A Computer Simulation of the Argument from Disagreement

Johan E. Gustafsson and Martin Peterson

January 18, 2010

Abstract. We offer new support to the Argument from Disagreement by putting it to test in a computer simulation. According to this argument, widespread and persistent disagreement on ethical issues indicate that there are no objective and epistemically accessible moral facts. Surprisingly, our simulation shows that if our moral opinions were influenced at least a little bit by moral facts, we would quickly have reached consensus, even if our moral opinions were affected by factors such a false authorities, external political shifts, and random processes. Therefore, since no such consensus has been reached, the simulation gives us increased reason to take seriously the Argument from Disagreement.

Keywords. Hegselmann-Krause; disagreement; simulation; metaethics; moral realism; opinion dynamics.

In this paper we offer new support for the Argument from Disagreement by putting it to test in a computer simulation. The Argument from Disagreement holds that there are no objective and epistemically accessible moral facts. This is because if objective and epistemically accessible moral facts were to exist, then there would be no more interpersonal and intercultural disagreement in ethics than in, say, chemistry or mathematics. However, there is in fact much more disagreement in ethics than in other areas, and therefore advocates of the Argument from Disagreement conclude that no such facts exist.

Sextus Empiricus was the first to offer a detailed account of the Argument from Disagreement.[1] In more recent years, J. L. Mackie's discussion has been particularly influential. According to Mackie, "[r]adical differences between first order moral judgements make it difficult to treat those judgements as apprehensions of objective truths." [2] This is because,

> the actual variations in the moral codes are more readily explained by the hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values.[3]

1. *Against the Ethicists*, 68–78, *Outlines of Pyrrhonism* III, 190.
2. Mackie (1977, p. 36).
3. Mackie (1977, p. 37).

Recent discussions of the Argument from Disagreement have revealed at least two major weaknesses. The first is exemplified in Mackie's treatment: He offers no or little support for the claim that variations in moral opinions are more readily explained by the non-realist hypothesis than its realist rival. Why, exactly, is it more reasonable to think that people would stop to disagree if moral facts were to exist? Moral opinions are affected by a large number of factors, including authorities and other opinions held by friends and relatives, so advocates of the Argument from Disagreement need to do more to convince us that the existence of moral facts would invariably lead to consensus.

The second weakness, highlighted by several contemporary moral realists, is that it is difficult to reconcile the Argument from Disagreement with the fact that we actually agree on many moral issues, such as the wrongness of torturing innocent children for fun.[4] More precisely put, it seems that we disagree on *some* issues, such as abortion and capital punishment, but not on all. So how can claims about disagreement disprove the existence of moral facts, given that we agree on many—or even most—moral issues?

The aim of this paper is to show that computer simulations can overcome both weaknesses described above. If correct, the simulations will thus offer increased support to the Argument from Disagreement and, indirectly, the rejection of non-sceptic moral realism. Non-sceptic moral realism is here taken to be the view that there are moral facts that influence our moral opinions.[5] The simulations offer a detailed explanation of why the existence of such moral facts would lead to consensus on moral issues, and why we do actually agree on some moral issues but not on others.

The structure of this article is as follows. In § 1 we explain the basic configuration of our computer model, which is thereafter extended, in § 2, in order to become more realistic. We model a scenario that includes moral facts and compare it to a scenario without moral facts. Since the scenario comprising moral facts turns out to have no or little resemblance with the world we are actually living in, unlike the scenario in which there are no moral facts influencing our moral opinions, we conclude that *either* there is something wrong with the model *or* non-septic moral realism is false. In § 3 we argue that the last disjunct of the foregoing sentence is the correct conclusion. We do this by showing that the model is very robust. Its outcomes look very much the same even if the initial assumptions are modified quite dramatically. Thereafter, in § 5 and § 6, we explain more in detail how our computer simulation overcomes the weaknesses of traditional versions of the Argument from Disagreement outlined above. In these sections we also try to reply to some of the critical comments on our approach we expect to receive.

4. See e.g. Shafer-Landau (2003) and Tersman (2006, p. xii).
5. Cf. the weaker claim that there are moral facts that *might* influence our moral opinions.

1. The basic model

A successful computer model of moral reasoning should be flexible about how moral opinions evolve and change over time. The model has to be able to account for the influence of a number of factors, apart from moral facts, many of which we know relatively little about.

To start with, we wish to make room for the possibility that one's moral opinions are influenced, at least to some degree, by opinions held by others. People who think abortions should remain legal may perhaps pay no attention to the most extreme views put forward by their opponents, but they are likely to take into account at least *some* arguments put forward by people holding opinions that are sufficiently similar to their own. We shall model this by adopting a generalized version of a model originally developed by Hegselmann and Krause.[6] Hegselmann and Krause assume that opinions can vary continuously in an interval of real numbers. It is convenient to let the interval be $[0, 1]$. Since a model is a purely mathematical construction, the outcome of the simulation will not depend on the interpretation of the interval. However, to give the numbers some intuitive meaning, it is helpful to suppose that 1 corresponds to a strong positive opinion about abortion and 0 to a strong negative opinion.

Fig. 1 shows a simulation of a moral discussion between a group of 50 individuals, all of which initially hold different and randomly distributed opinions about abortion. All individuals are aware of the opinions held by others, and in each round they modify their opinions by taking into account views put forward by others. In this example, each individual considers only opinions that are up to 0.3 units more positive or negative than his or her initial opinion. The term *confidence interval* refers to the set of similar views that each agent is prepared to consider. (In the mathematical literature the confidence interval is denoted by the Greek letter ϵ.) In the original version of the Hegselmann-Krause model, developed for non-moral purposes, it is assumed that disagreeing parties should "split the difference", that is, adopt the opinion that corresponds to the average opinion of the set of opinions within the confidence interval. However, in order to make our model more flexible—and thus less vulnerable to criticism—we shall assume that agents are allowed to be more conservative, in the sense that they may put more weight on their own opinion than those of others. In Fig. 1 each individual consider his own opinion to count for ten times as much as opinions held by others. Nevertheless, despite this strong conservatism, the group quickly reaches consensus. This is due to the relatively large confidence interval.

A striking fact about the Hegselmann-Krause model is that the outcome of the simulation is determined solely by each agent's willingness to adjust his or her opinion to similar

6. Hegselmann and Krause (2002).

opinions held by others. If each agent's opinion at each round is influenced *only* by his initial view at the beginning of the round and other views that are within the confidence interval, i.e. views that are sufficiently close to one's own for being worth considering, then consensus will arise just in case the confidence interval is sufficiently large in relation to the size of the population.

