Investigating the impact of cognitive dissonance reduction strategies on echo chamber formation

Joshua Krause (s3344274)
Max V. Valk (s3246922)
Tom C. Theuer (s3216748)

November 2020

Abstract

Madsen, Bailey, and Pilditch [MBP18] showed that the structure of a network causally contributes to echo chamber formation in a social network. However, it does not follow that an agents' behaviour cannot have an impact on the formation of echo chambers as well. To ensure that the implications from the study of network simulations generalise onto real social networks it is therefore necessary to model, at a suitable level of abstraction, the agents' behaviour more explicitly. One such attempt can be found in the work by Ke, Haiming, Gang, and Yucheng [Ke+20], who implemented an agent logic inspired by research on cognitive dissonance, as discussed by Festinger [Fes57]. We expand on the work by Ke, Haiming, Gang, and Yucheng [Ke+20] by integrating novel insights into the mechanisms underlying cognitive dissonance into the multi-agent model, thereby further enhancing its ecological validity [Bre00]. Furthermore, we show that such an ecologically more valid model allows to generate important insights into the implications of different recommendation strategies for the polarisation and belief distribution in a network. Hence, this study highlights the necessity to compliment traditional investigation techniques, such as social network surveys, with multi-agent model simulations. The latter allow to investigate the impact of manipulations of recommendation algorithms, which would be impossible to investigate, due to ethical considerations, in real social networks.

1 Introduction

In recent years, factors contributing to polarisation in social networks, both offline and online, have garnered substantial interest [MBP18; Ke+20; Gar+18]. Research on the factors contributing to polarisation highlight the role played by so called echo chambers (ECs) in explaining polarisation [MBP18]. ECs were first described by Jamieson and Capella [JC09] at the example of the American conservative media establishment. They can be defined as environments wherein individuals are subject only to beliefs or opinions similar to their own, which reinforces existing positions while discouraging alternative ones [JC09]. Individuals within ECs are typically subject to opinion polarisation, perpetually reinforcing the exclusivity of narratives and theories they are confronted with [JC09; MBP18; Ke+20]. This dynamic has not only been a major breeding ground for conspiracy theories but also affected how we interact with our own personal network on a daily basis [Gar+18]. Research therefore investigates the dynamics behind the formation and development of ECs [MBP18; Ke+20]. Beside the traditional techniques used by social scientists, such as surveying social network participants, multi-agent modelling in particular has shown great successes in explaining and investigating this phenomenon [MBP18; Ke+20].

1.1 Problem

Multi-agent modelling has contributed to our understanding of ECs by modelling and testing our assumptions about the mechanisms governing their development [MBP18; Ke+20]. As opposed to traditional research techniques such as surveying, simulation provides a direct evaluation tool allowing for precise control about the manipulation of different experimentally relevant factors. However, the validity of experimental findings generated by those multi-agent simulations hinges on the validity of the model underlying the simulation [Ke+20]. A model's implications for the real world are commonly referred to as as ecological validity [Bre00]. To ensure high ecological validity, it is necessary to incorporate the *relevant* processes, contributing to a phenomenon in the real world, in the multi-agent simulation model. Evidently, no model can effectively emulate a social network system by encompassing and implementing *all* intricacies of a social system. However, such a hypothetical model would also likely be suspect to over-fitting, eliminating the ability to derive meaningful predictions. Hence, abstractions and simplifications are not only necessary but even *desirable* when modelling complex dynamical systems such as social networks. The key difficulty to modelling those systems then arises from choosing the right level of abstraction, optimally balancing the trade-off between simplicity and ecological validity [And07]. Recent work by Madsen, Bailey, and Pilditch [MBP18] based their findings on a model with purely rationally acting agents, hence abstracting away cognitive biases, potentially influencing an agent's actions. Because this choice drastically simplified the model, it needs to be determined whether the findings remain consistent when modelling more complex human behaviour. Further, because of the limited ecological validity, it may be difficult to investigate promising intervention mechanisms to limit or reverse the formation of ECs.

