



Don't forget forgetting: the social epistemic importance of how we forget

Daniel J. Singer¹ · Aaron Bramson^{2,3,4} · Patrick Grim^{5,6} · Bennett Holman^{7,8} · Karen Kovaka⁹ · Jiin Jung¹⁰ · William J. Berger¹

Received: 30 November 2018 / Accepted: 24 September 2019
© Springer Nature B.V. 2019

Abstract

We motivate a picture of social epistemology that sees forgetting as subject to epistemic evaluation. Using computer simulations of a simple agent-based model, we show that how agents forget can have as large an impact on group epistemic outcomes as how they share information. But, how we forget, unlike how we form beliefs, isn't typically taken to be the sort of thing that can be epistemically rational or justified. We consider what we take to be the most promising argument for this claim and find it lacking. We conclude that understanding how agents forget should be as central to social epistemology as understanding how agents form beliefs and share information with others.

Keywords Agent-based modelling · Forgetting · Group deliberation · Social epistemology

We're very thankful for the feedback on aspects of this research from H el ene Landemore, Joe Halpern, and audiences at the University of Michigan, Dartmouth College, and the Wharton School at the University of Pennsylvania.

✉ Daniel J. Singer
singerd@phil.upenn.edu

¹ University of Pennsylvania, Philadelphia, USA

² Riken Brain Science Institute, Wakoshi, Japan

³ Ghent University, Ghent, Belgium

⁴ University of North Carolina at Charlotte, Charlotte, USA

⁵ University of Michigan, Ann Arbor, USA

⁶ Stony Brook University, Stony Brook, USA

⁷ Underwood International College, Yonsei University, Seoul, Korea

⁸ Faculty of Humanities, University of Johannesburg, Johannesburg, South Africa

⁹ Virginia Polytechnic Institute and State University, Blacksburg, USA

¹⁰ Claremont Graduate University, Claremont, USA

“...gaining knowledge and losing knowledge are natural duals, and indeed, forgetting is crucial to one’s sanity.”—Johan van Benthem (2011, p. 109).

1 Introduction

It’s a normal part of our epistemic practice to see what others believe as rational and justified, but it’s not standard to see what others forget as rational or justified. As Moss (2015, p. 174) puts it, if you forgot what you had for dinner last night, it would be odd to rebuke you with “how *irrational* of you!”. Williamson (2000, p. 219) agrees: “forgetting is not irrational; it is just unfortunate.”¹ The primary aim of this paper is to motivate the idea that, when it comes to social epistemology, believing and forgetting shouldn’t be so different. How we forget should be a topic that’s discussed alongside rational and justified belief, testimony, evidence, knowledge and the impacts of social structural features of our communities on what we know.

Using computer simulations of a simple model of group deliberation, we show that how we forget can have a profound effect on whether the groups we’re in can achieve optimal epistemic outcomes—an effect that’s at least on par with other paradigmatic topics in social epistemology, such as how others share their information. We’re inclined to see these effects as symptoms of forgetting being more or less rational or justified.² So, before we demonstrate how important forgetting is for social epistemic outcomes, we’ll first consider what we take to be the most promising argument against the idea that forgetting can be rational or justified. We’ll show that that argument is no stronger against forgetting being rational or justified than it is against believing being rational or justified. For this reason, we suggest that what we forget, like what we believe and how we share information, should be a central topic in social epistemology.

2 On rational or justified forgetting

What motivates the idea that what we forget can’t be rational or justified? A natural thought is that forgetting can’t be rational or justified because it’s outside our control. Whether or not we forget something doesn’t seem like something we can do at will. I can’t simply choose to forget what I ate for dinner last night. So, how we forget cannot be rational or justified if being rational and justified requires that kind of control.

¹ Outside of epistemology, it’s more common to see forgetting as the sort of thing that can be subject to rational criticism. Angela Smith’s (2005) paper about responsibility for attitudes, for example, has as its central example a case in which someone is blameworthy for forgetting a close friend’s birthday. It’s most natural to conceive of that case as one of practical or moral responsibility, rather than epistemic. Here we’ll restrict ourselves to epistemic evaluations of forgetting.

² This suggests that we’re committed to a consequentialist picture of epistemic normativity. While the primary author does explicitly endorse such a view, one need not be a consequentialist to feel the pull of this idea. One only needs to see the epistemic impacts as a *symptom* of the rationality of the forgetting method, not its (constituent) cause.

This argument for thinking forgetting can't be rational or justified closely mirrors Alston's (1988) discussion of the "deontological conception" of justified belief. Alston worried that a conception of epistemic justification in terms of "obligation, permission, requirement, blame, and the like" was untenable because we could only be justified in that sense if we control what we believe, and Alston thought we don't.

We take it that something about Alston's argument must be wrong since belief can be both rational and justified. In light of that, the analogous problem for forgetting shouldn't seem so compelling, since our control over what we forget seems to be on par with our control over what we believe. We can neither directly will ourselves to believe grass is purple nor directly will ourselves to forget that grass is green. But we can indirectly get ourselves to believe in God by surrounding ourselves with believers, and we can equally get ourselves to forget the pains of yesterday by distracting ourselves with the pleasures of today. Moreover, even though most believing and forgetting just happens to us in the sense that it happens without forethought or consideration, we do sometimes use strategies for both belief formation and forgetting.³ Just as a doctor can train herself to recognize a specific pattern of symptoms or a salesperson can train himself to recognize a likely client, we can train ourselves to remember certain kinds of facts and not others. We have strategies for this, which include repeating important facts to ourselves, writing them down, etc. Given those considerations, one shouldn't think that the control argument is any more compelling when applied to forgetting than to believing.

Moreover, many of the responses to Alston's argument against the deontological conception of justified belief work equally well against the analogous argument about forgetting. Feldman (2000), for example, denies that all obligations, permissions, and responsibilities require control.⁴ Feldman thinks that epistemic obligations are what he calls "role oughts"—'oughts' that apply to us in virtue of some role we occupy. Teachers and parents have certain obligations—to explain things clearly and protect their children, for example—that they're subject to independently of whether they are able to fulfill them. Feldman thinks that we can similarly make sense of belief being rational and justified since "forming beliefs is something people do" and "Anyone engaged in this activity ought to do it right" (88). That reasoning applies equally well to forgetting as it does to believing, though. Remembering some things while forgetting others is something we do to manage our limited memory, just like explaining is something we do to manage our teaching obligations. Why not think a role ought applies to us in our role as agents with limited memory?

Whereas Feldman responds to Alston by denying that epistemic obligations require control, another option is to claim that we do control (at least many of) our doxastic states. Ginet (2001), Hieronymi (2006), Steup (2008, 2011), and Weatherston (2008) all take some version of this line. Steup (2011), for example, argues that if we conceive of the relevant freedom in a compatibilist way, it looks like our

³ For more discussion of the kind of control we have over forgetting, see Bernecker (2018). Bernecker (2018) also argues that we can be responsible for what we forget, though the sense of responsibility in play there is a moral one.

⁴ For other examples of this kind of response to Alston, see Ryan (2003), McHugh (2012), and Peels (2017). For a related response, see Vahid (1998).

beliefs are as free as our actions, since both our beliefs and our actions are appropriately reasons-responsive (Fischer and Ravizza 1998).