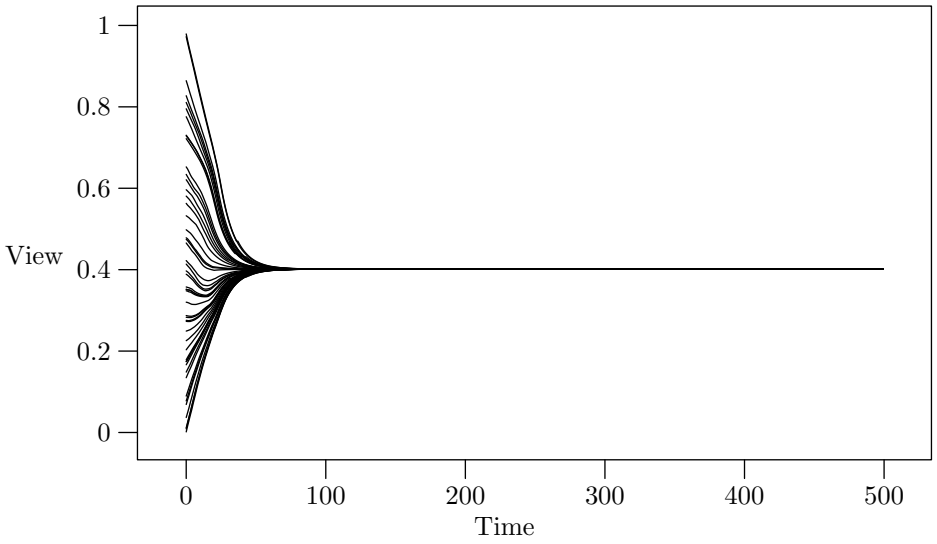

Figure 1: $\epsilon = 0.6$

In Fig. 2 and 3 the population is kept constant while the confidence interval is decreased. In Fig. 2 each individual considers opinions that are up to 0.2 units more positive or negative than his own, and in Fig. 3 the confidence interval is decreased to 0.1 units. Hegselmann and Krause refer to the three types of cases illustrated in Fig. 1–3 as *consensus*, *polarization*, and *plurality*. This strong correlation between the number of surviving opinions and the size of the confidence interval is typical of the Hegselmann-Krause model. The relationship is, however, not monotonic. It has been shown that in some intervals the number of surviving opinions increases as the confidence interval increases.[7] Furthermore, Hegselmann and Krause show that patterns similar to those in Fig. 1–3 arise even if the confidence interval is asymmetric, i.e. if the upward limit of the confidence interval is not equally far from the initial opinion as the downward limit.

A number of mathematicians have obtained analytic results about the Hegselmann-Krause model. Krause (2008) proves a convergence theorem for a limited number of cases, and Lorenz (2007) and Lorenz and Lorenz (2008) prove generalized versions of the theorem. It has furthermore been hypothesized, but not yet proved, that for every very small

7. Lorenz (2007).

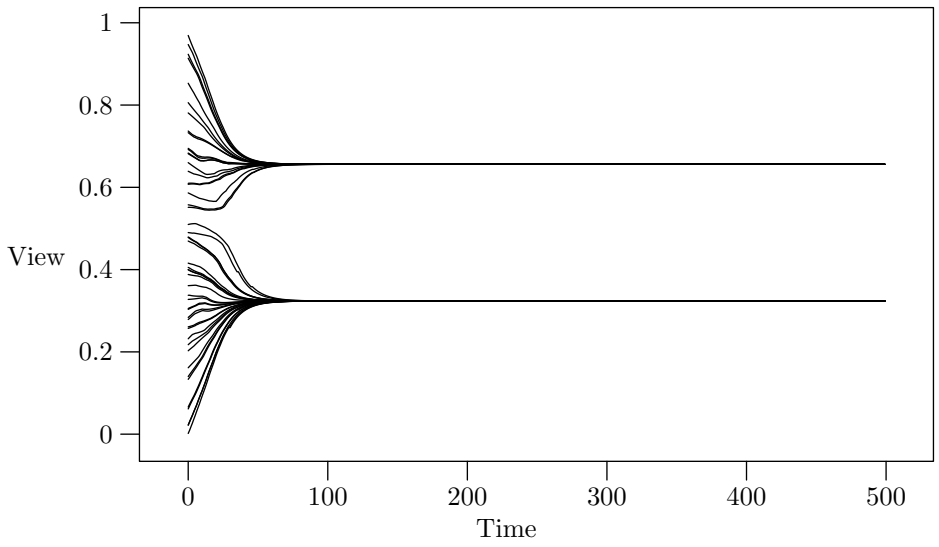

Figure 2: $\epsilon = 0.4$

confidence interval consensus will arise given that the size of the population is sufficiently large. The present paper makes no attempt to add new analytic results to the literature on the Hegselmann-Krause model.

Arguably, it is a striking feature of moral discussions on abortion, capital punishment, and a number of similar issues, that we never seem to reach consensus. It is therefore interesting to study what structural assumptions about moral discussions, if any, are consistent with persistent disagreement. In particular, it is interesting to investigate whether persistent disagreement is compatible with the assumption that moral opinions are affected by moral facts that we somehow perceive or get acquainted with.

2. The extended model

In the basic model persistent moral disagreement is possible only if the confidence interval is sufficiently small. Otherwise consensus will quickly be reached, as shown above. Therefore, in the rest of this paper, we assume that the confidence interval is sufficiently small to enable persistent disagreement, i.e. we will only study cases similar to that in Fig. 3. Although it cannot be excluded *a priori* that persistent disagreement could arise for larger confidence intervals once more features are added to the model, we have found that this is not the case. (For details, see § 3.) In order to be on the safe side, we also assume that individuals are even more conservative than before. In the extended model, each individual takes his own view to count for several hundred times as much as his moral peer. Although this is hardly