1.2 State of the art

1.2.1 Madsen et al. 2018

The above mentioned trade-off, sacrificing ecological validity in favour of abstraction, becomes particularly evident in the seminal study conducted by Madsen, Bailey, and Pilditch [MBP18], in which the authors investigated the primary causes of EC formation. Previous research on contributing factors to EC formation focused on two categories of primary mechanisms: how cognitive processes within individuals shape the interaction within social networks and how the structure and inherent features of the network itself contribute to the formation of ECs [Ke+20; MBP18]. Madsen, Bailey, and Pilditch [MBP18] focused on the latter and their results led them to conclude that even rational agents with perfect memory form ECs. Therefore, they claimed the structure of social networks to constitute the primary mechanism driving the emergence of ECs [MBP18]. While these results are compelling, several limitations arise from such a reductionist approach to understanding as complex a phenomenon as EC formation. Even if the network structure is causally sufficient for ECs to emerge, it does not follow that agent behaviour or interventions have no impact on this effect. Hence, it could be that while network structure is causally sufficient for ECs to emerge, specific interventions could again be causally sufficient to prevent their formation. The investigation of this hypothesis requires a more ecologically valid model which takes into account that humans are not inherently rational agents with perfect memory, [Nic98], thereby overcoming the limited ecological validity of [MBP18].

1.2.2 Ke et al. 2020

To achieve a more ecologically valid model one needs to move away from the assumption of rationality as made by Madsen, Bailey, and Pilditch [MBP18] since individuals are unlikely to behave rationally when deciding with whom they want to communicate and exchange information. For example, human individuals tend to seek out information that matches their own beliefs - a finding referred to as "confirmation bias" [Nic98]. Moreover, research has pointed out that being confronted with conflicting information can lead to distress and the resulting state of negative arousal has been called cognitive dissonance (CD) [Fes57; Mcg17]. CD typically arises in individuals who are confronted with opposing opinions or who are engaged in behaviour not conforming to their own beliefs [Mcg17]. The negative arousal that is inherently associated with the experience of CD has been shown to be detrimental to the well-being of an individual [Fes57]. Therefore, Festinger [Fes57], suggested that individuals are compelled to engage in strategies to reduce the negative arousal associated with CD. In the recent decades a tremendous amount of research has been dedicated to investigate such strategies [Mcg17]. Originally, Festinger [Fes57] introduced three strategies to ameliorate CD: changing cognitions, creating new cognitions, or adapt the importance of dissonant cognitions. Since their introduction, those strategies have been investigated, among others, under the theme of attitude change [SGB95; Mcg17] in the light of confrontations with conflicting information. However, the study of attitude change often neglects actual behavioural and cognitive strategies that are deployed by individuals to reduce CD [Mcg17]. However, a consideration of the actual strategies used by individual is crucial since those strategies are likely to affect how an individual prunes and expands their own personal network [Mcg17; Ke+20]. This idea was picked up by Ke, Haiming, Gang, and Yucheng [Ke+20], who attempted to create a more realistic model of opinion dynamics in social networks. Based on similar concerns about the limitations of existing models the authors decided to incorporate research on cognitive dissonance reduction strategies to provide their agents with a more plausible tool set to determine how to form and terminate new connections [Ke+20].

While the strategies (for a discussion of their implementation see method section) implemented in Ke, Haiming, Gang, and Yucheng [Ke+20] are a step in the right direction, the implementation is not reflective of the current state-of-the-art understanding of the processes underlying CD [Mcg17].

1.3 New idea

Recently, new theories emerged that attempt to expand on reduction strategies and the understanding of CD in general, such as the cognitive dissonance reduction model by Hardyck and Kardush [HK68] and Mcgrath [Mcg17]. In their review, the authors discuss that the negative arousal attributed to dissonance likely follows a continuous dimension. Thus, challenging less important cognitions should only lead to little arousal. The authors point out that this is likely to influence the strategy to reduce dissonance - with individuals falling back to more extreme strategies only in scenarios in which their core beliefs are challenged. Additionally, recent work by Harmon-Jones and Harmon-Jones [HH08] and Mcgrath [Mcg17] points towards a behaviour limiting effect of CD: increasing levels of CD interfere with the individual's ability to act. This is likely a result of the negative arousal inherently associated with states of CD [Fes57; Mcg17], which will at some point interfere with the normal functioning of a person. It follows that, should dissonance built up to a threshold at which the individual feels too limited in their control about their own actions, they would again fall back to more radical strategies to gain back control [Mcg17]. This perspective points towards a model of CD in which a single clash of two incompatible cognitions does not necessarily result in enough negative arousal to demand a reduction strategy by the individual [Mcg17]. Rather, this model treats dissonance as an accumulating process. It builds up with each negative encounter until it becomes unbearable [HH08; Mcg17].

