CANCER CARE AND SAVING PARROTS

Hilary Greaves (Oxford)

'Philosophical foundations of effective altruism' conference

St Andrews, 30 March 2016

"The EA questions"

- Two questions for would-be effective altruists:
  - (1) How much to give?
  - (2) Which organisations to give it to?
- Moral philosophers' versions of these questions
  - (1) How much is one *morally required* to give?
  - (2) What are the *moral requirements* concerning which organisations to give to?
- (Focus on giving *money*, for simplicity/concreteness. But the issues apply more generally.)

The austere answers

- The austere answers
  - (A1) You're morally required to give up to the 'point of equal marginal utility' (PEMU).
  - (A2) You're morally required to: give only to the most cost-effective charities.
- Concerns about these answers
  - Re (A1): Note (for whatever it's worth…) that *no-one actually does this*.
  - Re (A2): What about
    - Personal connections to benefactors: The neighbourhood child who knocks on your door asking for sponsorship
    - Personal connections to beneficiaries: Supporting the refugee camp you have just visited
    - Personal connections to causes: Feminism, cancer research

Alternatives to the austere answers

- Permissive answers:
  - You're permitted to give less than the austere answer demands (although giving more would be agent-neutrally better).
  - You're permitted to give to less cost-effective charities (although giving to more cost-effective charities would be agent-neutrally better).
- Scalar answers
  - Giving more (up to the maximisation point), rather than less, is better.
  - Giving to charities that are more cost-effective, rather than less, is better.
  - Eschew any notion of moral permissibility/requirement.
- A mixed answer
  - You're permitted to give less than the austere answer demands (although giving more would be better).
  - *But* you are required to: give only to the most cost-effective charities.

A mixed answer in the EA literature (maybe)

- "Should I have donated to the Fistula Foundation, knowing that I could do more to help people if I donated elsewhere? I do not think so… [and] similar thoughts apply to deciding what cause to focus on more generally. … By all means, we should harness the sadness we feel at the loss of a loved one in order to make the world a better place. But we should focus that motivation on preventing death and improving lives, rather than preventing death and improving lives in one very specific way. Any other decision would be <u>unfair</u> to those whom we could have helped more." (MacAskill 2015; emphasis added)

Outline

- Support for a mixed answer?
  - Parrots, and the 'argument from analogy'
- Three ways to deny (A1)
  - 'Long-game maximising consequentialism'
  - Rule-consequentialism
  - Agent-centred prerogatives
- …all of which about-equally rationalise denying (A2) too.
- The argument from analogy revisited
- Conclusions
  - A mixed answer to the 'EA questions' is motivationally unstable.
  - And the common intuitions about parrots may be wrong.

Support for the mixed answer?

- Analogy: The burning building

  You are outside a burning building. Inside, a person and a parrot are trapped. You can enter the building in an attempt to rescue either the person or the parrot, at some risk to yourself (the same risk for both rescue missions). Or you can stay outside.

- Arguably: in this case, it is permissible to save the person, and permissible to stay outside, but not permissible to save the parrot.

- But isn't this just like: it's permissible to give to the more-effective charity, and permissible to keep the money for yourself, but not permissible to give to the less-effective charity?

- If so, whatever justifies the intuitive answer to the parrot case will also justify a mixed answer to the EA questions.

Inspiration from parrots

- What *is* the explanation of the intuitive permissibility pattern in the parrot case? Two stabs at this:
  - It's both permissible to accept some cost to yourself, and to refuse that cost. But holding fixed the amount of cost-to-yourself that you have accepted, you're required to maximise aggregate benefit to others.
  - More generally: There are two morally relevant respects of betterness in play: betterness w.r.t. cost-to-oneself, and betterness w.r.t. aggregate-benefit-to-others. The choice of how to weigh these up against one another is (perhaps within reasonable limits) a matter of agential prerogative. But a choice is permissible only if it is rationalisable by some such weighing.
    - I.e.: maximise $\lambda V_{agent} + (1-\lambda) V_{others}$, for some reasonable value of λ.

Doubts about parrots

- It's not clear why betterness w.r.t. me, the parrot and the stranger cannot be *three* 'morally relevant respects', with agential freedom over how to weigh all 3.
- In particular, there doesn't seem to be anything *obnoxious* about someone who is more strongly motivated to rescue a parrot than a stranger. (Especially, but perhaps(?) not only, if it's his own parrot.)
  - "I just thought I'll have to go back and get [my parrot]. The fireman said 'you can't go in there', but it just had to be done." Kevin Ross, house fire survivor, 4 April 2015
- Still, those who have the standard "parrot intuition" seem committed to the view that, for whatever reason, there *are* only two morally relevant respects here.