Is what we forget similarly reasons-responsive? One might think that ideal unlimited agents would not have epistemic reasons to forget.⁵ But, when we're limited in our time or resources, we can have reasons that we wouldn't otherwise have. The fact that today is the last day of your short vacation in Paris might mean that you have a reason to try to see the Musée d'Orsay today. Were your time and energy in Paris literally unlimited, you might not have a particular reason to go today rather than tomorrow. And if you're down to your last dollar for feeding your family, you have a reason to carefully make the decision about which beans are most cost-effective, which is a reason you wouldn't have if you had unlimited funds. The case is similar with our limited memory: given that we can only remember so many things, we have more reason to remember that heavy medium-sized objects tend to fall than to remember the details of an unremarkable dinner three weeks ago.

Do we actually respond to these kinds of reasons? In many cases, it seems like we do. We usually forget details that are unimportant to our practical and epistemic lives (what we ate at the unremarkable dinner, exactly how many stop lights we passed on the way to the market, etc.). And we usually remember as much of what is practically and epistemically important as we can, such as general rules about how objects behave, where we live, and basic facts about how to care for ourselves. So, following much of the same argument that Steup (2011) uses for the voluntariness of belief (in a compatibilist sense), we should see forgetting as similarly voluntary.

More generally, it seems natural to think that if forming a belief can be rational or justified, then forgetting can be too, since forgetting and forming a belief are two sides of the same coin. We often talk about "beliefs" being the things that are (epistemically) rational or justified, but it's typically thought that those notions really apply more generally to doxastic states, which also include suspension, disbelief, and possible other states, such as credences. It is typically thought that whether a doxastic state is rational or justified is a question of whether it was (or would be) formed, maintained, and destroyed in the right ways, e.g. in a reliable way or in response to our evidence. The process of forgetting is just one way of destroying a doxastic state. So if forming a belief can be rational or justified—which surely it can—it seems like forgetting, when faced with a memory limitation, should be equally open to being rational or justified. Hence the objection that forgetting can't be rational or justified on the basis of our control of it shouldn't seem very compelling.

The overarching goal of this paper is to argue that forgetting should be considered on par with other paradigmatic elements of epistemology, like beliefs and the social institutions of knowledge. In the sections below, we'll argue that understanding what we forget is as essential in understanding epistemic outcomes as is what we believe, how we collect evidence, and how we share our information with others. As mentioned, we're inclined to see the importance of forgetting for epistemic outcomes as symptoms of cases of forgetting being more or less rational or justified. If you're not convinced though, for the arguments below it's possible to follow Schwarz (2017,

⁵ This might not be right though, as the example of forgetting yesterday's pain (mentioned above) shows. We're inclined to see that reason as non-epistemic, but the point here is just that limited agents have some reasons that unlimited agents don't have.

p. 25) and view forgetting badly as merely an epistemic “shortcoming” rather than irrational or unjustified.

For ease of exposition below, we’ll switch between talking about “what is forgotten” and “ways agents manage their memory” as the things that are rational or justified.⁶ By doing this, we’ll appear to conflate two different things: (1) whether what is forgotten is rational/justified to forget, and (2) whether an agent is rational/justified in managing their memory in some way. When it comes to belief, the analogues of these are distinct. We might ask, for example, (1) whether an agent’s belief in God is rational, or (2) whether it’s rational for an agent to surround themselves with believers to get themselves to believe. Following the literature on rational belief, we intend to focus on what is forgotten in each case, rather than the strategy an agent uses to effectuate in themselves the loss of doxastic states. Because English lacks an analogue of “rational beliefs,” which might have been something like “rational forgets,” we use the (misleading) language about “ways of managing memory” in some places below. When we use that locution though, note that we intend to refer to what is forgotten, not the way of managing memory itself.

3 A simple model of group information sharing

To explore how the memory limitations of individuals affects their epistemic states in groups, we’ll use a simple agent-based model of group information sharing. We’ll think of agents as having different bits of information or evidence that can combine to support conclusions. This model takes inspiration from instances of group information sharing like the one found in the iconic sequestered jury-room scenes from *12 Angry Men* (Lumet et al. 1956). In that screenplay, jurors share pieces of information with the goal of coming to an agreement about the guilt or innocence of a defendant. Collectively, the shared information is supposed to support some conclusion about the guilt or innocence of the defendant. That the defendant had a strong motivation to commit the crime, for example, would be a piece of information that by itself supports thinking the defendant is guilty. The combination of the information that the defendant left the house in a blue shirt and the information that the criminal was wearing a red shirt supports thinking the defendant is innocent.

In our model, agents remember and exchange pieces of information, which we’ll call “propositions”. Some sets of propositions constitute *arguments*, which support some conclusion with some strength. We’ll assume that our agents are rational in that they believe whatever is, on balance, supported by the arguments constructible from the information they have. On each round of the model, an agent will share one of the propositions they have with everyone in the group, and everyone will update their stock of propositions with the newly shared proposition.

⁶ We’ll also drop talk of justification here, simply so that we don’t have to repeat “rational or justified” many times, but nothing we say below hinges on whatever difference there may be between rational and justified beliefs/things forgotten.

This model is an extension of the simpler model we used in earlier work (Singer et al. 2019). In the earlier model, agents exchange reasons for beliefs, where a reason for belief is treated as a pair of (1) a supported conclusion and (2) a (real-valued) strength at which the reason supports that conclusion. Agents believe the conclusions that are supported by the reasons they have, where a set of reasons is taken to support a conclusion when the sum of the reasons for the conclusion in the set, weighted by their strength, is greater than the weighted sum of the reasons against. Group deliberation happens by agents sharing reasons with the other members of the group.

One worry about our earlier model is that it can't represent agents engaging in discussion where information combines non-additively. All that agents can do is share a reason with a fixed strength of support for a conclusion. The reasons each person has simply sum to their belief. So, there is no way, for example, for agents to contribute two premises of an argument in support of P where neither premise supports P on its own.

This new model allows agents to contribute to the discussion different propositions that collectively support a conclusion, despite neither supporting the conclusion on its own. They do this by each contributing different propositions in an argument. Agent A might contribute the conditional, while Agent B contributes the conditional's antecedent. They thereby collectively engage in *modus ponens*. The new model is also compatible with arbitrary logics, including non-monotonic logics.⁷ So, our new model of group deliberation is significantly more powerful than its predecessor in Singer et al. (2019), despite only a small increase in complexity.

For the discussion here, we will only use a small amount of the model's power. We will assume that the groups of agents are considering only one main question, like whether the boy committed the crime. We will also assume that there is a fixed (though potentially very large) class of propositions that can be had by agents and which combine to form arguments for or against views about the main question—though which reasons are had by agents will change over time. We also assume that which sets of propositions constitute arguments is fixed, as is what the arguments support about the main question and the strength of that support.

To assess the effect of memory limitations on otherwise fully ideal and rational groups of deliberating agents, we'll assume that an agent's belief about the main question is exactly what is supported by the balance of the arguments they can construct from the propositions they have. For example, if an agent has propositions p_1, \dots, p_8 and $\{p_1, p_3\}$ is an argument with weight 2 to believe the boy is guilty, $\{p_3, p_4\}$ is another argument for guilty of weight 0.5, and $\{p_1, p_3, p_8\}$ is an argument of weight 5.5 for not-guilty, the agent will all-things-considered believe that the eighteen-year-old is not guilty (and will do so with strength 3).