Hence, if and which strategy is ultimately chosen to reduce dissonance depends on multiple factors [Mcg17; HK68; HH08; Fes57]: Not only is the importance of the cognition or belief that is challenged by incoming information of relevance, but also the expectation of success of a given strategy, the costs associated with a strategy, and the availability of the strategy. All those factors are likely to influence which and whether a dissonance reduction strategy is deployed [Mcg17]. Additionally, the context of a situation as well as the personality and experiences of an individual crucially influence whether dissonance reaches unbearable levels and compels an individual to act [Mcg17].

As mentioned in the earlier section, these new insights are not yet included in the implementation by Ke, Haiming, Gang, and Yucheng [Ke+20]. However, by modelling dissonance reduction strategies more appropriately, we can better explore their impact on the formation of ECs within social networks. Additionally, ensuring an optimal trade-off between a parsimonious model and a more realistic treatment of CD allows for a better investigation of possible interventions designed to reduce the formation of ECs, since results generated using such a model will more easily generalise to real social networks due to a higher ecological validity.

2 Research Question & Hypotheses

Before evaluating our model and investigating the impact of reduction strategies on the formation of ECs, we investigated whether our model, despite the added complexity, would still support the basic premise of the results by Madsen, Bailey, and Pilditch [MBP18]: ECs reliably emerge in social networks. Subsequently, we investigated the following research questions.

2.1 Research Questions

1. How is the formation of ECs related to the maximum possible difference in belief-states considered by a person with a perfect openness score?
2. How does a decrease in individual openness for increasing levels of dissonance affect the formation of ECs?
3. Do different recommendation systems impact the behaviour and formation of ECs in social networks of agents utilising cognitive dissonance reduction strategies?
   (a) How long does it take the network to converge on different opinions?
   (b) Do recommendation strategies that promote a broader belief set lead to higher levels of cognitive dissonance?
   (c) Do the answers to those questions depend on whether or not individuals become more narrow minded when experiencing cognitive dissonance?
4. Do specific recommendation systems reduce the formation of ECs?

To investigate those questions in a testable way, we formulated the following hypotheses.

2.2 Hypotheses

1. The dissonance levels in a network should remain low if a reinforcing recommendation strategy is deployed - reflecting the benefit of the dissonance reduction strategy to access confirming information.
2. ECs should form in a network independent of the recommendation strategy deployed.
3. The polarisation in a network should decrease over time if a random recommendation strategy is deployed, ensuring that each agent has access to a broad range of opinions.
4. In a network in which the randomising strategy is applied, agents will be confronted with conflicting information more frequently and therefore they should experience higher levels of dissonance.

3 Method

3.1 Simulation model

The study by Ke, Haiming, Gang, and Yucheng [Ke+20] includes a formal set of rules to update the belief (opinion) state of agents in a network. Based on the belief state, they formulate additional rules to both expand and prune a social network simulation that are inspired by the theory on CD. Specifically the update strategy by Ke, Haiming, Gang, and Yucheng [Ke+20] for the belief of an agent at any time point (from now on referred to as epoch) assumes that agents have a certain "confidence set" of agents they trust and are willing to listen to. If a connected agent Y shares an information that is too extreme for agent X, then agent X will behave dogmatically and reject the information without consideration. Agents therefore have an accepted range of beliefs [Ke+20; MBP18]. Ke, Haiming, Gang, and Yucheng [Ke+20] assume that at any epoch, all agents communicate with all their connections and send them information based on their current belief state. This prevents them from defining parameters for message quality or remembering the point of origin for messages, which was criticised as an arbitrary choice commonly made in studies on EC formation [MBP18].