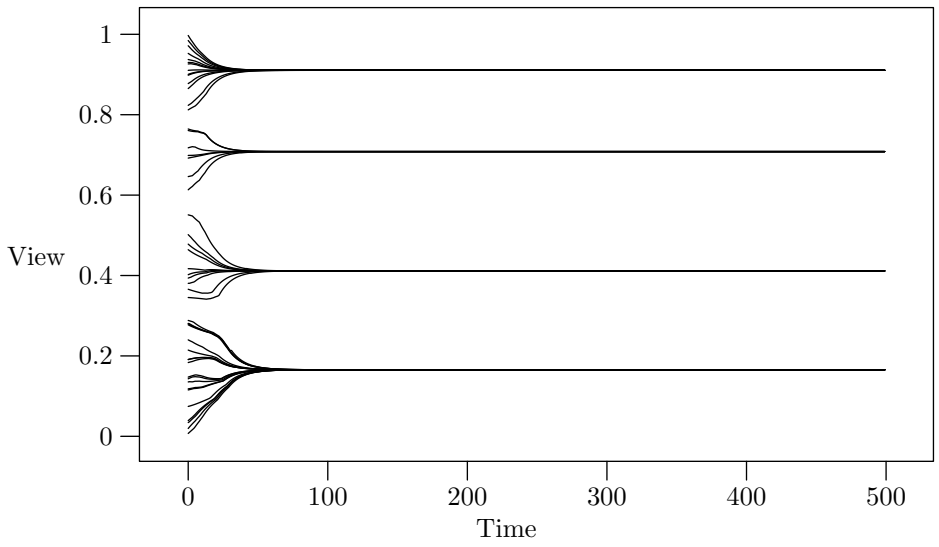

Figure 3: $\epsilon = 0.2$

rational, this makes the model less vulnerable to criticism, as will be shown below.

Needless to say, it would be implausible to assume that moral opinions are affected *merely* by opinions held by others. Other factors may also play a role. In the extended model three such additional factors are incorporated into the simulation, viz. authorities, external shifts, and random processes. Naturally, the more factors that can be incorporated into the model without affecting the overall result of the simulation, the less vulnerable will our conclusion be to criticism. So we do not wish to claim that all moral opinions are always affected by authorities, external shifts, and random processes. The point is, on the contrary, that our conclusion will hold water *even if* these factors affect how moral opinions change over time.

An *authority* can be conceived of as an influential person, organization, political party, book, or philosophical movement. Typical examples include the Pope, Greenpeace, the German Nazi party in the 1930's, Peter Singer's book *Animal Liberation*, and Aristotelian virtue ethics. All these entities might affect our moral opinions, at least to some degree. However, they cannot all be right about what they preach, because on many moral issues they advocate inconsistent views. Furthermore, each authority tends to attract opinions over an extended period of time, but the influence is seldom constant over time. In our model, we assume that each authority, or rather each period in which an authority affects our opinions, extends over a randomized interval of 10 to 30 rounds. (See § 3.) Of course, some authorities, such as the Church, tend to be influential during several periods. As will be shown in § 4 our conclusion will go through even if our moral opinions are influenced by one or more constant moral authorities, i.e. constant attractors of moral opinions that

may differ from the truth. In Fig. 4 and Fig. 5 authorities are represented as relatively short, thick horizontal lines. Contrary to what one might have expected, the presence of a large number of authorities lead to more consensus rather than less. Moreover, our simulations also show that the result will look the same no matter whether the strength of each authority is constant or decreases with respect to the distance of the opinion.

Sometimes *external shifts*, such as the political shift towards the left in 1968, may affect our opinions quite dramatically during a relatively short period of time. From a strict mathematical point of view, authorities and external shifts are very similar: they can be represented as restricted areas of the figure that affect other, nearby opinions. The main difference is that authorities affect our opinions to a limited degree over a relatively long period of time, while external shifts take place within just one or a small number of rounds. However, during those rounds they have a relatively large influence on our moral opinions. In Fig. 4 and 5 external shifts are represented by thick vertical lines.

The third new factor in the extended model is *random processes*. Sometimes opinions change for no good (or bad) reason at all. This may very well be irrational, but since the Argument from Disagreement starts from the assumption that people do *actually* disagree, it seems wise to also take the possibility of random shifts into account. In our extend model we have stipulated that opinions change randomly every 30 to 70 rounds, and that no random shift is more than twice the size of the confidence interval. These numerical assumptions are of course arbitrary. However, as will be shown below, nothing really hinges on them. The robustness analysis presented in § 3 shows that our conclusions does not depend very much on what assumptions we make about the numerical values of our variables.

Fig. 4 shows a simulation in which the three factors introduced above have been added to the Hegselmann-Krause model. The most important conclusion to be drawn here is that no consensus is reached, not even after a large number of rounds. In the next section we discuss the assumptions we have made about the values of the underlying variables and the mechanism for aggregating their influence, and we show that the result in Fig. 4 is surprisingly robust. This means that the figure will look pretty much the same no matter what is assumed about the size of the population (relative to the confidence interval), the number of authorities, their position on the scale of opinions, their extension in time, the number of external shifts, their strength, and the influence of random processes.

Enter moral facts. We shall take for granted that if they were to exist, they would somehow affect our moral opinions. Furthermore, if moral facts were to exist, they would be invariant with respect to time—they would always be there.[8] In the extended model these

8. This assumption must be interpreted with care. Naturally, similar acts may have different moral status at different points in time due to changes in the external world.

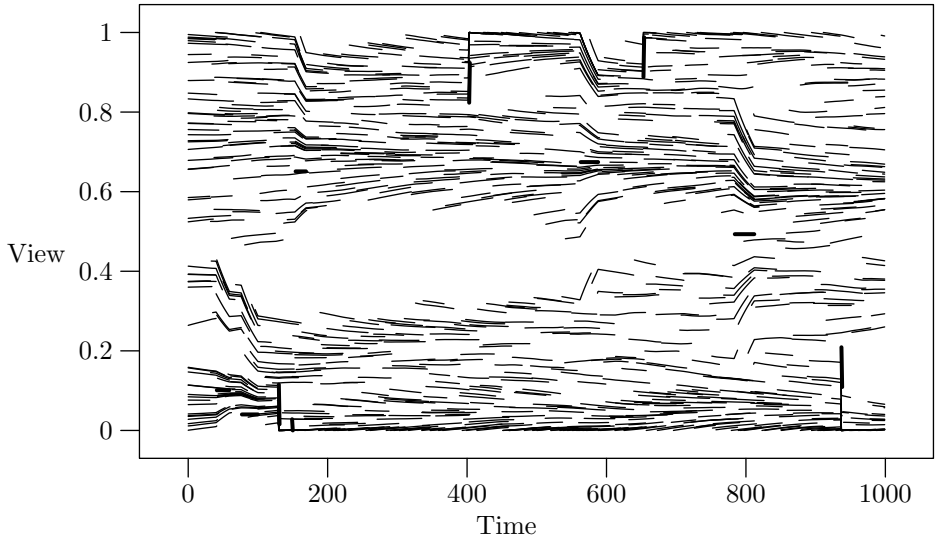

Figure 4: The extended model without moral facts

assumptions will be modelled by assuming that moral facts attract our opinions very much in the same way as authorities do. This is to say, they exercise moderate influence on our opinions over an extended period of time. For simplicity, we assume that the influence of moral facts is constant over time. Furthermore, we assume that there is only one fact of the matter about each moral issue; hence, it cannot be true that, say, abortion is both entirely right and entirely wrong at the same point in time.