In our model, all of the agents start out with the same number of randomly-selected propositions from the collection of all propositions, although different

⁷ Since the sets of propositions that constitute arguments in this model can be arbitrary sets, there can be sets that support some conclusion to a certain degree despite a superset of it supporting the negation of the conclusion to a greater degree.

agents may start out with different propositions. Then on each round of the model, a random agent is selected to share one of the propositions they have. We'll describe how the agent decides which proposition to share below, but after they share the proposition, we assume every other agent 'hears' and 'accepts' that proposition by adding it to their stock of propositions and updating their belief about the main question in light of the arguments the agent is able to form with their new stock of propositions.

Agents with unlimited memory have it easy when it comes to how to update their stock of propositions upon learning a new one: they simply add it in. But what about an agent with limited memory? Here we'll consider three different simple ways an agent might manage their limited memory.⁸ Each way of modelling memory treats the agent's memory as fixed in size. So when an agent has a full memory, in order to remember something new, they must forget something else.⁹

First, we'll consider agents who manage their memory in the simplest way possible. 'Simple-minded' agents forget a proposition at random when their memory is full. If the agent has a memory limited to 7 propositions, when the agent gets an 8th proposition, they lose one of the 8 propositions at random and keep the rest. The agent will then reassess what they believe on the basis of the propositions they still have.¹⁰ It's probably obvious that forgetting this way isn't ideal, but it's instructive to see how it compares to the other ways of managing memory below.

A more plausibly rational way to manage a limited memory is to forget the propositions that are the least informative for the agent. Intuitively, the informativeness of a proposition can be measured by how much one's beliefs would change when that proposition is added to or removed from their stock of propositions. So to measure how informative a proposition is for an agent who has it, we compare what their belief would be if they lost the proposition to what their belief is with it. When their memories are full and they hear a new proposition, 'weight-minded' agents forget the proposition that is the least informative in this sense.

⁸ The ways of forgetting we use were also in used in our earlier work (Singer et al. 2019). Like in that work, here we keep fixed and uniform how many propositions the agents can remember, but the results we discuss here do not depend on this. In other work studying agents with limited memories, the norm has been to model limitations on memory by limiting the number of states a finite automaton that represents the agent can be in. Those models clearly differ greatly from the model we discuss here. See Singer, et al (2019), note 11 for more information.

⁹ Of course, this idealizing assumption is not realistic, but since the role of the model is to help us *understand* the effect of memory limitations on social epistemic outcomes for groups, the complexities of a more realistic memory model would likely detract from the more general usefulness of the model.

¹⁰ For each forgetting method, agents will reassess what they believe after forgetting a proposition. One might think it's more realistic for agents to continue believing what was supported by their information even after they forget the supporting information. Agents like that would be much more doxastically complicated than the kinds of agents we consider, since our agents can be represented by just the supporting information they have, not any additional belief state that would itself need updating methods, etc. One reason to think that forgetting methods like ours might be realistic in some cases is their ability to explain phenomena witnessed in the real world, like group polarization (see Singer et al. 2019). Moreover, modeling limited agents as fully removing forgotten information from their beliefs is standard. See Halpern et al. (2014) for a survey.

A third way that agents might manage their memory prioritizes the role the proposition plays in their belief state, rather than how informative it is. ‘Coherence-minded’ agents prioritize remembering the propositions that show most in favor of the view that seems right to them in light of all the propositions they have. When this kind of agent hears a proposition that takes them over their memory limit, if they have any propositions that don’t factor into any arguments, they first forget one of those (at random). If not, they consider what view about the main question is supported by *all* of the propositions they have (including the new one). They then forget a proposition that on balance shows for an opposing view. “Why ought I prioritize remembering misleading arguments in favor of an apparently false view?” these agents might ask. These agents try to respect the informativeness of the propositions they have though, so when they forget a proposition that on balance supports the opposing view, they start with the *least* informative. So, even though these agents care about the coherence of their belief state, they don’t biasedly maximize that by actively ignoring the primary arguments for the opposing view. They forget first the propositions that support the opposing view in the least consequential way and only after that do they forget more informative propositions. If the agent has no propositions that on balance show against their view, they forget a proposition that shows in favor of their view, but again, they start with the least informative.

To show that the different ways of forgetting have important effects on the epistemic prospects of groups, we’ll compare the epistemic impacts of different ways of managing memory to the impact of different ways individuals might share their information with others. We focus on ways to share information as a comparison class because a large literature about group opinion dynamics across a range of disciplines has highlighted the importance of these. A central example of this is the work on the hidden profile paradigm in social psychology, which shows that groups tend to repeat and discuss their mutually-shared information rather than bring new information to the forefront (see Stasser and Titus 1985, 1987). In light of that research, many have recommended that individuals “speak up” when they disagree or have something new to add to the discussion and have tested methods for encouraging people to do so.¹¹ The assumption is that it’s epistemically better for the group if individuals contribute more of their private, non-shared information to the group than if they keep it to themselves. Other sophisticated accounts of how individuals should contribute to groups can be found in many places, including much of the work of Cass Sunstein (particularly Sunstein 2002). Sunstein and Hastie (2014), for example, give six bits of advice to groups to encourage better information sharing. These include that the group include a “red team”, who tries to adduce arguments against a certain proposal, and openly assign different roles to members of the group so information is more effectively elicited from the members in their roles.

¹¹ See Cruz et al. (1997), for example. That article lists a number of other tested interventions intended to increase information pooling in groups. For a survey of the relevant literature, see Wittenbaum et al. (2004).

We'll consider three different ways agents might share their information with a group. Like above, the first method is the simplest: agents share one of the propositions they have at random.

A second (plausibly more helpful) way to share information would be to contribute to the group discussion the proposition with the highest informativeness in light of what has already been said. 'Influential' speakers will use this method. When they are chosen to speak, they share the proposition that is the most informative of the ones they have. Here, the informativeness of a potentially shared proposition is measured in terms of how much it would shift what it rational to believe on the basis of what is already part of the shared group information.¹² This is measured in terms of difference between (1) the sum of the strengths of new arguments for the conclusion that can formed with the new proposition, and (2) the sum of the strengths of new arguments against the conclusion that can formed with the new proposition. When influential speakers don't have any propositions to share that would influence the group in any direction, they just share a proposition at random. This is one way to flesh out the idea of group members 'speaking up' when they have something to add.

Another way to flesh out the idea of members of a group 'speaking up' when they have something to add is to have members contribute the proposition to the discussion that they see as pushing the group most towards what they see as the right conclusion. Whereas influential agents tried to maximize the informativeness of their contribution, 'biased' speakers share the proposition that they take to be the most telling in favor of the position they hold. Here again, the measure used is how much it would shift the view supported by the arguments that could be formed on the basis of the previously shared information. When they don't have any propositions that would push the group in their direction, biased speakers share a proposition at random, as long as that proposition doesn't cut against their view in light of what has been previously shared (and if there are no such propositions, they share nothing).

All together then, we have three ways that agents can manage their memory and three ways agents can choose which pieces of information to share with the group. Those methods are summarized in Table 1.

To evaluate each of these ways of sharing information and managing memory, we'll compare the performance of groups using the different methods. Since we assume there is a fixed class of arguments, a fully informed agent who knows all of the propositions would have access to all of the arguments and would thereby have the view supported by all of the potentially available information. One way we'll measure the performance of a group is by asking what percentage of group members have the view that this fully informed agent would have. It would be too strong to also require that the agents' attitudes towards the proposition have the same strength as the fully informed agent's, since limited agents will lack the memory space needed to collect all of the information needed to do that. So on this measure of the

¹² When influential and biased speakers share information, they measure the informativeness of their information based on all of the information that has been shared, not just what they remember of it. We can think of this like the agent doing a literature review before publishing: the expectation is that they will add to the discussion that is being had by everyone, not just the aspects of it they happen to be familiar with.