Here we decided to diverge from the implementation in [Ke+20] for the first time. Instead of only receiving information from connected agents we decided to implement a recommendation set that can utilise different strategies to recommend to any agent a set of other individuals in the network the agent is not connected to. This choice reflects what is the reality in most prominent social networks such as Facebook or LinkedIn. The inclusion of the recommendation set does not change the agent's confidence set, it still consists solely of other agents that lie within the agent's accepted range of beliefs o (e in Ke, Haiming, Gang, and Yucheng [Ke+20]). A similar parameter exists in [MBP18], where it is used for identical purposes. We denote the accepted range of beliefs as o, to highlight the connection to the personality trait Openness which is included in the "Big Five" personality model where it reflects a desire for new experiences, variety, and even liberalism [MC87]. In our model o is a compound score depending on two parameters: a maximum belief range which is defined by a global parameter applying to the entire network, similar to the parameters used by Ke, Haiming, Gang, and Yucheng [Ke+20] and Madsen, Bailey, and Pilditch [MBP18], and an agent's individual openness value sampled from a non-asymptotic distribution containing values between 0 and 1. Hence, an individual agent's openness score acts as a weight for the maximum belief range, which itself then represents the maximum distance for beliefs considered acceptable by an agent with a perfect openness score of 1. Therefore, especially "narrow-minded" individuals will only consider information very close to their own belief, while especially "open-minded" individuals will consider a broader range of information reflecting the interpretation of the parameter discussed earlier [MC87].

The confidence set is then used to update each agent's belief for the next epoch. In the study by Ke, Haiming, Gang, and Yucheng [Ke+20] each agent's information (corresponding to the agent's belief) in the confidence set receives a weight of 1 divided by the number of agents in the confidence set. Hence, each agent in the confidence set contributes equally to the updated belief state of the agent to which the confidence set belongs. Here we again decided to diverge from the implementation described in [Ke+20], by adding the agents own current belief to the confidence set as well. This choice reflects the temporal dependence that likely exists between one's own belief at the current and last epoch [FT16]. Evidently, this adaptation still hinges on assumptions, such as the fact that the weight of the own prior belief is equal to the weight of another trusted agent.

As mentioned in the beginning, what is special about the work by Ke, Haiming, Gang, and Yucheng [Ke+20] is that they include updating rules for an agent's connection that are inspired by CD reduction