A moral fact is represented by a dotted line in our plots. In Fig. 5 we assume that the objectively true opinion is, say, to fully reject abortions. To start with, we assume that moral facts affect our moral opinions to the same degree as authorities, i.e. to the same degree as the Church or Peter Singer. Fig. 5 clearly shows that if moral facts were to have such a great influence on our opinions, then consensus would have been reached within a relatively small number of rounds. Suppose that a "round" in the model is equivalent to, say, one month of discussion. Then consensus on abortion would have been reached within a few years after the discussion began. This is, however, not the case. Abortion is still a very controversial moral issue, on which people hold very different moral views—despite the fact that free and democratic discussions have been going on for at least a century. Therefore, the assumption about the strong influence of moral facts on our opinions made in Fig. 5 is hardly a correct representation of what the world is really like.

But what if moral facts have far less influence on our moral opinions? In Fig. 6 the influence of moral facts is less than 0.5 % of the total influence (as defined in the next section) put on our moral opinions. Despite this very low sensitivity to moral facts, it nevertheless

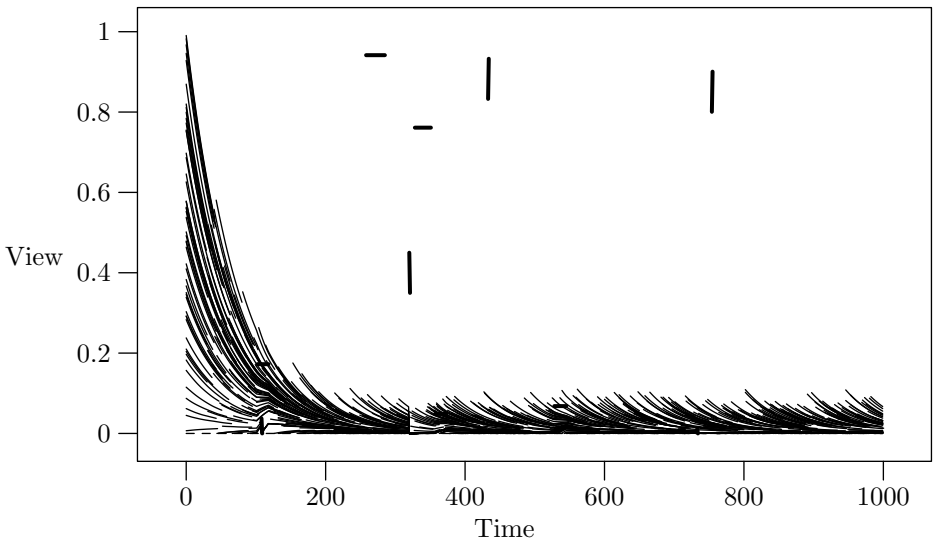

Figure 5: The extended model with strong moral facts

turns out that consensus will be reached within a moderate number of rounds. Moreover, this conclusion also holds true if the number of authorities is dramatically increased, (to 50 rather than 5) as can be seen in Fig 7. Therefore, since no consensus has yet been reached on abortion, it follows that moral facts about abortion do not exist, or at least do not affect our opinions to any significant degree.

3. Further details about the extended model

In the foregoing sections we have tried to give a brief and non-technical overview of how our model works and why it is relevant to the Argument from Disagreement. The main point is that our moral opinions are not affected by any objective moral facts, because if they were, we would have reached consensus a long time ago. In this section we offer a detailed presentation of the technical assumptions underlying the model, and discuss their plausibility. The basic message is that the model is surprisingly robust. Our conclusion will not change dramatically if the simulation is repeated a large number of times, or if the initial values in each simulation are altered.

Opinions are represented by real numbers, which are updated in each round depending on the distance to other opinions, authorities, external shifts, random processes, and facts. Time is measured through the number of rounds the process is iterated. It is of course difficult to tell exactly how long each round lasts. However, in order to be on the safe side, we assume that each round lasts no longer than one year, i.e. that individuals are willing

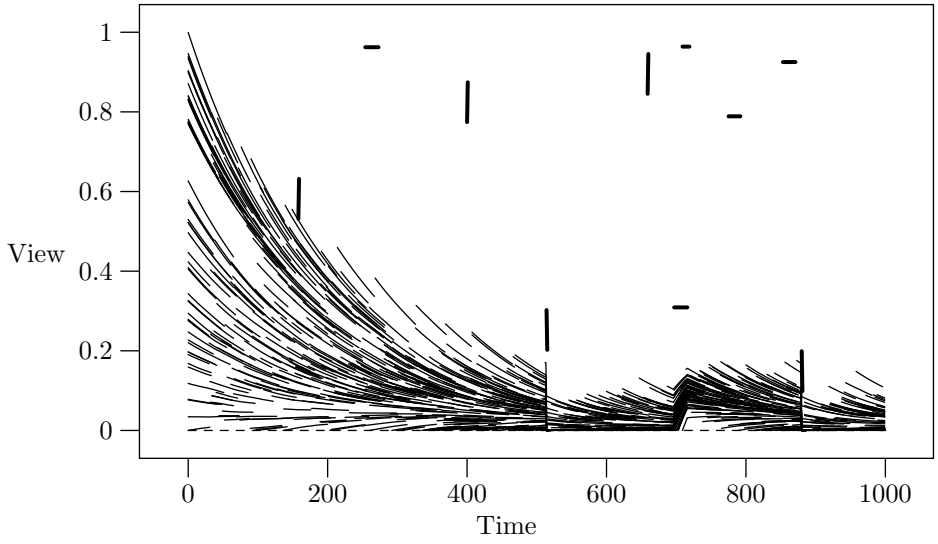

Figure 6: The extended model with weak moral facts

to reconsider their moral opinions at least once a year. Arguably, this is a very conservative assumption. If opinions are actually reconsidered more frequently, say once a month (as we hypothesized in § 2), our case becomes even stronger. This is because consensus would then be reached much quicker. If each round were to last much longer, then people would be too conservative for the model to be realistic. Therefore, it is not a problem that the unit of time is somewhat arbitrary, as long as the assumptions we base the simulations on are sufficiently conservative.

