The tradeoff problem and two-level epistemic consequentialism

Hilary Greaves (Oxford)
Workshop on *Epistemic consequentialism: Problems and prospects*
LSE, 21 November 2014

1) Epistemic consequentialism
   a) Project: justify the (intuitive) norms of epistemic rationality by showing that following these norms [tends to] promote some measure of 'epistemic value'.
   b) Motivations
      i) Fundamentally nonconsequentialist norms of (any kind of) rationality are mysterious
      ii) Need to explain what epistemic rationality has to do with *truth.*
2) Clarifications; preliminary objections and replies
   a) ∃ 'full-belief' and 'credence' versions
   b) ∃ 'subjective' and 'objective' versions
   c) 'Epistemic value': just degree of match to the *truth* (not e.g. degree of justification)
   d) Objection: it's not epistemically rational[ly required] to count blades of grass
      i) Reply: epistemic consequentialism evaluates only *epistemic* acts
   e) Objection: Epistemic voluntarism is false
      i) Reply: Neither *evaluation* nor *vindication of internalised procedures* presupposes voluntarism
3) A worse(?) problem for epistemic consequentialism: The tradeoff problem
   a) A warm-up version of the tradeoff problem
      i) The case: You are being interviewed by the Templeton Foundation for a very important grant. If you succeed in getting the grant, you will be able to carry out a programme of research that will reveal to you many of the most important truths about the universe. Unfortunately, the Templeton Foundation issues grants only to believers.
      ii) Datum: It is not *epistemically* rational for you to believe in God.
      iii) Worry: Epistemic consequentialism disagrees?
      iv) Reply: time-relative epistemic consequentialism (e.g. Foley)
   b) A worse tradeoff problem: Suppose there is some proposition A and action ϕ such that:
      i) You currently have very low credence in A (and 'rightly so')
      ii) You think that performing ϕ would be good if A is true, bad if A is false; thus you think that believing A would cause you to decide to ϕ, while disbelieving A would cause you to decide not to ϕ
      iii) The future is much more easily predictable if you perform ϕ than if you refrain from performing ϕ
      Is it then epistemically rational to believe A? (Example: A = my eldest child behaved particularly well yesterday, ϕ = giving the ice-cream to the eldest child)
   c) Numerous other tradeoff cases: 'truth fairy' (Jenkins), 'epistemic imps' (Greaves), paradox case (Berker), higher-order beliefs case (Berker)
      i) Note: these cases have widely *varying* degrees of 'outlandishness'
4) Possible responses to the tradeoff problem
   a) Pessimism: give up on epistemic consequentialism (Berker)
   b) Revisionism: give up the nonconsequentialist particular-case intuitions (caviliers)

c) Perseverance: Amend epistemic consequentialism so that it no longer yields the problematic (pro-tradeoff) judgments (But how? And is this ad-hoc?)

5) Two-level consequentialism
   a) The basic idea:
      i) Everyday practice employs rules/norms of a non-consequentialist character. (Both *decision-making* practice, and *evaluative* practice.)
      ii) *But* there is a consequentialist story explaining why (a) everyday practice does not directly employ consequentialism, and (b) why everyday practice employs the *particular* nonconsequentialist rules/norms that it does.
   b) This allows the would-be consequentialist who is (however) faced with strong and apparently non-consequentialist intuitions to "have her cake and eat it".

6) Two-level consequentialism in ethics
   a) The self-defeatingness objection to (straightforward) ethical consequentialism
      i) Calculation time
      ii) Personal bias
      iii) (Alienation)
   b) Reply: "two-level-ism" (e.g. Hare)
      i) *Agree* that because we are (i) imperfect calculators and (ii) imperfectly impartial, it will lead to better consequences if we, on an everyday basis, generally guide our decisions by nonconsequentialist norms, than if we try to do full consequentialist calculations for every decision.
         (1) 'Because we are imperfect': two-level-ism doesn't apply to 'archangels'
         (2) 'On an everyday basis': *sometimes* (even) we should return to full consequentialist calculation, e.g. when stakes are high, and in occasional 'moments of cool reflection'
      ii) But this doesn't mean that consequentialist principles have no role to play in ethics…
   c) …Making two-levelism precise: two proposals (aside: which is better?)
      i) 'Rule-consequentialist' way: direct-consequentialist evaluation of rules, indirect evaluation of acts *via* prior evaluation of rules
      ii) 'Global-consequentialist' way: direct-consequentialist evaluation of acts but also of rules

7) Two-level consequentialism in epistemology?
   a) A way of having and eating our epistemic cake? (A species of either 'perseverance' or (now misnamed) 'pessimism')
   b) How this might work
      i) Part I: In part because of the scarcity/absence of situations relevantly similar to the above ice-cream case, the consequentialist pressure to internalise genuinely consequentialist norms is (anyway) weak – non-consequentialist norms that agree with the consequentialist ones on *the majority of actually-encountered cases* perform almost as well.
      ii) Part II: Because of imperfections in us, we would actually attain a *lesser* degree of fit between our beliefs and the truth overall if we attempted to internalise epistemic-consequentialist epistemic norms.
   c) Objections
      (1) Objection to Part I: Not all cases of "consequentialist wins" involve truth-fairies. There's nothing outlandish or unusual about the ice-cream case above, or Berker's cases.

(2) Objection to Part II: unlike the ethical case, in the epistemic case there's no plausible story about what the relevant 'imperfections in us' might be, or how they would lead to the stated consequence.

  (a) "We're imperfect calculators, and therefore…

      (i) "…The time it would take us to perform the consequentialist calculations would be too high a cost"?

          1. Not (or not obviously) an *epistemic* cost

      (ii) "…We'd more often get the wrong answers from calculations than from following simpler rules"?

          1. Example?

  (b) "We're psychologically incapable of internalising the epistemic-consequentialist norms, so there would not be good epistemic consequences of *trying* to do so." (Elstein & Jenkins)

      (i) But why think this? There's no (independent) evidence for it.[1]

8) Conclusions

  a) Epistemic consequentialism does not, at least not *directly*, recover intuitive verdicts of epistemic irrationality (the tradeoff problem).

  b) But (i) a *fundamentally* non-consequentialist normative theory is mysterious, and (ii) pending further enlightenment, we should be wary of *simply* revising our intuitive judgments.

  c) "Two-level" epistemic consequentialism, *if* it could be defended, would avoid both horns of this dilemma.

  d) But defending two-level epistemic consequentialism requires explaining why we either cannot or should not [try to] internalise consequentialist epistemic norms. This appears to be significantly harder than the analogous task in the ethical case, and (anyway) has not been done.

---

[1] Elstein and Jenkins write that "the point that we aren't able to respond to Truth-Fairy-like offers in a way that wins the epistemic prize on offer" [i.e., we aren't able to form our beliefs in accordance with epistemic consequentialism in tradeoff-problem cases] "tells us that rules which recommend acceptance of the relevant proposition in such cases… are very difficult, even practically impossible, to internalise". But this is either just mistaken, or question-begging.