidance regarding *which* methods do certain tasks well. Such information is vital when our philosophical goal is to engineer concepts and methods that lead us more efficiently to epistemically desirable ends. Similar studies explore the fruitfulness of other aggregation methods in other contexts. For example, these same researchers (with additional collaborators; Wintle et al. 2013) present the early results of a long-term study observing the forecasting performance of group predictions, aggregated using structured (*Delphi*-style) deliberation in groups of diverse individuals.

### 3. Other Topics and Further Reading

This chapter has surveyed work in which formal and experimental methods provide mutually informative means to epistemological ends. Experiments have been used to test the extent to which formal epistemological explications satisfy important desiderata, similarity and fruitfulness. And such explicative research has, in turn, advanced epistemically-relevant experimental work, allowing for exact and precisely testable new theories of puzzling phenomena.

The above survey is both simplified and incomplete. Interested readers should follow relevant citations to dig more deeply into any discussed topic. Other subjects this chapter might just as well have covered include work on causal reasoning and modeling



(Danks, this volume; Sloman and Park, this volume), reasoning with conditionals beyond the Wason task (Douven, this volume), the learning and emergence of conventions (Devetag, Hosni, and Sillari 2013; Bruner et al. Forthcoming), and general issues related to logic (Ripley, this volume). Another field of research I have sidestepped (as any useful discussion of it would take up an entire chapter of itself) uses experiments to compare inductive logics with “everyday rationality.” The reader might refer to (Chater et al. 2009) for an overview of this literature, to (Oaksford and Chater 2007) for a defense of the predominant Bayesian paradigm, and to (Unterhuber and Schurz 2013) for one recent experimental criticism of the Bayesian approach.

## References

- Bonawitz, Elizabeth Baraff, and Tania Lombrozo. 2012. "Occam's Rattle." *Developmental Psychology*, 48: 1156–1164.
- Bruner, Justin, Cailin O'Connor, Hannah Rubin, and Simon Huttegger. Forthcoming. "David Lewis in the Lab." *Synthese*.
- Carnap, Rudolf. 1950. *Logical Foundations of Probability*. Chicago: University of Chicago Press.
- Chater, Nick, Mike Oaksford, Ulrike Hahn, and Evan Heit. 2009. "Inductive Logic and Empirical Psychology." In *Handbook of the History of Logic, volume 10: Inductive Logic*, edited by Dov M. Gabbay, Stephan Hartmann, and John Woods, 553–624. Amsterdam: North Holland.
- Crupi, Vincenzo, Branden Fitelson, and Katya Tentori. 2008. "Probability, Confirmation, and the Conjunction Fallacy." *Thinking and Reasoning*, 14: 182–199.
- Crupi, Vincenzo, Katya Tentori, and Michel Gonzalez. 2007. "On Bayesian Measures of Evidential Support." *Philosophy of Science*, 74: 229–252.
- Devetag, Giovanna, Hykel Hosni, and Giacomo Sillari. 2013. "You Better Play 7." *Synthese*, 190: 1351–1381.
- Fitelson, Branden. 1999. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science*, 66: S362–S378.
- Fitelson, Branden, and James Hawthorne. 2010. "The Wason Task(s) and the Paradox of Confirmation." *Philosophical Perspectives*, 24: 207–241.
- Harris, Adam J. L., and Ulrike Hahn. 2009. "Bayesian Rationality in Evaluating Multiple Testimonies." *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35: 1366–1373.
- Levi, Isaac. 2004. "Jaakko Hintikka." *Synthese*, 140: 37–41.
- List, Christian. 2012. "The Theory of Judgment Aggregation." *Synthese*, 187: 179–207.