As explained in the foregoing section, authorities are denoted by thick horizontal lines, whereas external shifts are denoted by vertical lines. However, from a technical point of view they share the same structure; they can be conceived of as *attractors* that attract opinions over a shorter or longer period of time. Moral facts are different, in that they attract opinions continuously. The strength of each attractor is measured as the percentage of the overall influence each entity has on the individual's new opinion.[9] Hence, the claim that convergence emerges if moral facts m account for at least 0.5 % of the total influence means that if we have a list of attractors $a, b, c, \ldots m$, then

$$\frac{m}{a+b+c+\ldots+m} \geq 0.005$$

We are now in a position to describe the algorithm that was used for performing the

9. We are aware that this makes the measure of strength dependent on each agent's degree of conservatism. Of course, it is possible to propose other measures of strength that do not rely crucially on how conservative the agent is.

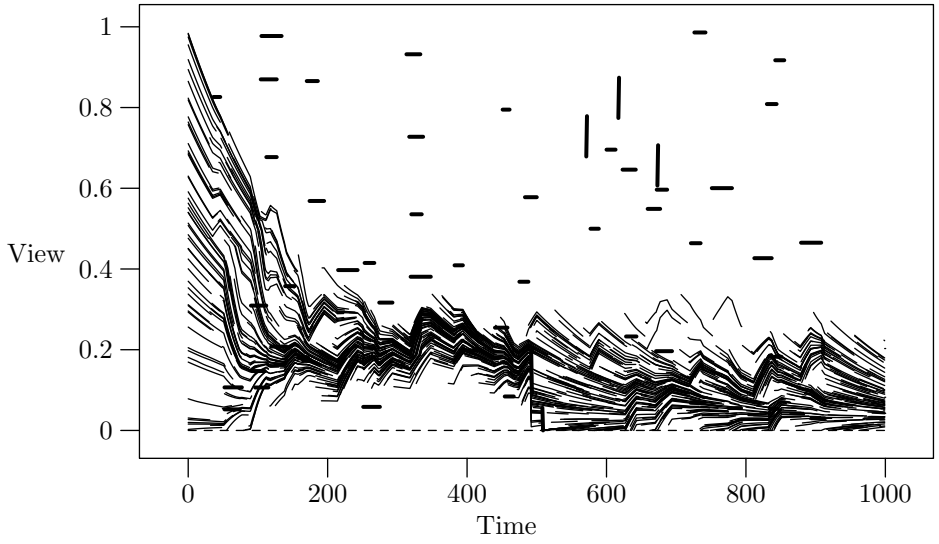

Figure 7: The extended model with weak moral facts and many authorities

simulations. (The names of the parameters should be self-explanatory, but see Appendix A for detailed explanations.)

The Simulation Algorithm for the Extended Model: Given the parameters defined in Appendix A, simulate and plot the changes in views for n persons over R rounds.

1. [Initialize round count] Set r to 1.
2. [Initialize persons.] Generate the initial views and rounds left to live for the n persons in round 1, with the view randomized between 0 and 1 and the rounds left to live randomized between *personMinAge* and *personMaxAge*. For all persons P_i, store the rounds left to live in $roundsLeftToLive_i$.
3. [Generate authorities.] Generate *authoritiesCount* authorities with a duration randomized between *authorityMinDuration* and *authorityMaxDuration* beginning at a round randomized between 1 and R.
4. [Generate shifts.] Generate *shiftsCount* shifts at a round randomized between 1 and R originating from a view randomized between 0 and 1, with a positive direction if the it originates from a view over 0.5, otherwise it has a negative direction.
5. [Increase round count] Increase r by 1.
6. [Conservatism] For all persons P_i, set $conservatismPoint_i$ to the view P_i had last round.
7. [Calculate influence from friends] For all persons P_i, calculate the viewpoint their friends influence them toward, and store it in $friendsPoint_i$. This viewpoint is cal-

culated as the average view last round of all persons within ϵ of the view P_i had last round. If P_i has no friends then $friendsPoint_i$ is set to the view P_i had last round.

8. [Calculate influence from authorities] For all persons P_i, calculate the viewpoint active authorities influence them toward, and store it in $authoritiesPoint_i$. This viewpoint is calculated as the average view of all authorities active this round within *authoritiesRange* of the view P_i had last round. The number of active authorities within *authoritiesRange* of the view P_i had last round is stored in $activeAuthorities_i$.

9. [Calculate influence from shifts] A person p_i is within range of a shift with negative direction if $conservatismPoint_i$ is less than the view the shift originates from. Similarly, a person p_i is within range of a shift with positive direction if $conservatismPoint_i$ is greater than the view the shift originates from. For all persons P_i set $shiftsPoint_i$ to 0. For all persons P_i and for all shifts during this round that P_i is within range of, increase $shiftsPoint_i$ with the view shifted *shiftLength* in the direction of the shift from $conservatismPoint_i$. For all persons P_i, set $activeShifts_i$ to the number of shifts during this round that P_i is within range of.

10. [Age persons] For all persons p_i, decrease $roundsLeftToLive_i$ by 1.

11. [Calculate new view for survivors] For all persons P_i, if $roundsLeftToLive > 0$ then calculate the the view for P_i with the following formula:
$(conservatism * conservatismPoint_i + friendsInfluence * friendsPoint_i$
$+ authorityInfluence * authoritiesPoint_i + factInfluence * fact + shiftInflunce * shiftsPoint_i)$
$/(conservatism + friendsInfluence + authorityInfluence * activeAuthorities_i$
$+ shiftInfluence * activeShifts_i + factInfluence)$

12. [Generate successors] For all persons P_i, if $roundsLeftToLive = 0$ then generate successor for P_i. To generate a successor set the new view for P_i to a view randomized within *mutationRange* of the view P_i had last round, and set $roundsLeftToLive_i$ to a number randomized between *personMinAge* and *personMaxAge*.

13. [Another round?] If $r < R$ then go back to step 5 otherwise continue to step 14.

14. [Plot] For all persons P_i, plot the view of P_i for rounds 1–R. ❙

For the simulation visualized in Fig. 4 we set the values of the parameters listed in this footnote.[10] For the simulations visualized in Fig. 5 and Fig. 6 exactly the same parameters were used, except that *factInfluence* was set to 15 instead of 0 in Fig. 5, and set to 5 instead

10. $R = 1000$, $n = 50$, $authorityCount = 5$, $shiftCount = 5$, $fact = 0$, $conservatism = 1000$, $friendsInfluence = 5$, $authorityInfluence = 15$, $shiftInfluence = 5000$, $factInfluence = 0$, $authoritiesRange = 0.35$, $\epsilon = 0.1$, $mutationRange = 0.15$, $shiftLength = 0.1$, $personMinAge = 30$, $personMaxAge = 70$, $authorityMinDuration = 10$, $authorityMaxDuration = 30$.

of 0 in Fig. 6.