Table 1 Different ways agents manage their memory or what they share with the group

Method	Manages	Summary
Random-memory	Memory	Agents forget at random
Weight-minded	Memory	Agents forget least informative proposition
Coherence-minded	Memory	Agents forget propositions for other views, starting with least informative, then forget least information for their own view
Random	Sharing	Agents share at random
Influential	Sharing	Agents share most informative given what has been shared, random if all are non-informative
Biased	Sharing	Agents share most in favor of their view given what has been shared, random if none can influence the group towards their view

epistemic success of the group, we'll measure the percentage of agents who have the view about the main question that someone who had all the potentially available information would have, regardless of the strength of that attitude.

Why think this is a good measure of epistemic success? It's plausible, we think, that having the attitude that's supported by all the potentially available information might itself be an epistemic good. By appeal to a fairly weak anti-skeptical assumption, however, we can also say that what is supported by all the potentially available information is likely the truth. So veritists like Goldman (1999) and Singer (2018), who conceive of true belief as a final epistemic end, should also see our measure as measuring genuine epistemic value. Similarly, one might think that having the attitude that's supported by all the potentially available evidence is sufficient for one's belief to be justified, which may itself be of final epistemic value. If not, having that attitude may be instrumentally valuable in the pursuit of knowledge or some other epistemic good. The measure we provide is thus one that tracks many conceptions of what is epistemically good. We'll say that the attitude supported by all the information is the 'right' attitude, leaving the question of how to fill out the details to the reader.

In addition to the percentage of individuals that have the right attitude, there are a number of other measures of group performance we might be interested in. We'll also measure when (or whether) the group converges on a single attitude, whether the group's beliefs stop evolving independently of convergence, and whether the group has the right attitude when it does converge. As such, we're tracking what Zollman (2007) calls 'speed' and 'reliability' as well as the 'consensual' and 'veritistic' aspects of group performance measured by Betz (2012). We now turn to using these measures to see the effects of different ways of sharing information and managing memory on groups of deliberating agents.

4 When and how much forgetting matters

The model described above was simulated in NetLogo. We collected data from 1000 runs of every combination of way to share information and manage memory for groups of 25, 50, and 100 agents, 100 propositions, 100 and 400 arguments, and memory capacities of 5, 10, 20, and 50 propositions. We'll start by considering the results from the runs with 100 arguments and 25 agents with memory capacity 10. At the beginning of each run, each agent was given 10 random propositions (with different agents potentially getting different ones). From the set of 100 propositions, 100 random subsets of size 1–3 were designated as arguments. Each of the arguments was randomly assigned to support either one view about the main question or its negation (e.g. the boy is guilty or the boy is not guilty). The strength of the support was randomly assigned by an exponential distribution with mean 1. Because of the way arguments are created, we can expect the vast majority of them to have a low strength (i.e. they're weak arguments points in favor of a view, like the boy having the hair color reportedly seen by a witness).¹³

Consider what happens when the agents have unlimited memories. Figure 1 shows what percent of the 25 agents have the right belief as time progresses (averaged over 1000 runs). What we see is that after 1000 steps of the model, the different ways agents share information have a major effect on what percent of the population has the right belief. At 1000 steps, in groups sharing randomly, an average of 77% of the agents have the right attitude. In groups using biased sharing, only 63% of the group has the right attitude on average. But groups using influential sharing perform a lot better: 88% of the group has the right attitude after 1000 steps. Here, how the agents share information highly affects the outcome. Influential sharing significantly outperforms the other two information sharing methods ($p < 2.902 \times 10^{-11}$ for two-sided t tests, $p < 2.2 \times 10^{-16}$ for Wilcoxon rank-sum tests, Cohen's $d > 0.299$),¹⁴ which intuitively makes sense since influential sharing spreads the weighty information first, rather than potentially hiding it (like biased sharers) or waiting for it to be randomly shared (like random sharers). Even after just 50 steps of the model, influential sharing does significantly better than the other two (t tests with $p < 8.6 \times 10^{-9}$ for both pairs, Cohen's $d > 0.25$).

The results show that there is a plausible mechanism by which encouraging group members to 'speak up' when they have something to say can importantly influence the outcome for the group. As such, the results can be seen to vindicate the work in social psychology on the hidden profile paradigm and the work of authors like Sunstein and Hastie (2014) mentioned above. In particular, what these results suggest is that agents who have information that bears on the main question in any direction

¹³ The strengths were assigned in this way to mimic the idea that few arguments are "clincher" arguments—ones which totally win the day—while there are many arguments that lend some support to one view or another.

¹⁴ When they're appropriate, we use parametric tests, like t -tests, and non-parametric tests, like the Wilcoxon rank-sum test (Mann–Whitney U test) and Kolmogorov–Smirnov test ("KS-test"). Cohen's d is a measure of effect size, with < 0.2 being generally recognized as small and > 0.8 being generally recognized as large. When it's reported with an inequality here, we're summarizing more than one test.

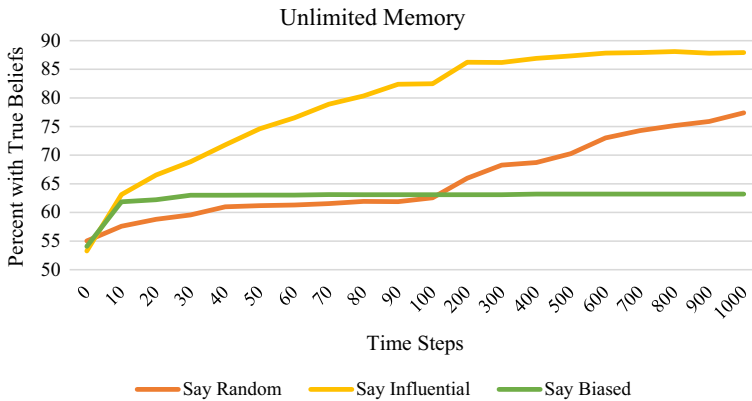


Fig. 1 Percent of agents with unlimited memory who have the attitude supported by all the propositions. Shown for each of the three saying methods averaged over 1000 runs at various time steps

should share it. Having groups where agents only advocate for their own view (even as it changes through deliberation) can be expected to do significantly worse—even worse than agents who just share information at random.

This last result, that groups of agents advocating for their own view (the biased agents) do worse than agents who share information at random, might be surprising.¹⁵ What it suggests is that there is at least one natural interpretation of ‘speaking up’ that we should avoid. One might think the poor performance of the biased agents is explainable by those agents sharing at random when they have nothing to say that helps their side. Might the adversarial method be diluted by agents adding random information to the mix? This turns out not to be the case. We compared biased groups to what we might call ‘very biased’ groups, where agents do not share anything at all when they don’t have anything with positive informational content to add. These very biased agents can be thought of as contributing only when they can influence the group, in contrast to biased agents who always contribute but prefer to contribute information that supports their side. In 1000 runs, the performance of very biased groups with these parameters is not significantly distinguishable from the biased groups ($p=0.19$ in a t test and the Wilcoxon rank-sum test, $p=0.24$ in a KS-test). So it isn’t that feature of biased groups that explains their poor performance when the agents have unlimited memories.¹⁶

¹⁵ It’s worth noting that the distributions are very different. At step 1000, the biased agents all have either 100% of the agents on the right side or 100% on the wrong side. For random agents, roughly the same number of runs of random agents have 100% on the right side, but the rest of the runs are roughly evenly spread through the rest of the space. The difference in mean is significant though ($p < 3.2 \times 10^{-13}$ in a t test).