- Lombrozo, Tania. 2007. "Simplicity and Probability in Causal Explanation." *Cognitive Psychology*, 55: 232–257.
- Lyon, Aidan, Fiona Fidler, and Mark Burgman. 2012. "Judgement Swapping and Aggregation." In *AAAI Fall 2012 Technical Report on Machine Aggregation of Human Judgment*. AAAI Press.
- Lyon, Aidan, and Eric Pacuit. 2014. "The Wisdom of Crowds." In *Springer Handbook for Human Computation*. Springer.
- Millgram, Elijah, and Paul Thagard. 1996. "Deliberative coherence." *Synthese*, 108: 63–88.
- Nickerson, Raymond S. 1996. "Hempel's Paradox and Wason's Selection Task." *Thinking and Reasoning*, 2: 1–31.
- Oaksford, Mike, and Nick Chater. 1994. "A Rational Analysis of the Selection Task as Optimal Data Selection." *Psychological Review*, 101: 608–631.
- Oaksford, Mike, and Nick Chater. 2007. *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford: Oxford University Press.
- Pacer, Michael, Joseph Williams, Xi Chen, Tania Lombrozo, and Thomas L. Griffiths. 2013. "Evaluating Computational Models of Explanation using Human Judgments." *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*, 498–507.
- Read, Stephen J., and Amy Marcus-Newhall. 1993. "Explanatory Coherence in Social Explanations." *Journal of Personality and Social Psychology*, 65: 429–447.
- Read, Stephen J., and Lynn C. Miller. 1993. "Rapist or 'Regular Guy'." *Personality and Social Psychology Bulletin*, 19: 526–541.
- Rusconi, Patrice, Marco Marelli, Marco D'Addario, Selena Russo, and Paolo Cherubini. Forthcoming. "Evidence evaluation." *Journal of Experimental Psychology: Learning, Memory and Cognition*.

- Schank, Patricia, and Michael Ranney. 1991. "The Psychological Fidelity of ECHO." In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 892–897. Hillsdale, NJ: Erlbaum.
- Schubach, Jonah N. 2011. "Comparing Probabilistic Measures of Explanatory Power." *Philosophy of Science*, 78: 813–829.
- Schubach, Jonah N. 2012. "Is the Conjunction Fallacy Tied to Probabilistic Confirmation?" *Synthese*, 184: 13–27.
- Schubach, Jonah N. Unpublished. "Experimental Explication."
- Sides, Ashley, Daniel Osherson, Nicolao Bonini, and Riccardo Viale. 2002. "On the Reality of the Conjunction Fallacy." *Memory and Cognition*, 30: 191–198.
- Simon, Dan, and Keith J. Holyoak. 2002. "Structural Dynamics of Cognition." *Personality and Social Psychology Review*, 6: 283–294.
- Strawson, P. F. 1963. "Carnap's Views on Constructed Systems versus Natural Languages in Analytic Philosophy." In *The Philosophy of Rudolf Carnap*, edited by Paul Arthur Schilpp, 503–518. La Salle, IL: Open Court.
- Tentori, Katya, and Vincenzo Crupi. 2012. "How the Conjunction Fallacy is Tied to Probabilistic Confirmation." *Synthese*, 184: 3–12.
- Tentori, Katya, Vincenzo Crupi, Nicolao Bonini, and Daniel Osherson. 2007. "Comparison of Confirmation Measures." *Cognition*, 103: 107–119.
- Tentori, Katya, Vincenzo Crupi, and Daniel Osherson. 2007. "Determinants of Confirmation." *Psychonomic Bulletin & Review*, 14: 877–883.
- Tentori, Katya, Vincenzo Crupi, and Selena Russo. 2013. "On the Determinants of the Conjunction Fallacy." *Journal of Experimental Psychology: General*, 142: 235–255.
- Thagard, Paul. 1989. "Explanatory Coherence." *Behavioral and Brain Sciences*, 12: 435–502.
- Thagard, Paul. 2000. *Coherence in Thought and Action*. Cambridge, Mass: MIT Press.

Tversky, Amos, and Daniel Kahneman. 1983. "Extensional versus Intuitive Reasoning." *Psychological Review*, 90: 293–315.

Unterhuber, Matthias, and Gerhard Schurz. 2013. "The New Tweety Puzzle." *Synthese*, 190: 1407–1435.

Wintle, Bonnie, Steven Mascaro, Fiona Fidler, Marissa McBride, Mark Burgman, Louisa Flander, Geoff Saw, Charles Twardy, Aidan Lyon, and Brian Manning. 2013. "The Intelligence Game." In *Proceedings of the 5th Australian Security and Intelligence Conference*, 14–26.

## Cross-references

SEE CHAPTER XX (Danks on "Causal Search and Causal Modeling")

SEE CHAPTER XX (Douven on "Conditionals")

SEE CHAPTER XX (Lombrozo on "Explanation")

SEE CHAPTER XX (Ripley on "Logic")

SEE CHAPTER XX (Sloman and Park on "Causal Reasoning and Learning")

## Biographical Note

Jonah N. Schupbach is Assistant Professor of Philosophy at the University of Utah. His research interests pertain to the study of human reasoning. He has published numerous articles in top journals, including *The British Journal for the Philosophy of Science*, *Erkenntnis*, *Philosophical Studies*, *Philosophy of Science*, and *Synthese*.