As emphasized above, the values of the parameters are to some extent arbitrary. Therefore, in order to neutralize the objection that our results depend too heavily on these values, we have performed a series of simulations that help to assess the robustness of our results. Briefly put, it turned out that even if the values of the parameters are significantly altered, the outcome of the simulations will be more or less the same. We iterated the simulation 10 000 times, under the assumptions that moral facts exist, each time with new arbitrarily chosen values that were up to fifty percent higher or lower than the initial values described above. At the end of each simulation, the distance between the opinions being furthest away was measured. Fig. 8 shows the relative number of times the distance between the opinions being furthest away were between 0 to 0.1 units, and between 0.1 to 0.2 units, and so on. Fig. 9 shows the same plot under the assumption that moral facts do not exist. It seems undeniable that the non-existence of moral facts is a necessary condition for persistent moral disagreement.

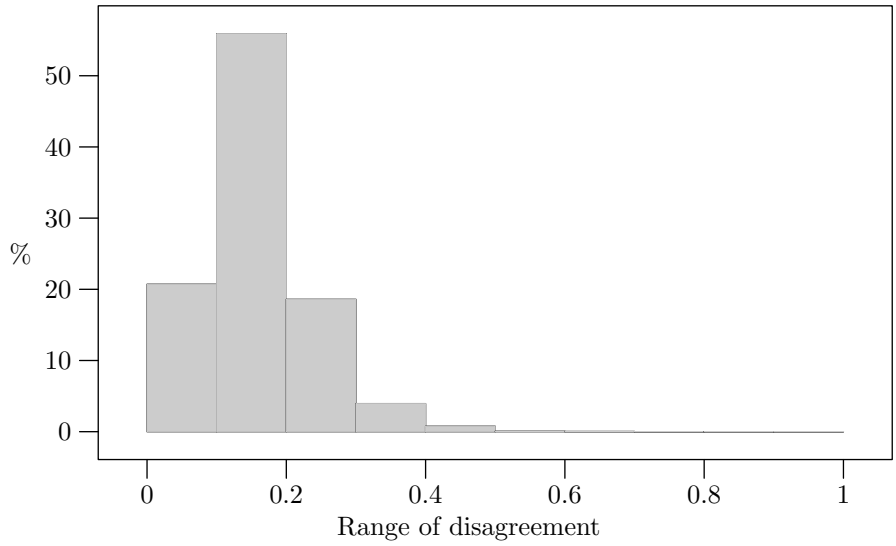

Figure 8: Range of disagreement after 1000 rounds in 10 000 randomized simulations with moral facts

4. Constant authorities

In § 2 we briefly mentioned that some authorities, such as the Church, may influence our moral opinions over very long periods of time. From a modelling perspective, the Church can be conceived as a *constant* authority, i.e. as an authority that affects our opinions in all

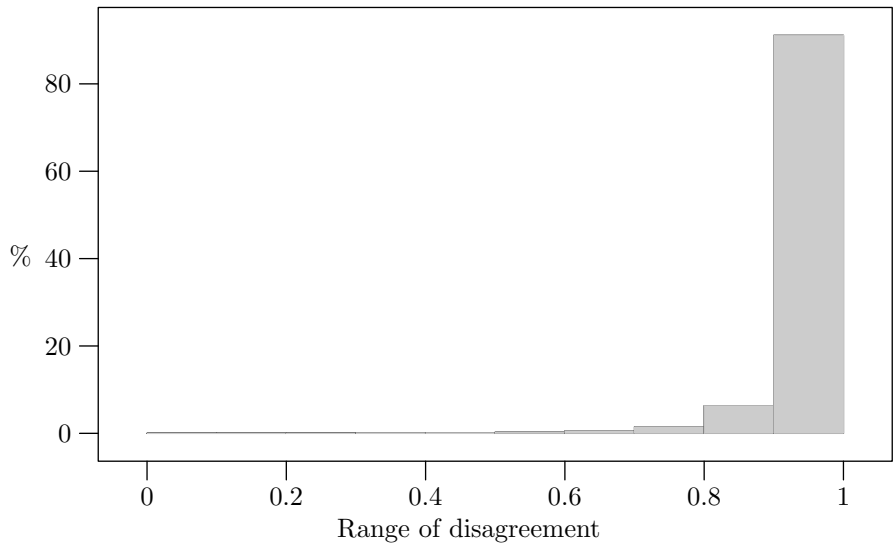

Figure 9: Range of disagreement after 1000 rounds in 10 000 randomized simulations without moral facts

rounds. How does the presence of one or more constant authorities affect the robustness of our conclusion?

Fig. 10 shows a simulation with a moral fact at 0 and a constant authority near 1. A possible example could be a situation in which it is, say, wrong not to allow gay marriages although the Church holds the opposite view. Naturally, at most two opinions will survive in the long run, no matter the relative strength of the fact and the constant authority. (In this particular example, the second surviving opinion is just below 1. This is due to the relative strength of the moral fact and the opinion preached by the Church.) However, this is clearly not a correct account of how moral opinions evolve over time. Our opinions do not tend to converge to two extreme views. On the contrary, nearly *all* opinions between the two extremes tend to be represented when people disagree on moral issues. Moreover, in Fig. 10 all change is change towards an extreme view, and once that view has been reached the individual never departs from it in the future. This is implausible. Therefore, Fig. 10 is not a convincing picture of how people do actually reason about moral issues. This indicates that the non-sceptic moral realist has no reason for thinking that the disagreement we observe is due to the existence of some (possibly false) constant authorities, such as the Church.[11]

11. For the simulation in Fig. 10 the following parameters where used: $R = 1000$, $n = 50$, *authorityCount* = 5, *shiftCount* = 5, *fact* = 0, *conservatism* = 1000, *friendsInfluence* = 5, *authorityInfluence* = 15, *shiftInfluence* = 5000, *factInfluence* = 5, *authoritiesRange* = 0.35, $\epsilon = 0.1$, *mutationRange* = 0.15, *shiftLength* = 0.1, *personMinAge* = 30, *personMaxAge* = 70, *authorityMinDuration* = 10, *authorityMaxDuration* = 30, *constantAuthorityCount* = 1, *constantAuthorityInfluence* = 50, *constantAuthorityRange* = 0.35.