¹⁶ One might expect biased sharers to perform equally as well as random sharers in the long run. This turns out not to be true because biased sharers end up hiding information, since when they have nothing to support their own view, they share a proposition at random, *as long as that proposition doesn’t go against their view, in light of what has been shared.*

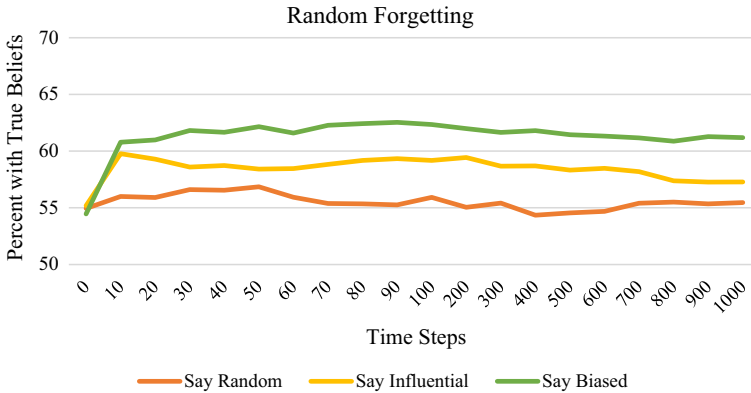


Fig. 2 Percent of agents who have the attitude supported by all the propositions in groups of agents with a memory limit of 10 and who forget at random. Shown for each of the three saying methods averaged over 1000 runs at various time steps

Understanding the dynamics of biased groups becomes even more complex when we look at limited agents. Figure 2 shows the percent of agents with the right attitude over time, again averaged over 1000 runs, when the agents have a memory that's limited to 10 propositions and forget at random when their memory overflows. Note that the scale of the graph has changed from the previous figure. What we see is that when the agents' memories are limited in this simple way, the order and relative impact of the sharing methods completely changes. Here, agents who share information biasedly do significantly *better* than agents who use either of the two other systematic ways of sharing information ($p < .05$ in both cases for t tests and $p < 0.005$ in Wilcoxon rank-sum tests, Cohen's $d > 0.088$). Random and influential speakers aren't statistically distinguishable (t test $p = 0.35$ and Wilcoxon rank-sum $p = 0.17$). Given such a result using a sample this large, we should think that if there is a difference in the mean performance of random and influential speakers, the magnitude of that difference is quite small.

Notice that with these limited agents, the way agents share information had a very muted impact on the group outcome compared to the impact it had for unlimited agents. For unlimited agents, how agents share information made the difference between 63% of the group getting it right and 88% getting it right (a 25 percentage point difference). For these limited agents, the difference is less than 6 percentage points. The same pattern generalizes to the other ways limited agents might manage their memory. Figures 3 and 4 show the effect of the different ways agents share information for weight-minded and coherence-minded agents, respectively.

For weight-minded agents with these parameters, the effect of how agents share information is harder to detect statistically than it was for agents who forget at random. The Wilcoxon rank-sum test, which roughly tracks whether one distribution has a higher median than the other, doesn't statistically significantly distinguish any of the ways of sharing from the others ($p > 0.33$ in all cases). Random speakers do better than biased speakers ($p = 0.02695$), but the effect is small (Cohen's $d = 0.099$).

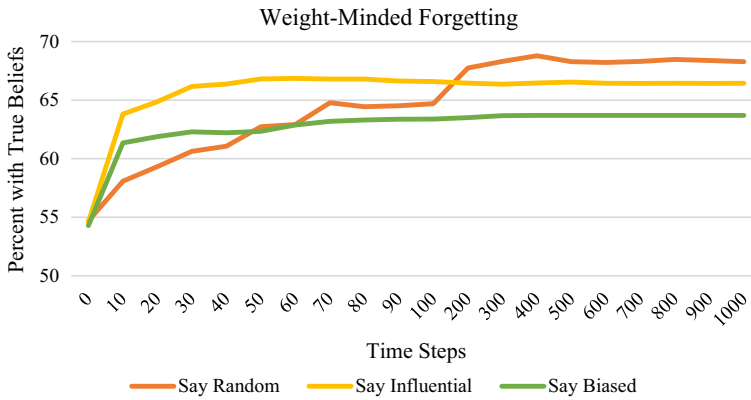


Fig. 3 Percent of weight-minded agents who have the attitude supported by all the propositions in groups of agents with a memory limit of 10. Shown for each of the three saying methods averaged over 1000 runs at various time steps

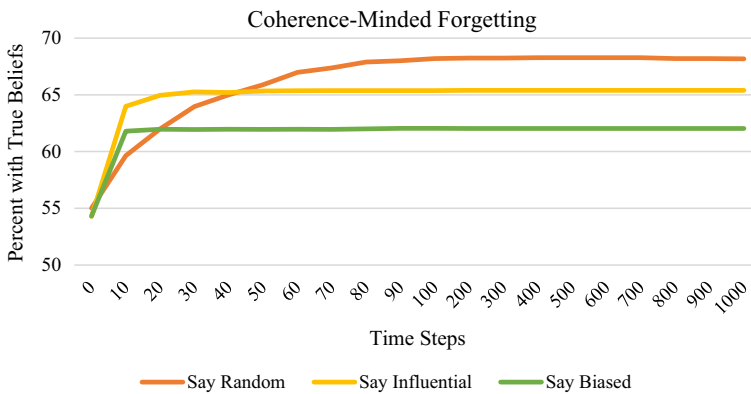


Fig. 4 Percent of coherence-minded agents who have the attitude supported by all the propositions in groups of agents with a memory limit of 10. Shown for each of the three saying methods averaged over 1000 runs at various time steps

There are no statistically significant differences between influential speakers and either of the other two ways of sharing ($p > .19$). The total difference made by how agents share information for weight-minded agents is 4.6 percentage points.

For coherence-minded agents, the range of influence of how agents share information was similarly small at 6.2 percentage points. For those agents, biased sharing is statistically distinguishable from random in a t -test ($p = 0.0018$, $t = 3.1265$) but not a Wilcoxon rank-sum test ($p = 0.5843$). The effect size was also small (Cohen's $d = 0.14$). None of the other pairs of biased agents were statistically significantly distinguishable by either t tests or Wilcoxon rank-sum tests ($p > 0.1$).

We can also look at whether there's a consensus, a majority, only a minority, or none of the population on the truth at 1000 time steps. For each method of sharing information, Fig. 5 shows how each forgetting method performs in these respects.

