Conceptual Preferences in Formal Epistemology

Vincent F. Hendricks
Department of Philsophy and Science Studies
Roskilde University
Denmark

GLLC XV: The Dynamics of Preferences and Intentions / Amsterdam 02/08

Epistemology: 5 Questions

Vincent F. Hendricks & Duncan Pritchard (eds.), Automatic Press / VIP, October 2008, ISBN 978-87-92130-07-5

The goal with these questions is to gain a sense of what contemporary epistemology is up to and its broader intellectual environment , with particular emphasis on the meeting point between mainstream and formal epistemology.

Interviewees

Horacio Arlo-Costa, Sergei Artemov, Johan van Benthem, Luc Bovens, Lorraine Code, Stewart Cohen, Fred Dretske, Robert Fogelin, Peter Gärdenfors, Alvin Goldman, **Clark Glymour**, John Greco, Alan Hajek, Joe Halpern, Sven Ove Hansson, John Hawthorne, Jaakko Hintikka, Wiebe van der Hoek, Keith DeRose, Martin Kusch, Kevin T. Kelly, Jon Kvanvig, Isaac Levi, Rohit Parikh, John Pollock, Hans Rott, Krister Segerberg, Ernest Sosa, Wolfgang Spohn, Barry Stroud, Moshe Vardi, Michael Williams, Timothy Williamson, Linda Zagzebski

⟶ **Clark Glymour** on the question: "What do you think the future of epistemology will (or should) hold?"

1 Conceptualities and Technicalities

- Ian Hacking once noted that the hallmark of any fertile research program is *plethora* – the ability of the program to produce new and interesting phenomena.

- Formal epistemology certainly qualifies as a fertile reserach program in the plethoric sense.

- Some buy into the idea that although the discipline is held together *technically*,

- it is not in any philosophically interesting sense held together *conceptually.*

- Buying this idea is going to leave one with buyer's regret.

2 Going Philosophical Formally

- Crucial philosophical concepts like
  - reliability analysis
  - first vs. third person perspectives in inquiry, even
  - downright conceptual analysis
  - ...

- are unilaterally shared by approaches ranging from belief revision theory, dynamic epistemic logic, Bayesianism to formal learning theory

- Today we will consider some of the conceptual preferences in formal epistemology and their philosophical impact.

2.1 Reliability Analysis

- Question: What does it mean for a method or a process to be reliable?
  - William James: Avoiding error vs. gaining truth – where, how often, how many?

- Mainstream epistemologists have had a host of anwers to this question. A few examples:
  - Alvin Goldman: The justificational status of a belief is a function of the reliability of the process or processes that cause it, where reliability consists in the tendency of a process to produce beliefs that are true rather than false. ... Again, the degree of justifiedness is a function of reliability ... Returning to the categorical conception of justifiedness, we might ask how reliable a belief forming process must be in order that its resultant beliefs be justified. A precise answer to this question should not be expected. Our conception of justification is vague in this respect. It does seem clear, however, that perfect reliability isn't required. (1976)

- **Not so good**
  - The Gettier-problem re-emerges due to ignoring possibilities of error which in turn undercut the reliabilist solution to the Gettier-paradox: Consider a case in which an arbitrary hypothesis h is a reliable indicator of yet another hypothesis h' in the following way:

$$P(h' \mid h) = .95 \text{ and } P(\neg h' \mid \neg h) = .95. \quad (1)$$

    A method basing its conjecture about $h \vee h'$ given h' alone will be reliable. The probability of error when h occurs is 0 because h again entails $h \vee h'$. The probability of error when h' occurs is .05. So the over-all unconditional probability of mistake pertaining to $h \vee h'$ is given by

$$P(h) \cdot 0 + P(\neg h) \cdot .05 \leq .05. \quad (2)$$

    The method is accordingly highly reliable. Nevertheless (2) is exactly a Gettier-case!

- Another try:
  - Robert Nozick: To know is to have a belief that tracks the truth. Knowledge is a particular way of being connected to the world, having a specific real factual connection to the world: tracking it. (1981)

In other words: Ξ *knows* h *iff*

1. h *is true*,
2. Ξ *believes that* h,
3. $\neg h \rightarrowtail \neg$(Ξ *believes that* h),
4. $h \rightarrowtail$ (Ξ *believes that* h).

where '$\rightarrowtail$' denotes the counterfactual conditional.

- **Much better**
  - The Gettier-problem is solved,
  - and *reliability* is defined in terms of
    * refutation, $\neg h \rightarrowtail \neg(\Xi$ *believes that* $h)$, or
    * verification, $h \rightarrowtail (\Xi$ *believes that* $h)$
    * (i.e. decision, a recursive procedure)
    * in all close worlds circumscribed by the counterfactual

- Nozick in the long run ..."If the universal hypothesis were false, Ξ would not believe it 'now'"

- This provides a lead as to how reliability should be assessed.