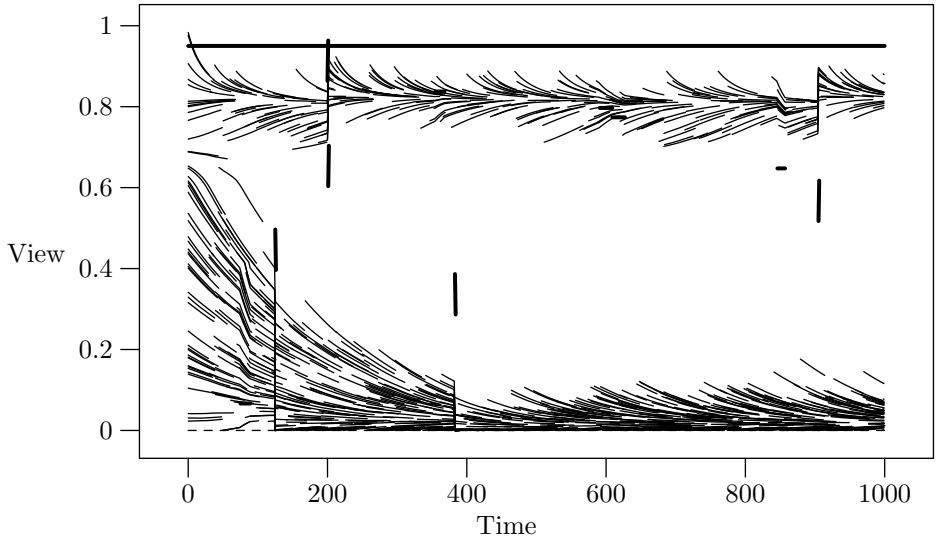

Figure 10: Disagreement and facts with a constant authority

At this point it might be objected that non-sceptic moral realists could try to make their claim about moral facts compatible with the prevalence of widespread moral disagreement by simply increasing the number of constant authorities. Fig. 11 shows a scenario with seven constant authorities and a single fact. Note that in this example no individual will ever discover the moral fact (at 0) due to the strong influence of the constant authorities. Unfortunately, this figure also fails to give a convincing account of how moral opinions evolve over time.[12] Although we can now observe considerable disagreement, people in this simulation simply never *change* their opinions as much as they sometimes do in real life. Moreover, in Fig. 11 all change is change towards a constant authority, and once the individual holds the same view as the authority, he or she will never depart from it in the future. This is implausible. We therefore conclude that this is not a convincing non-sceptic moral realist picture of moral disagreement.

Finally, let us consider yet another way in which the non-sceptic moral realist could try to rescue his position. This time we assume that constant authorities are much less influential than before and that people are more likely to randomly change their opinions; see Fig. 12.[13] During the first couple of rounds this figure appears to give a plausible account of how disagreement could be possible despite the existence of an objective moral fact that

12. For the simulation in Fig. 11 we used the same parameters as in Fig. 10, but with the following modifications: *mutationRange* = 0.08, *constantAuthorityCount* = 7, *constantAuthorityInfluence* = 500, *constantAuthorityRange* = 0.08.

13. For the simulation in Fig. 12 we used the same parameters as in Fig. 11, but with the following modifications: *mutationRange* = 0.15, *constantAuthorityInfluence* = 150.

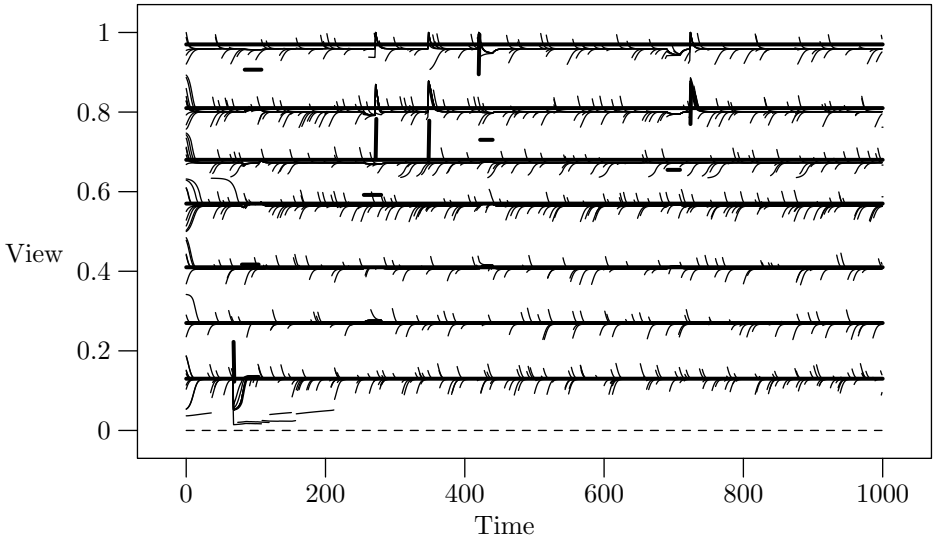

Figure 11: Plurality and facts with many constant authorities

influence people's opinions. As can be seen in the figure, people do actually change their opinions over time. However, the problem for the non-sceptic moral realist is now a different one: After a few hundred rounds more and more people are drawn towards the truth. Note that the entire north-east quarter of Fig. 12 is empty. Again, this is not a convincing picture of moral disagreement. In moral debates on abortion and capital punishment nearly *all* moral opinions are represented, contrary to what we are told in Fig. 12.

The upshot of all this is that non-sceptic moral realists have no reason for thinking that the widespread and persistent moral disagreement we observe in our everyday life is due to the influence of constant authorities. If such constant authorities were to exist they would indeed give rise to some disagreement, but the structure of the disagreement would not be the right one.

5. A better foundation for the Argument from Disagreement?

In the introduction we identified two weaknesses of traditional treatments of the Argument from Disagreement, namely, (i) the absence of an explanation of why the non-existence of moral facts leads to moral disagreement, and (ii) the absence of an explanation of why we disagree on some moral issues but not on all.