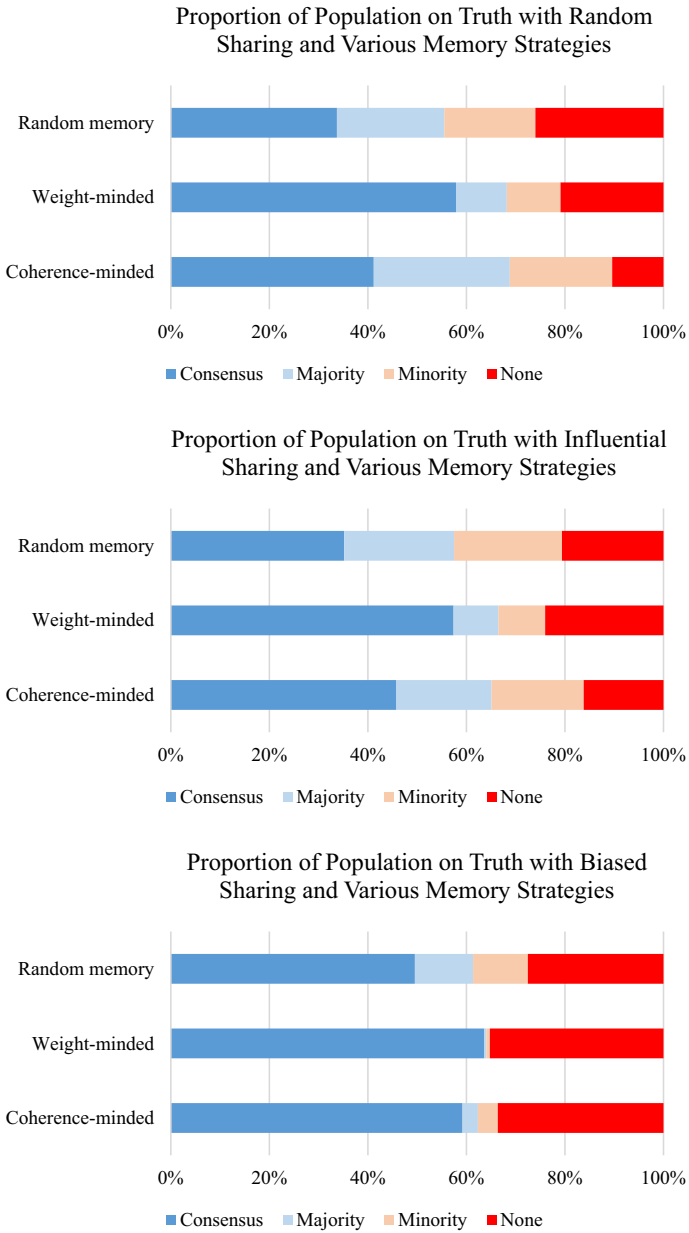


Fig. 5 Proportion of 1000 runs with a consensus, majority, minority, or none of the population having a true belief after 1000 steps for each forgetting method paired with each sharing strategy. Results shown for 1000 runs with 25 agents, 100 propositions, 100 arguments, and a memory of 10

By looking at this collection of graphs, we see that the pattern of how the forgetting methods perform is roughly preserved when the sharing method is changed. One exception is that with biased information sharing, we see a pushing-to-the-edges effect, which decreases the proportion of non-extremal outcomes. That said, it's clearly the forgetting methods that make the largest and most consistent differences. Whereas the pattern of outcomes is relatively consistent as the sharing method changes (i.e. between the graphs), the patterns are much more varied with changes in forgetting method (i.e. within each graph).

When we look at all of the data from 1000 runs of each of the 288 different combinations of parameters, we see more generally that how agents forget has a bigger and more consistent impact on veritistic group outcomes than how agents share information. A Tukey multiple comparison of means test shows that there are, in most cases, statistically significant though minor differences in how groups perform using different ways of sharing information, holding fixed how agents forget. Table 2 shows these differences. What these data show is that, for random rememberers, biased sharing does best (on average 6.3% more of all the agents have the right attitude with biased sharing than with influential) and random does the worst (1.9% fewer have the right attitude than influential). For weight-minded agents and coherence-minded agents, biased sharing does the worst and the other two ways of sharing information are hard to distinguish (and the differences are quite small).

If we instead hold fixed how agents share information and look at how different ways of forgetting information affect the outcome, we see much larger and more consistent differences. Table 3 shows this data. For all ways of sharing, weight-minded memory is best, coherence-minded is second, and random memory does worst. As the table shows though, the magnitude of these differences is influenced by how agents share their information. Whereas for random and influential sharers, the average performance of groups with different memory rules differs by 15.4% and 12.8%, respectively, for biased speakers, that difference is only 3.7%. What we see then is that biased sharing of information mitigates the impact of how agents manage their memories.

All together, these data suggest that at the 1000th time step how agents forget is a greater factor in the overall performance of the group than how agents share information. This is further confirmed by comparing the partial eta-squared measure of percent of variance explained by a given factor, where we see that the amount of variance attributable to differences in forgetting is much higher than the amount attributable to differences in sharing: how agents forget explains 79 times as much variance in the outcome as how agents share information.¹⁷

But, it would be too quick to conclude from these statistics that how we forget is overall more important to what percentage of the group gets the right answer than how information is shared in the group. All of the data described above are for groups after 1000 time steps in the model. After 10 time steps, the story is different.

¹⁷ This was computed using partial eta-squared in a ANOVA type III, where both of these factors only explained a tiny amount of the variance in the model. We believe the extreme overall variance is explained by the wildly different initial conditions that are possible in the model.

Table 2 Average differences in means (as a percentage value) of group performance for each of the saying rules holding fixed the memory management rule, including adjusted p values

Memory rule	Saying rule comparison	Winner	Diff in means	95% CI	p value
Random Memory	Random vs. Biased	Biased	8.1	[7.2 9.0]	<2.2E-16
	Influential vs. Biased	Biased	6.3	[5.4 7.2]	<2.2E-16
	Influential vs. Random	Influential	1.9	[1.0 2.8]	2.60E-06
Weight-Minded	Random vs. Biased	Random	3.5	[2.6 4.4]	<2.2E-16
	Influential vs. Biased	Influential	2.8	[1.9 3.7]	<2.2E-16
	Influential vs. Random	Random	0.7	[-1.6 0.2]	0.1879
Coherence-Minded	Random vs. Biased	Random	1.0	[0.2 1.8]	0.0136
	Influential vs. Biased	Influential	1.6	[0.8 2.4]	1.03E-05
	Influential vs. Random	Influential	0.6	[-0.2 1.5]	0.1598

Bold text indicates insignificant differences

Table 3 Average differences in means (as a percent value) of group performance for each of the memory rules except unlimited, holding fixed the sharing rule, including adjusted p values

Saying rule	Memory rule comparison	Winner	Diff in means	95% CI	p value
Random	Random vs. Coherence	Coherence	11	[10.2 11.8]	<2.2E-16
	Weight vs. Coherence	Weight	4.4	[3.6 5.2]	<2.2E-16
	Weight vs. Random	Weight	15.4	[14.6 16.2]	<2.2E-16
Influential	Random vs. Coherence	Coherence	9.7	[8.9 10.6]	<2.2E-16
	Weight vs. Coherence	Weight	3.1	[2.2 3.9]	<2.2E-16
	Weight vs. Random	Weight	12.8	[12 13.7]	<2.2E-16
Biased	Random vs. Coherence	Coherence	1.8	[0.9 2.8]	1.37E-05
	Weight vs. Coherence	Weight	1.9	[0.9 2.8]	7.80E-06
	Weight vs. Random	Weight	3.7	[2.8 4.7]	<2.2E-16

At that early point, there is no discernable difference between groups of coherence-minded and weight-minded agents (adjusted $p=0.97$ in a Tukey test) and both are better than simple-minded memory management by only 2.7% (95% CI [2.3% 3.2%], adjusted $p < 2.2 \times 10^{-16}$). The difference made by how agents share information is higher here though, with a 5.4% mean difference between random and biased speakers (95% CI [5.0% 5.9%], adjusted $p < 2.2 \times 10^{-16}$), a 4.1% difference between random speakers and influential speakers (95% CI [3.7% 4.6%], adjusted $p < 2.2 \times 10^{-16}$), and a 1.3% difference between influential and biased speakers (95% CI [0.8% 1.7%], adjusted $p < 2.2 \times 10^{-16}$). Comparing the partial eta-squared also shows the relatively higher influence of sharing after 10 time steps, with how agents share information explaining roughly 3.3 times as much variance as how agents forget. In contrast to what happens in the long run (at 1000 steps), in the short run (at 10 steps), it appears that how agents share information is more important than how they manage their memory.