The argument from analogy

- In the parrot case, [for some perhaps-unknown reason] it is permissible to save the person and permissible to stay outside, but impermissible to save the parrot.
- The analogue of "permissible to save person or neither, but impermissible to save parrot" is the mixed answer to the EA questions ("permissible to give to most cost-effective charities or not to give, but impermissible to give to less cost-effective charities").
- The two cases *are* analogous in all relevant respects.

Therefore,

- The mixed answer to the EA questions is correct [for some perhaps-unknown reason].

Against the mixed answer (however)

- Setting aside parrots for now…
- Basic strategy: Argue that each of the available ways of denying (A1) also rationalises denying (A2). Therefore a mixed answer is motivationally unstable.
- First task: what are the available ways of denying (A1)?
  - 'Long-game maximising consequentialism'
  - Rule-consequentialism
  - Agent-centred prerogatives

First route: Long-game maximising consequentialism

- The austere account would give the correct account of what maximising utilitarianism implies *in a 'last action' case*, i.e. if there were no implications for one's future actions of how much one gives now.

- But the actual case is not last-action.

- Likely detrimental consequences, for future giving, of giving too much now:
  - You get depressed. Among other things, this makes you less productive at work, hence less able to give in the future.
  - You get demoralised. This makes you more likely to just abandon the whole giving project.

- Difficult question: What amount of giving *is* 'long-game optimal'?
  - Probably a *lot* less than 'to the point of EMU'.
  - But probably still a lot more than almost anyone does give.

Second route: Rule-consequentialism

- Rule-consequentialist account: the facts about the moral permissibility of giving track the following condition for the optimal rule: whichever rule is such that the consequences of *people regarding it as the true standard of moral permissibility* are better than for any other rule.

- On this account, the facts about moral permissibility roughly track what it is strategically best for e.g. GWWC to suggest (in this domain, 'truth collapses into expediency-to-assert').

  - Maybe "give 10%" is the optimal answer to (1) (in this country, now).

Third route: Agent-centred prerogatives

- A more overtly non-consequentialist moral theory
  - Although considerations of overall good are relevant to permissibility, we are not in general morally required to maximise the good. There are often *agent-centred prerogatives* to choose actions that lead to less good from an impartial point of view, when the impartially-suboptimal action in question is sufficiently (proportionately) superior in terms of self-interest. (Scheffler, 'The rejection of consequentialism')
  - Maximise $\lambda V_{agent} + (1-\lambda) V_{impartial}$, for some reasonable value of λ.
- Precisely because maximising impartial good when choosing one's donation-amount would be so costly for the agent, this theory will include a moral permission to give significantly less.
  - (Again it is an open question how much less.)

Long-game maximising consequentialism on the second EA question

- It's not only *spending more money on oneself* that can reduce the probability of depression and/or demoralisation: giving some of the funds in question to causes that are closer to one's heart could so too.
  - Up to a point. 'Long-game maximising consequentialism' probably doesn't rule out pledging £5 by way of sponsorship for the child on your doorstep, but probably does prohibit e.g. sending £100/£1000/etc. to Marie Curie.
  - So perhaps we get something *close* to (A2) here.
  - But note that this first route also got us something relatively close to (A1).

Rule-consequentialism on the second EA question

- There's a limit to how much a moral theory can require people to give *to the most cost-effective charities* before people will just ignore the theory in question. But it's plausible that people will be willing to give to causes closer to their heart, beyond that limit.

- Plausibly-optimal rule (?): Give 10% to the most cost-effective charities; in addition, give as much as you want beyond that to whatever other charities you want to support.

- This denies (A2), so this isn't the sought-after justification for a mixed answer either.

Third route: Agent-centred prerogatives

- Actually, the binary division into 'impartial value' and 'self-interest' is too crude – if 'self-interest' is taken literally.

- A more plausible model (and indeed the one Scheffler proposed) allows the agent to assign importance greater than the impartially justified amount to both persons *and causes* that are close to the agent's heart, not only to the agent herself. (V_{agent} is not just about *self*-interest.)

- But this model then very clearly permits support for less cost-effective charities/causes, in much the same way as it permits keeping more than the impartially optimal quantity 'for oneself'.
  - (In the here-relevant sense, 'keeping for oneself' is just 'giving' to a particularly cost-ineffective 'charity' that is particularly 'close to one's heart'.)

Parrots revisited

- The argument from analogy can also be run as a *modus tollens*.
- Our analysis of why the mixed answer is incorrect also suggests reasons why the common intuitions about parrots may be incorrect.
  - There could be an agent-centred prerogative to save the parrot (whether because it's your own pet parrot, or just because you particularly like parrots).

Tentative conclusions

- A mixed answer is motivationally inconsistent. Anyone who rejects (A1) should also reject (A2).

- Note that these arguments only support rejecting (A2) *in order to permit preference for a suboptimal cause for which the donor has special concern*. I haven't defended the permissibility of 'gratuitous' deviations from maximising cost-effectiveness (also an important issue for EA).

- Perhaps it is permissible to save the parrot rather than the human after all.