The simulations presented in this paper offer new insights into both these issues. Our simulations clearly show why the absence of moral facts is a necessary although not a sufficient condition for persistent disagreement (of the right structure). If moral facts were to

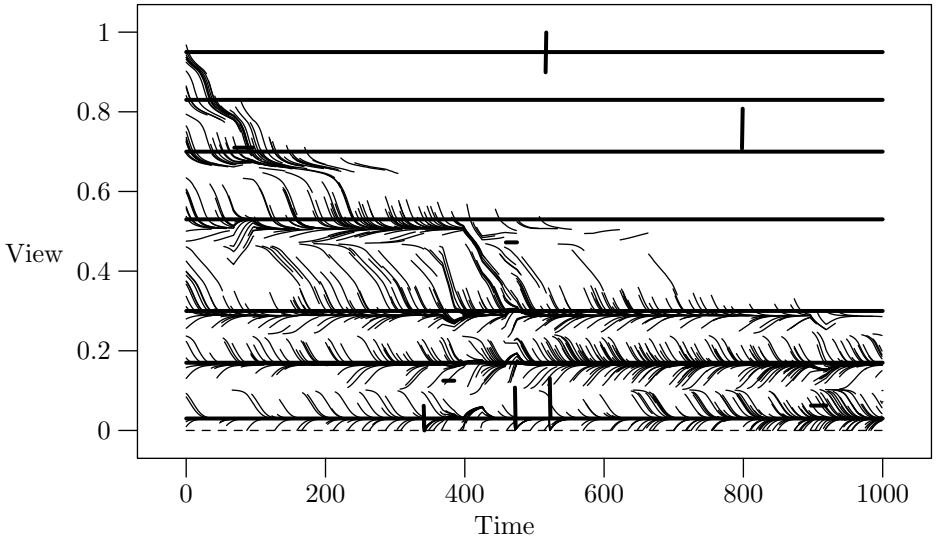

Figure 12: Constant authorities and changing views leads to consensus

exist, and if our opinions have long been affected by them to at least some degree (0.5% or more in each round), then consensus would have emerged quickly. Moreover, if our moral opinions were also affected by some constant authorities the resulting disagreement would not have the right structure, as explained in § 4. This means that the basic premise of the Argument from Disagreement has gained additional support by our simulations: If moral facts existed we would not disagree widely on controversial issues, such as abortion and capital punishment. However, as we all know, we disagree widely on many such issues. This justifies the conclusion that it cannot be the case that moral facts exist and somehow affect our opinions.

The observation that the absence of moral facts is merely a necessary and not a sufficient condition for persistent disagreement also helps to explain why we disagree on some moral issues but not on all. The point is that consensus can arise even if no moral facts exist. The simulations have revealed a number of such consensus-inducing mechanisms apart from moral facts.

Having said all this, it might be objected that our model proves too much. There is also considerable disagreement about whether there is life in other galaxies, whether there is a god, and on exactly what Aristotle actually tried to tell us about the virtues. So does not our model predict that agreement should have emerged a long time ago on all these controversial issues? Our reply is that our model merely predicts that we would quickly reach consensus *if* our opinions were affected at least to some degree towards the fact of the matter. In the examples mentioned above most of us our opinions are presumably not affects positively by

the fact of the matter. Of course, this does not imply that there cannot be any true answer to these questions. All we are entitled to conclude from our simulations is that the truth is positively affect our opinions whenever there is persistent disagreement.

One might also object that our computer simulations are self-defeating.[14] This is because we seem to disagree not just on issues related to normative ethics, but also on many issues in metaethics, including the question whether non-sceptic moral realism is true. So if our model is correct, then there cannot be any fact about whether non-sceptic moral realism is correct that influences our metaethical opinions, since such a fact would have lead to agreement. One might therefore wonder how our model could support this metaethical fact. The answer is that our model merely implies that the fact about whether non-sceptic moral realism is correct has not affected our opinions in the past. This does not rule out that our ability to let our future opinions be affected by the relevant facts might increase as we adopt new methodological approaches to metaethics. One such new approach is the computer simulations presented in this paper. Therefore, our simulation is not self-defeating.

The same move is of course open to the moral realist. While our simulations might have shown that there does not exist any moral facts that have influenced our moral opinions in the past, there might exist moral facts that will do so in the future. Perhaps some new methods in normative ethics will enable us to access the truth. It seems unlikely that any argument from past disagreement can rule out this possibility.

6. Conclusion

Showing how the Argument from Disagreement can be modelled in a computer simulation, brings more clarity to this influential argument. As explained above, we believe that our findings lend additional support to the rejection of non-sceptic moral realism, that is, that there does not exist moral facts that influence our moral opinions. However, the most important conclusion is perhaps that the methodology we use appears to be fruitful for moral philosophers wishing to discuss the metaethical significance of moral disagreement. By giving up some of the seemingly plausible assumptions we use in our model, it may very well be possible to construct *some* version of moral realism that is not sensitive to the Argument from Disagreement—if so, this would then tell us something important about the structure of a plausible form of moral realism. We thus conclude that the computer has given us new insights into the structural limitations of the debate over moral realism in metaethics.

14. This objection was suggested to us by Jonas Olson.

Appendix A: Definitions of parameters

authorityCount	the number of authorities in the simulation
shiftCount	the number of shifts in the simulation
constantAuthorityCount	the number of constant authorities in the simulation
personMinAge	the minimum length of a persons life
personMaxAge	the maximum length of a persons life
authorityMinDuration	the minimum duration of an authority
authorityMaxDuration	the maximum duration of an authority
fact	the fact of the matter
ϵ	the maximum distance in views over which persons influence each other
authoritiesRange	the maximum distance in views over which an authority attracts
constantAuthorityRange	the maximum distance in views over which a constant authority attracts
friendsInfluence	the factor of how much persons are influenced by persons within ϵ
authorityInfluence	the factor of how much persons are influenced by an authority within *authoritiesRange*
constantAuthorityInfluence	the factor of how much persons are maximally influenced by a constant authority within *constantAuthoritiesRange*
conservatism	the factor for how much persons are influenced by their view from last round
shiftInfluence	the factor of how much persons are influenced by a shift
factInfluence	the factor of how much one is influenced by the fact of the matter
mutationRange	the maximum distance in views between predecessor and successor
shiftLength	the distance away that a shift pulls to

References

Hegselmann, R. and U. Krause: 2002, 'Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation'. *Journal of Artifical Societies and Social Simulation (JASSS)*

5(3).

Krause, U.: 2008, 'Compromise, consensus and the iteration of means'. *Elemente der Mathematik* **63**, 1–8.

Lorenz, J.: 2007, 'Repeated Averaging and Bounded Confidence Modeling, Analysis and Simulation of Continuous Opinion Dynamics'. Ph.D. thesis, Universität Bremen.

Lorenz, J. and D. A. Lorenz: 2008, 'On Conditions for Convergence to Consensus'. *Arxiv preprint arXiv:0803.2211*.

Mackie, J. L.: 1977, *Ethics: Inventing Right and Wrong*. Penguin.

Shafer-Landau, R.: 2003, *Moral Realism: A Defence*. Oxford University Press.

Tersman, F.: 2006, *Moral Disagreement*. Cambridge University Press.