So we shouldn't conclude that how we forget universally influences what percent of the group has the right belief more than how the group shares information. But, these data do support thinking that how agents manage their memory should be seen as akin to how agents share information in terms of overall effects on the dynamics of the group deliberation. Early on, how agents share their information has a larger effect (by roughly 3.3 times). At about 30 time steps, the two have similar influence (measured both in terms of differences in means as well as explained variance). Beyond that point, how agents manage their memory has a greater effect. By 1000 time steps, memory management has a much greater impact (79 times) than how agents share information.

What might explain the larger impact of forgetting than sharing on group outcomes in the long run? We think the answer lies in understanding how sharing and forgetting differentially impact deliberating groups. Consider a jury where only one agent has a key piece of information. If the agent doesn't share the key piece of information on some round, the information is still a candidate for being shared in future rounds. And when the information shared on some round, it can still be forgotten by that agent or others. In that way, sharing only has impermanent effects on where information exists in groups. Forgetting, on the other hand, has much more permanent impacts. When a piece of information is forgotten by everyone who has it, that information is gone and cannot influence what any member of the group thinks anytime in the future. This can obviously have a big impact when the one member of jury with the key information forgets it, but since some methods of forgetting induce correlations in what is forgotten (like when everyone forgets their least influential proposition), we should expect forgetting to permanently remove information from a group's collective intellectual resources sometimes even when that information is held by many members of the group. In this way, how agents forget information can have a greater, more permanent, impact on the long term informational future of groups than how they share information.

What about forgetting's impact on other metrics of group performance? If we look at whether and how quickly the groups form a consensus, whether the group has the right attitude when it does reach convergence, and whether and how quickly members of the group stop changing their belief (regardless of whether there's a consensus), we get the same basic result. In each case, how individuals in a group manage their memory has an impact that's at least roughly on par with the impact of how agents share their information.

With regard to whether agents stop changing their beliefs at any point before 1000 steps, how agents forget explains 56.2 times the variance explained by how agents share information. Memory management explains 3.8 times the variance in how quickly that stop happened, when it did. How agents forget explains 1.5 times the variance in whether groups converged on an answer in less than 1000 steps compared to sharing, and it explains only slightly less of the variance as does how agents share information in how quickly groups converged, if they did converge in less than 1000 steps. Moreover, for runs that did converge in 1000 steps or fewer, how agents forget explains 2.4 times the variance in whether the group converged on the right attitude but only 0.6 times the variance in the time it took to get convergence on

the truth in those groups.¹⁸ So, our model results indicate that how agents forget matters a lot for understanding the dynamics of group deliberation, often more than how agents share information, and in each of these cases, it seems like it's the more permanent informational effect of forgetting than sharing that engenders that result.

5 The place of forgetting in epistemology

What does the importance of forgetting in our model say about rational group deliberation for real people? We take it that many arguments in normative domains like epistemology, ethics, and political philosophy appeal either implicitly or explicitly to models of idealized agents and social systems. A common argumentative form is to draw out the entailments of antecedently plausible assumptions about how rational or idealized agents or systems work using idealized models of agents. Consider Dutch book arguments for probabilism in formal epistemology, for example (see Vineberg 2016). Those arguments work by appeal to assumptions about rational reasoners (like that rational agents wouldn't accept a series of bets that insures their sure loss and that rational agents can be understood in terms of a real-numbered model of both degreed beliefs and utilities) and draw the implication that rational agents' degreed beliefs must be representable by a probability function. Arguments about rational decision theory often have this form too.

We take ourselves to be employing a similar strategy here, though with different idealizations. We start with an idealized model of rational group deliberation in which agents are represented as having certain information and group deliberation is represented as agents sharing that information with other agents. Though our model uses a very simplified notion of memory limitations, it doesn't completely idealize away those limitations (like Bayesian models do). What the analysis here brings out is that, in such a model, how agents manage their memory has significant effects for whether groups achieve optimal epistemic outcomes. So, insofar as our model tracks what explains the epistemic outcomes for real groups, we should expect there to be better or worse ways for real agents in groups to forget.

Does our model track the important elements of group deliberation? If we think of real group deliberation as a process of dissemination and social processing of information, the model looks pretty good. The model's agents, like real groups of deliberating agents, start out with different pieces of information. As the model progresses, agents share different aspects of the information they have. In our model, and plausibly in real group deliberation, combining the different agents' information is the best guide to the truth, so how agents share their information is important to whether members of the group get the right view. Information can only be spread and analyzed if some agent has that information to share.

¹⁸ As above, this was measured by the partial eta-squared in a type III ANOVA. Convergence was tested for only every 10 steps from 0 to 100 steps and then every 100 steps from 100 to 1000.

If no agent remembers some information, that information is no longer available in the collective discussion. So, the core result from our model—that there are non-trivial questions of how it is best for agents to forget information for the sake of the group outcomes—should apply to real deliberating groups.

There were two main goals of this paper: to establish the importance of how we forget for understanding group epistemic outcomes, and to undermine the line of argument that purports to show that forgetting can't be individually rational or justified because of our lack of control. In defense of the first goal, we presented the model and analysis above. It would be too fast to infer directly from our analysis specific lessons about how it's rational for individuals to forget, since it's well-known that what's best for groups may not be best for individuals and vice versa (see, for example, Mayo-Wilson et al. 2011).¹⁹ That said, we can draw particular lessons about what's best for a group in terms of how its members forget, and these can suggest lines of potential research about what might be individually rational or justified.

For example, one of the more intriguing results comes from comparing the effect of different memory strategies for agents who biasedly share information in the model. Biased sharing can be thought of as adversarial in that agents share information with the group that is most likely to push the group towards the view held by that agent. What we saw above is that biased sharing mitigates the influence of how agents forget on the group outcome. As Table 3 shows, the differences made by changes in forgetting is much lower for biased sharing than for other ways of sharing information. After 10 steps, the average worst outcome for the group is significantly higher for biased sharers than the minimum for any other way of sharing information (64.3% for biased, 57.3% for random, and 61.2% for influential, $p < 2.2 \times 10^{-16}$). At 1000 steps, the average worst outcome for biased sharing is 65.2% (with random memory), whereas it's substantially worse at 57.1% and 59.0% for random and influential sharing, respectively (also with random memory, $p < 2.2 \times 10^{-16}$).²⁰ So biased sharing can be seen as a safe (maximin) way to share information, one which protects the group from the worst possible outcome, regardless of how agents manage their memory. This suggests a possible regret-minimization defense of adversarial methods in legal, political, and educational contexts. Fleshing out that defense is beyond the scope of this discussion, but it does offer an example of the type of thing that could be learned using this technique.