- Reliability is assessed with respect to two parameters
  - Sense of success:
    - verification / refutation / decision / discovery ...
    - with certainty / with n-mind changes / in the limit / gradually ...
  - Range of worlds:
    - actual world
    - normal worlds
    - close worlds
    - defined by binary accessibility relation in Kripke-style semantics
    - or some other topology ...

- One may discriminate between two types of reliability often found in the literature:

> Categorical reliability—A concept of reliability is **categorical** if defined in terms of unequivocal success for some convergence criterion in the actual world or relevant possible worlds

> Stochastic reliability—A concept of reliability is **stochastic** if defined in terms of a success over failure ratio for some convergence criterion in the actual or relevant possible worlds

- Nozick is of the first type, Goldman of the second, and so it goes for Bayesianism, formal learning theory, epistemic logic, belief revision theory and most of the mainstream proposals.

	Gain truth	Avoid error
1a. [Goldman 96]	w	$-$
1b. [Goldman 96]	w	w'
2. [Goldman 86]	w, w'	$-$
3. [Goldman 79]	w	$-$
4. [Goldman 73]	w	w'
5. [Goldman 67]	w	$-$

Table 1: Goldman's views on forcing over 30 years

- Canons of inductive rationality or justification imposed on methods of inquiry can interfere with reliability.

- Call a method for assessing belief in a hypothesis *consistent* if the belief in the hypothesis is taken back as soon as the hypothesis has been refuted by the observed evidence

- The recommendation of consistency sounds like a rational prescription;

- It does not seem rational to continue to believe in a hypothesis refuted by available evidence.

- Ideal agents not limited by effective or computational constraints have no problem with the consistency requirement.

- Even if the scope of reliable inquiry is not limited for ideal agents by consistency it may for agents computationally bounded cause effective problems:

- Kelly and Schulte have proved that for special computable epistemic assessment problems, consistency stands in the way of getting to the truth. There is an epistemic problem which in fact is computably refutable with certainty by a computable method, but no consistent method, even with infinitely effective powers (in the sense that it is hyper-arithmetically defined), can as little as gradually solve this problem in terms of refutation or verification.

- A discovery by Osherson, Stob and Weinstein shows that the consistency prescription restricts the set of identifiable recursive enumerable Σ_1^0-sets to the recursive Δ_1^0-sets for effective hypothesis discovery methods. This result also has an impact on effective belief revision AGM-identifiers [Hendricks 97].

- Arguing that principles of rational inquiry are normative in terms of getting to the truth should be done cautiously.

- The principles advocated may just be in the way of the truth when it can be proved that there is a method violating the principles which could have gotten there.

- One should accordingly distinguish between two types of methodology;

Categorical methodology—a methodological recommendation is advanced to its own end regardless of finding the truth

and a methodology committed to finding the truth through the proper means and methods insofar the truth can be found:

Hypothetical methodology—a methodological recommendation is advanced in the aim of finding the truth

- Putnam provided one of the first computational epistemological (learning theoretical) results:

- For any algorithm of extrapolation based on a Carnapian theory of confirmation, there is an epistemic problem which the Carnapian extrapolator cannot acquire knowledge of even when fed all possible instances of the problem.

- Using confirmation theory is not necessarily a truth-conducive methodological strategy.

2.2 First vs. Third

- As epistemologists we are not only in the business of ascribing ourselves knowledge,

- but equally much in the business of ascribing knowledge to other agents.

- Lewis has pointed out that there is a significant difference between one agent ascribing himself knowledge in his local epistemic situation, and us ascribing him knowledge given the situation we are in.

  - Definition: Agent Ξ knows A iff Ξ's evidence eliminates all possibilities in which $\neg A$ – Psst – Except those possibilities *we* are properly ignoring.

- The two situations do not necessarily always coincide.

- **Lewis**: The logic of knowledge may be
  - strong for the first person, but
  - weak on the third person perspective
- **Nozick**: The logic of knowledge is
  - weak on the first person perspective
- **Levi**:
  - To gain truth and avoid error beliefs should be chosen carrying the highest 'epistemic utility' (epistemic utility embodies truth as well as content)
  - Significant possibilities of error are forgivable just the agent settles for the belief with the highest epistemic utility in the particular context.
  - This may not exactly add up to real knowledge but it is good enough for decision and action.

- A couple of examples
  - $K_{\overline{=}}A \rightarrow K_{\overline{=}}K_{\overline{=}}A$
    - Levi's epistemological program is a version of a first person perspective emphasizing a distinction between *the logic of truth* and *the logic of consistency*—though related the two distinctions are not exactly the same.
    - Levi denies the validity of various epistemic axioms as axioms of an epistemic logic of truth.
    - This crudely means to reject these axioms as axioms for a third person knowledge operative.
    - The KK-thesis is here valid as an axiom serving regulative purposes of maintaining consistency for a rational epistemic agent.
    - The logic of truth for an epistemic agent on the other hand is not necessarily regulated by a principle like the KK-thesis.

