This essay first identifies the component of morality that Harman claims to arise from agreements between individuals, and then outlines his subsequent arguments for this claim, discussing them along the way. The primary positive argument given by Harman comes in the form of an argument by best explanation, where he argues constructivism is the only theory that satisfactorily accounts for two moral puzzles. In final discussion, I note that although Harman seemingly offers the best explanation relative to the alternatives considered, it is arguable that he has not considered enough reasonable alternatives, thus I ultimately conclude Harman’s arguments for constructivism fail.

Generally, moral contractarians claim ‘that moral norms derive their normative force from the idea of contract or mutual agreement’[[1]](#footnote-1) and these contracts can be hypothetical or actual. Harman defends agreements of the latter kind[[2]](#footnote-2). In Harman’s view moral oughts can be formulated by the four-place predicate ‘Ought (*A, D, C, M*)’[[3]](#footnote-3). For Julius to say: “Herod, you morally ought not to kill innocent children” Julius implies that Herod (*A. An agent*) has motivating reasons (*M*), all things considered (*C*) to not kill children (*D. A type of act*) and Julius himself endorses these reasons. Secondly, for Julius to say Nero (A) is evil for having murdering children, Julius implicitly shows the motivational reasons (*M*) that hehas for *not* killing children, could not have been reasons Nero had, because reasons motivate those who have them, but Nero was motivated *to* kill children. These motivating reasons for action ‘must have their source in goals, desires, and intentions’[[4]](#footnote-4) which Harman in turn argues are ‘intentions to adhere to a particular agreement on the understanding that others intend to do so’[[5]](#footnote-5). Moral ‘inner’[[6]](#footnote-6) judgements are therefore only properly made relative to these shared intentions to adhere to an agreement. Those outside the shared intention to adhere to *M*, are ‘beyond the motivational reach of the relevant moral considerations’[[7]](#footnote-7) as Nero was in the above illustration. I proceed now to outline Harman’s positive arguments that these *M* are intentions to adhere to an agreement.

Harman hypothesises that his theory provides the only account for two moral puzzles[[8]](#footnote-8), thus, he must reject all alternative accounts before, by Harman’s own standard[[9]](#footnote-9), being justified in making this claim. I return to this later. The first moral puzzle is: (1) the duty not to harm is regarded as more stringent than the duty to help others; for example, we tend to believe that we should not murder one person and take their organs to save others who need them[[10]](#footnote-10). Harman explains this as follows: ‘morality derives from an agreement among people of varying powers and resources’[[11]](#footnote-11) and these agreements are ‘reached through a process of mutual adjustment and implicit bargaining’[[12]](#footnote-12). When various individuals in society are bargaining about their intentions regarding harm, everyone knows how susceptible they are to being harmed; rich, poor, strong and weak alike. Therefore, most people will intend to keep an agreement to not harm others, on the condition others similarly intend. The duty to help, however, is one of greater importance for the weak and poor, whereas the rich and strong have fewer reasons to agree to a stringent duty of help, because the duty primarily falls on them, with little reciprocation. Thus, only a weaker duty of help gains ‘general acceptance’[[13]](#footnote-13) after we compromised down to a duty that everyone could intend to keep. Here Harman provides a reasonable explanation for the variance in stringency and further this same logic can be applied, *mutatis mutandis*, to theft, dishonesty etc. accounting again for our perception of their various stringencies as being the product of bargaining among different individuals. So far, Harman’s argument for constructivism is reasonable.

The second of the moral puzzles is as follows: (2) we all feel that ‘each person has an inalienable right to self-defence and self-preservation’[[14]](#footnote-14). Harman, focusing on these rights as incapable of being surrendered, argues that you cannot, ‘except in very special circumstances, rationally form the intention not to try to preserve your life if it should ever be threatened’[[15]](#footnote-15). We cannot now intend, when tortured later, to not resist, because we will never be motivated to adhere to this intention. In other words, we will disregard all prior (nominal) intentions to not preserve ourselves when tortured. Harman has given a reasonable explanation for the inalienable right to self-preservation and further, given (virtually) everyone else will similarly intend, our perception of this right as one we cannot take from others, is also accounted for. Additionally, this argument is logically valid from Harman’s premises, because motivating reasons for action must have their source in goals, desires, and intentions, and postulating that we will (virtually) never have a goal, desire, or intention to not preserve ourselves, then we will never have a motivating reason to not self-preserve, and therein always will self-preserve. Therefore, the right is inalienable.