The larger aim of this discussion is to highlight the role of forgetting for understanding the topics we care about in epistemology, and we're certainly not the first to do this. Work on epistemic logics have long recognized that agents can forget (see, for example, Fagin and Halpern 1988; Hill 2010; Van Ditmarsch et al. 2009; van Benthem and Velázquez-Quesada 2010; van Benthem 2011). Other modelers of group interaction and game theorists have seen the importance of memory too.

¹⁹ Moreover, the methodology we use here would not translate naturally to testing forgetting methods for individuals, since incoming propositions only come from other agents in the model as it was used here.

²⁰ The percentages given here come from the all of the data we have on limited groups, but the same pattern can be seen in the tables above for those particular parameters.

For example, Barrett and Zollman (2009) show the importance of forgetting for language learning in signaling games, Halpern and Pass (2010), Wilson (2014), and Singer, et al. (2019) show how group polarization may be a rational outcome for memory-limited agents, and it's well-known that forgetful agents can enable groups to avoid non-ideal outcomes like converging on bad social norms or getting stuck in suboptimal pooling equilibria. Bernecker and Grundmann (2019) even argue that forgetting itself can be a source of knowledge.

We take our work to extend the literature showing that epistemologists shouldn't forget forgetting in two ways. First, our agent-based model of group deliberation draws out the epistemic role of forgetting in deliberation and the relative epistemic importance of forgetting compared to other elements of deliberation. And second, by undermining a natural objection to treating forgetting as the kind of thing that can be rational or justified, we hope to have taken a step in securing forgetting as a central element of epistemology—one that can be rational or justified and one that should have a place in the heart of epistemology along with things like rational and justified belief, evidence, testimony, knowledge, and the social structural features of epistemic communities.

References

- Alston, W. P. (1988). The deontological conception of epistemic justification. *Philosophical Perspectives*, 2, 257–299.
- Barrett, J., & Zollman, K. J. S. (2009). The role of forgetting in the evolution and learning of language. *Journal of Experimental & Theoretical Artificial Intelligence*, 21(4), 293–309.
- Bernecker, S. (2018). On the blameworthiness of forgetting. In K. Michaelian, D. Debus, & D. Perrin (Eds.), *New directions in the philosophy of memory*. Abingdon: Routledge.
- Bernecker, S., & Grundmann, T. (2019). Knowledge from forgetting. *Philosophy and Phenomenological Research*, 98(3), 525–540.
- Betz, G. (2012). *Debate dynamics: How controversy improves our beliefs* (Vol. 357). Berlin: Springer.
- Cruz, M. G., Boster, F. J., & Rodriguez, J. I. (1997). The impact of group size and proportion of shared information on the exchange and integration of information in groups. *Communication Research*, 24, 291–313.
- Fagin, R., & Halpern, J. (1988). Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34(1), 39–76.
- Feldman, R. (2000). The ethics of belief. *Philosophy and Phenomenological Research*, 60(3), 667–695.
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge University Press.
- Ginet, C. (2001). Deciding to believe. In M. Steup (Ed.), *Knowledge, truth and duty* (pp. 63–76). New York, NY: Oxford University Press.
- Goldman, A. I. (1999). *Knowledge in a social world*. Oxford University Press.
- Halpern, J. Y. & Pass, R. (2010). I don't want to think about it now: Decision theory with costly computation. In *Twelfth international conference on the principles of knowledge representation and reasoning*.
- Halpern, J. Y., Pass, R., & Seeman, L. (2014). Decision theory with resource-bounded agents. *Topics In Cognitive Science*, 6(2), 245–257.
- Hieronymi, P. (2006). Controlling attitudes. *Pacific Philosophical Quarterly*, 87, 45–74.
- Hill, B. (2010). Awareness dynamics. *Journal of Philosophical Logic*, 39, 113–137.
- Lumet, S. (Director), Fonda, H., Rose, R. (Producers), Rose, R., & Hopkins, K. (Writers). (1956). *12 angry men* [Motion picture]. United States: United Artists Corp.

- Mayo-Wilson, C., Zollman, Kevin J. S., & Danks, D. (2011). The independence thesis: When individual and social epistemology diverge. *Philosophy of Science*, 78(4), 653–677.
- McHugh, C. (2012). Epistemic deontology and voluntariness. *Erkenntnis*, 77(1), 65–94.
- Moss, S. (2015). Time-slice epistemology and action under indeterminacy. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (pp. 172–194). Oxford: Oxford University Press.
- Peels, R. (2017). *Responsible belief: A theory in ethics and epistemology*. Oxford: Oxford University Press.
- Ryan, S. (2003). Doxastic compatibilism and the ethics of belief. *Philosophical Studies*, 114, 47–79.
- Schwarz, W. (2017). Evidentialism and conservatism in bayesian epistemology. Unpublished. <https://www.umsu.de/papers/conservatism.pdf>. Accessed 31 Oct 2018.
- Singer, D. J. (2018). How to be an epistemic consequentialist. *Philosophical Quarterly*, 68(272), 580–602.
- Singer, D. J., Bramson, A., Grim, P., Holman, B., Jung, J., Kovaka, K., et al. (2019). Rational social and political polarization. *Philosophical Studies*, 176, 2243.
- Smith, A. M. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics*, 115, 236–271.
- Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussions. *Journal of Personality and Social Psychology*, 48, 1467–1478.
- Stasser, G., & Titus, W. (1987). Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of Personality and Social Psychology*, 53, 81–93.
- Steup, M. (2008). Doxastic freedom. *Synthese*, 161, 375–392.
- Steup, M. (2011). Belief, voluntariness and intentionality. *Dialectica*, 65(4), 537–559.
- Sunstein, C. (2002). The law of group polarization. *Journal of Political Philosophy*, 10(2), 175–195.
- Sunstein, C., & Hastie, R. (2014). Making dumb groups smarter. *Harvard Business Review*. <https://hbr.org/2014/12/making-dumb-groups-smarter>. Accessed 15 Apr 2018.
- Vahid, H. (1998). Deontic vs. nondeontic conceptions of epistemic justification. *Erkenntnis*, 49(3), 285–301.
- van Benthem, J. (2011). *Logical dynamics of information and interaction*. Cambridge: Cambridge University Press.
- van Benthem, J., & Velázquez-Quesada, F. R. (2010). The dynamics of awareness. *Synthese*, 177(1), 5–27.
- Van Ditmarsch, H., Herzog, A., Lang, J., & Marquis, P. (2009). Introspective forgetting. *Synthese*, 169(2), 405–423.
- Vineberg, S. (Spring 2016 Edition). Dutch book arguments. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/spr2016/entries/dutch-book/>. Accessed 30 Apr 2018.
- Weatherston, B. (2008). Deontology and Descartes' demon. *Journal of Philosophy*, 105, 540–569.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Wilson, A. (2014). Bounded memory and biases in information processing. *Econometrica*, 82(6), 2257–2294.
- Wittenbaum, G. M., Hollingshead, A. B., & Botero, I. C. (2004). From cooperative to motivated information sharing in groups: Moving beyond the hidden profile paradigm. *Communication Monographs*, 71(3), 286–310.
- Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of Science*, 74(5), 574–587.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.