- – $\neg K_{\Xi} A \rightarrow K_{\Xi} \neg K_{\Xi} A$
  - * Stalnaker has argued that it should be possible for a player Ξ to know what a player Θ is going to do.It should be rendered possible in case Θ only has one rational choice, and Ξ knows Θ to be rational, that Ξ can predict what Θ is going to do.
  - * This should not imply however that it is impossible for Θ to act differently as he has the capacity to act irrationally.
  - * What is needed is a counterfactually possible world such that (i) Θ acts irrationally, but (ii) is incompatible with what Ξ knows.
  - * Ξ's prior beliefs in that counterfactual world must be the same as they are in the actual world for Θ could not influence Ξ's priors beliefs by making a contrary choice.

  - Then it has to be the case in the counterfactual world, that Ξ believes he knows something (e.g. that Θ is irrational) which he in fact does not know.

  - This is incompatible with **S5**.

- In general, one should carefully distinguish between:

> First person perspective—A perspective on scientific inquiry is **first person** if it considers what an agent can solve, can do or defend considering the available means for an end given the epistemic environment he is sunk into

Third person perspective—A perspective on scientific inquiry is **third person** if it considers what an agent could solve, could do or defend considering the best means for an end independently of the epistemic environment he is sunk into

- This distinction plays a severe role with respect to which axioms are valid for different knowledge operators and perspectives on inquiry.

- Do not criticize an apple for not being and orange!

- Question: Do there exist epistemic axioms separating the first person perspective from the first person perspecttive?

		COE	LE	MOE
N:	$\frac{A}{K_{\equiv}A}$	$1^s/3^s$	$1^s/3^s$	3^s
K:	$K_{\equiv}(A \to A') \to (K_{\equiv}A \to K_{\equiv}A')$	$(1^s)/(3^s)$	$1^s/3^s$	3^s
T:	$K_{\equiv}A \to A$	$1^s/3^s$	$1^s/3^s$	3^s
4:	$K_{\equiv}A \to K_{\equiv}K_{\equiv}A$	1^s	$1^s/3^s$	3^d
5:	$\neg K_{\equiv}A \to K_{\equiv}\neg K_{\equiv}A$	(1^s)	3^s	

- COE: Contextual Epistemology; LE: Logical Epistemology (epistemic logic); MOE: Modal Operator Epistemology.

- 1: First person perspective; 3: Third person perspective.

- s: synchronicity; d: diachronicity; (,): context-sensitive validity.

2.3 Downright Conceptual Analysis

- One often hears the characterization that philosophy is much about *conceptual analysis*.

- Conceptual analysis in mainstream epistemology is enjoying a revival these days.

- The goal of these conceptual exercises is to spell out and elucidate some of the epistemologically significant notions like knowledge, justification and rationality that ordinary folk use on a daily basis.

- An integral part of the elucidation process is to stretch the usage of these concepts to the *max* in order to reveal their limitations and what these limits in turn reveal about the nature of human cognition.

- Jackson: "Consulting intuitions about possible cases".

- An often recited charge against formal epistemology is that it doesn't encompass, and doesn't care about, this important aspect of epistemological inquiry

- This is plainly false for a number of reasons.
  - Formal epistemology has occasioned mainstream epistemology:
    * Lewis' epistemology may be seen as an attempt to provide general rules for the construction of what computer scientists using epistemic logic call 'runs' (vectors of possibilia with a specific range)
    * Stalnaker's use of belief revision to get out of the Gettier-problem is another case in point.
    * Computational epistemology has widened the general understanding of the problem of induction together with questioning intuition-driven norms often proposed for 'rational' inquiry.

    * Epistemic logic systematically addresses the modality of knowledge, sets up measures of its strength, sharpens mainstream discussions of rationality, infallibility, strategies for winning games and other pertinent epistemological issues.

    * Bayesianism has had a profound influence on the epistemological notion of confirmation and coherence.

    * Belief revision theory has influenced current philosophical discussions of theory change, etc.

– Formal epistemology *is* conceptual analysis too

    * If conceptual analysis is partly about stretching the concepts of epistemology as far as they can be taken, then

    * computational, modal operator epistemology, belief revision and epistemic logic are really quint-essential examples of this very practice.

* Considering what it takes to solve an epistemic problem independently of an agent's local circumstances, *a priori* means-ends analysis, examining the combinatorial structure of solvable epistemic problems in terms of complexity based characterization theorems for different notions of convergence and procedural success,

* providing methodological hypothetical imperatives like "*If your goal is validate modal system* **SX***, then use method* Ξ"

* considering the validity of knowledge in the limit, stretching the concept of knowledge all the way to entail logical omniscience and so forth

* are all rigorous ways of doing *conceptual analysis*

* and,

* *those are just some of our preferences and intentions in formal epistemology.*