To exhaust the alternative hypotheses, Harman illustrates their insufficiency in accounting for (1) and (2). Firstly, if ‘moral feelings derive from sympathy and concern for others’[[16]](#footnote-16) we could not give a satisfactory account of (1). This is so, given moral feelings would have us kill one healthy individual, to save five others, because our sympathy for five people is greater than our sympathy for one. However, this leads us to an (arguably) absurd conclusion, *i.e.,* not (1). Thus, Harman dismisses the possibility of morality being grounded in sympathy and concern for others. Further, Harman uses (1) to dismiss utilitarianism as an alternative account. Utilitarians claim we ought to maximise the overall good (where good is identified with pleasure) and everybody’s happiness is equally valued[[17]](#footnote-17). However, given that we maximise total pleasure by killing one and saving five, utilitarians conclude: not (1). Therefore, although most[[18]](#footnote-18) agree that (1) holds, the alternatives hypotheses fail to account for it as they logically conclude not (1), a *reductio ad absurdum[[19]](#footnote-19)*. Harman further claims that Utilitarians cannot ‘satisfactorily’[[20]](#footnote-20) justify (2), thus in sum, neither alternative explains (1) or (2), yet constructivism has.

Furthering this, Harman argues that ‘Anyaspect of morality that is susceptible’[[21]](#footnote-21) to utilitarian or sympathy-based explanations, is also accounted for by implicit agreements, but not vice versa. Harman concedes Utilitarianism offers explanations for many aspects of our moral views[[22]](#footnote-22), but argues this in turn can be explained by implicit agreements. Considerations of utility, benevolence etc. could influence ‘our implicit agreements, so that the appeal [made by utilitarians, to maximise the good] is to a shared intention to adhere to those agreements [of utility, benevolence and sympathy etc.]’[[23]](#footnote-23). Therefore, whatsoever is appealed to, and accounted for by utilitarian/ sentimentalist principles, is so agreed upon because considerations of utility, benevolence etc. have influenced the implicit agreements we base our morality upon, and we would never agree to said principles had utility/sympathy not been considered. Here again, the alternatives hypotheses are validly rejected as explanations of the source of moral norms, given what they do explain is in turn explained as proceeding from considerations of utility or sympathy in our implicit agreements. This is strong. By conceding the alternatives offer explanations for some aspects of our morality, but then in turn showing that this is of no damage to his position is a virtue of Harman’s argument. In short, he has shown both alternatives to be insufficient in certain regards, and that any moral explanations they do provide, is of no avail to their defence. Further, using this same framework of “What theory ‘so-and-so’ explains, is in turn explainable in terms of agreements” can be applied, *mutatis mutandis*, to other alternative hypotheses, thus strengthening Harman’s argument for constructivism.

Firstly, let us critically discuss a foundational element of Harman’s argument *qua* argument by best explanation. Harman limited himself to only discussing two potential alternatives. Therefore, an argument by best explanation is only valid if these are the only reasonable alternatives, and any other hypotheses do not need consideration[[24]](#footnote-24). Thus, his argument is as follows: A, B or C. Not A. Not B. Therefore C. That said, additional alternative accounts for (1) and (2) arguably exist, e.g. natural law theory, and without a reasoned qualification that theories beyond A and B can be ignored, Harman’s argument, *qua* argument by best explanation, is weakened. This alternative account can be briefly illustrated as follows:(2) one is, by one’s own nature, inclines towards that which is good, the preservation of one’s life is good, thus we are all inclined to preserve our own being[[25]](#footnote-25). Therein, by force of our nature as rational humans, we cannot surrender our right to self-preservation. Further, (1) is accounted for by the foundational principle of natural law, that ‘good is to be done and evil avoided’[[26]](#footnote-26), a precept ‘universally knowable by nature’[[27]](#footnote-27); Murdering a man is an evil act, and thus is to be avoided, even if the good of saving others would come. That this arguably reasonable alternative hypothesis exists, offering an account of both (1) & (2) and Harman did not argue against it, or show why this, or the multiple other alternatives beyond Utilitarianism and Sentimentalism, can reasonably be ignored, weakens his argument *qua* best explanation[[28]](#footnote-28). We can say Harman’s argument would be stronger, and avoid this question begging objection, if it was structured as follows: (P1) X or Y, where X is naturalism and Y non-naturalism. (P2) If Y then G. (P3) G is absurd. (C1) Not Y. (C2) Therefore, X. (P4) A, B and C are the only reasonable theories within X. This can then proceed as it was shown to before. Lastly, if one did, as Harman does, assume ‘A’ and ‘B’ are the only alternatives, this weakness in his argumentation for constructivism subsides, because we could then reasonably infer that Constructivism better explains moral norms than Utilitarianism and Sentimentalism.

To summarise, Harman presents two primary arguments for constructivism. The first was an argument by best explanation of two moral puzzles; In the second, Harman showed he can explain that which is explained by the alternative theories, in terms of his own theory. In discussion, I noted that the argument used by Harman requires a reasoned rejection of every plausible alternative to be valid, yet he arguably does not do so and absent a blanket defeater of widely regarded alternatives (‘Not Y’), this weakened his argument. I conclude Harman’s arguments for constructivism are strong, yet they are too narrow to warrant the claim that constructivism is true. In short, Harmans arguments for constructivism fail.

Bibliography

Aquinas, T. (No date) *Aquinas*. Available at: <https://aquinas.cc/> (Accessed: 27 December 2023).

Cudd, A. and Eftekhari, S. (2021) *Contractarianism*, *Stanford Encyclopedia of Philosophy*. Available at: [https://plato.stanford.edu/entries/contractarianism/](https://plato.stanford.edu/entries/contractarianism/%20) (Accessed: 13 December 2023).

Driver, J. (2014) *The history of Utilitarianism*, *Stanford Encyclopedia of Philosophy*. Available at: [https://plato.stanford.edu/entries/utilitarianism-history/](https://plato.stanford.edu/entries/utilitarianism-history/%20) (Accessed: 19 December 2023).

Harman, G. (1965) *The inference to the best explanation*, *JSTOR*. Available at: <https://www.jstor.org/stable/2183532> (Accessed: 03 January 2024).

Huemer, M. (2013) ‘The Hypothetical Social Contract Theory’, in *The Problem of Political Authority: An examination of the right to coerce and the duty to obey*. Basingstoke: Palgrave Macmillan, p. 36.

McInerny, R.M. (1997) Ethica Thomistica: The Moral Philosophy of Thomas Aquinas. Washington (D.C.): The Catholic University of America Press.

Murphy, M. (2019) *The Natural Law Tradition in ethics*, *Stanford Encyclopedia of Philosophy*. Available at: [https://plato.stanford.edu/entries/natural-law-ethics/](https://plato.stanford.edu/entries/natural-law-ethics/%20) (Accessed: 03 January 2024).

Shafer-Landau, R., Cuneo, T. and Harman, G. (2007) ‘Moral Relativism Defended’, in *Foundations of Ethics: An anthology*. Malden: Blackwell Publishing, pp. 84–92.

1. (Cudd et al, 2021) [↑](#footnote-ref-1)
2. (Shafer-Landau et al., 2007, p.84-85) [↑](#footnote-ref-2)
3. (*Ibid*, p.87) [↑](#footnote-ref-3)
4. (*Ibid*) [↑](#footnote-ref-4)
5. (Shafer-Landau et al., 2007*,* p.88). Cf. (*ibid*, p.86-87) [↑](#footnote-ref-5)
6. (*Ibid,* p.85) [↑](#footnote-ref-6)
7. (*Ibid*, p.86) [↑](#footnote-ref-7)
8. (*Ibid*, p.88) [↑](#footnote-ref-8)
9. (Harman, 1965, p.89) [↑](#footnote-ref-9)
10. (Shafer-landau et al, 2007, p.88) [↑](#footnote-ref-10)
11. (*Ibid*) [↑](#footnote-ref-11)
12. (*Ibid*, p.89) [↑](#footnote-ref-12)
13. (Shafer-Landau et al., 2007, p.89) [↑](#footnote-ref-13)
14. (*Ibid*, p.89) [↑](#footnote-ref-14)
15. (*Ibid*) [↑](#footnote-ref-15)
16. (Shafer-Landau et al., 2007, p.88) [↑](#footnote-ref-16)
17. (Driver, 2014). [↑](#footnote-ref-17)
18. (Shafer-Landau et al., 2007, p.88) [↑](#footnote-ref-18)
19. A ‘*reductio’* in so far as, the alternatives contradicted that which is generally agreed upon *i.e.,* (1) and (2). [↑](#footnote-ref-19)
20. (*Ibid*,p.89) [↑](#footnote-ref-20)
21. (*Ibid*) [↑](#footnote-ref-21)
22. (*Ibid*) [↑](#footnote-ref-22)
23. (Shafer-Landau et al., 2007, p.89) [↑](#footnote-ref-23)
24. This is an inference on my part, but Harman does say beyond implicit agreements, there is ‘no other way to account for’ (1) and (2) and thus he regards that beyond Utilitarianism and Sentimentalism, he need not consider other alternatives. [↑](#footnote-ref-24)
25. (Aquinas, *n.d.,* ST. I-II. Q94.A2.C.3); Cf (McInerny, 1997, p. 44-46) [↑](#footnote-ref-25)
26. (Murphy, 2019b, Sect. 1.3); Cf. (Aquinas, *n.d.,* ST IaIIae 94, 2) [↑](#footnote-ref-26)
27. (Murphy, 2019b, Sect. 1.2); Cf. (Aquinas, *n.d.,* ST IaIIae 94, 4; 94, 6) [↑](#footnote-ref-27)
28. N.B: I am not arguing *for* this theory, nor suggesting this is a full account of it. My objective here is providing reasons to doubt that Harman can infer implicit agreements best explain (1) and (2) when other accounts exist, and these have not been reasonably rejected. [↑](#footnote-ref-28)