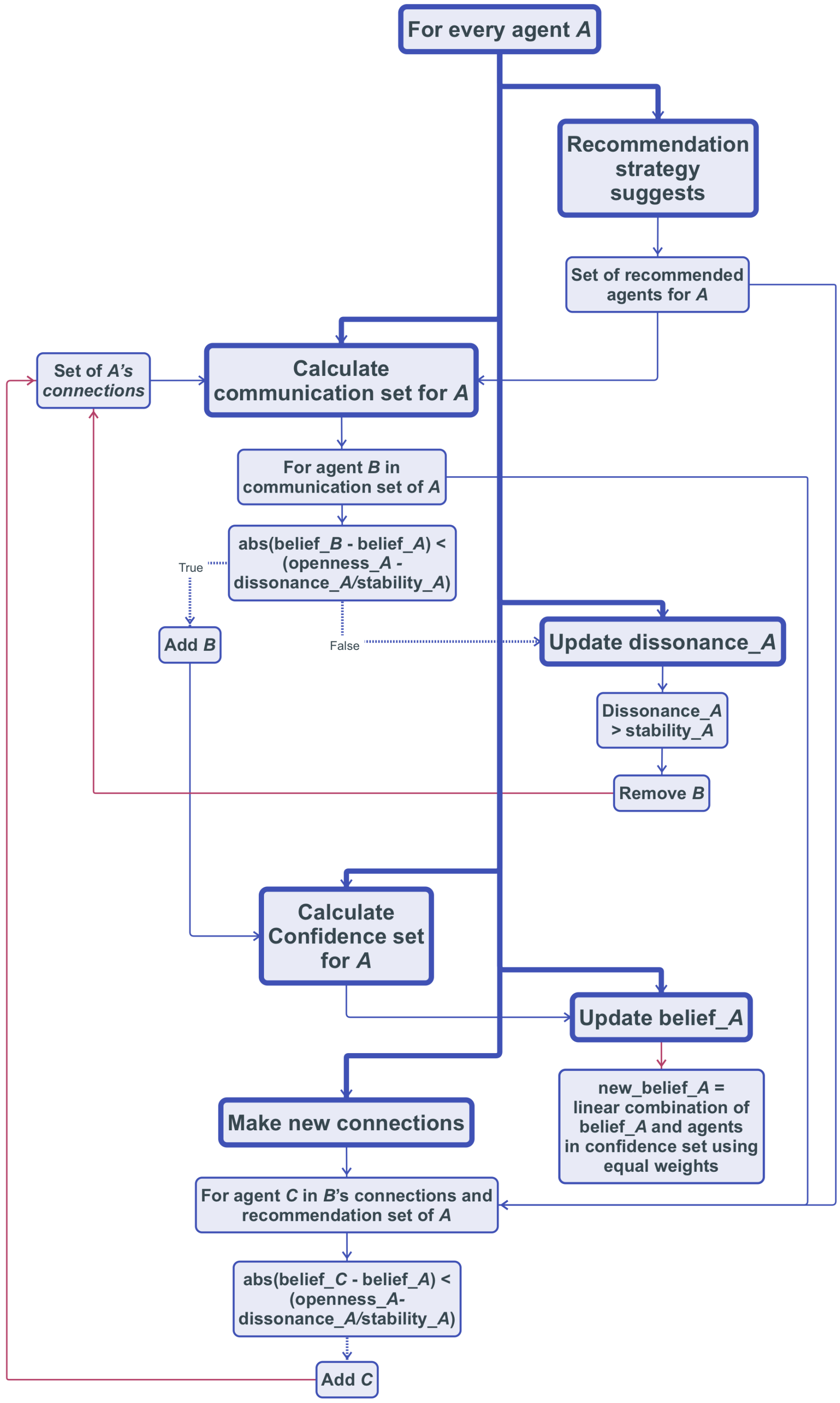


Figure 1: Flow diagram of the multi agent model displaying order of computations (from top to bottom) for current epoch (in blue), calculations that affect the calculations during the next epoch (in red), and conditional updates to parameters (dashed lines)

strategies. One of the three initial strategies by Festinger [Fes57] included adding new confirming cognitions. According to this strategy, the individual might actively seek out information that supports their position (see also confirmation bias [Nic98]). However, this strategy also works the other way around: the individual can attempt to reduce exposure to conflicting information [Fes57; Mcg17]. The pruning strategy in Ke, Haiming, Gang, and Yucheng [Ke+20] effectively reflects this second interpretation of the original dissonance strategy by Festinger [Fes57]: the individual agent X reduces the access to non-confirming information by terminating connections with any agent Y that at any epoch is connected to agent X but not included in its confidence set. However, implementing the strategy like this is not in line with the more recent findings on CD that were discussed before. According to those, a single exposure to conflicting information would be unlikely to result in terminating the connection. Rather this choice should depend on both, an agent's current level of dissonance which builds up over repeated encounters with conflicting information and an agent's resilience to CD [HH08; HK68; Mcg17]. Therefore, we decided to update the existing pruning strategy by Ke, Haiming, Gang, and Yucheng [Ke+20], taking into account those two additional parameters. In our implementation, an agent's current dissonance level increases every time an agent is confronted with conflicting information. Only if the agent's dissonance level exceeds a certain threshold, will the agent terminate the connection. This action will only have a small immediate beneficial effect on the agent's current dissonance. However, in addition to this small immediate effect cognitive dissonance will decline with time, reflecting the remarkable psychological resilience of human individuals that are capable to adjust to changes in their stress level reasonably well [FT16]. The dissonance parameters can optionally have an additional effect in our model. Instead of assuming that for an individual agent o remains static over time, the model discussed here allows the use of the ratio between an agent's current dissonance and the agent's dissonance threshold parameter (from now on referred to as "dissonance ratio") to penalise an agent's o parameter. Hence, the model allows for agents to become temporarily less open-minded when they experience increasing levels of distress in the form of dissonance. The penalising impact of the dissonance ratio on o can be regulated via a weight between 0 and 1: if this dissonance ratio weight is set to 1 the dissonance ratio can effectively reduce o by a value of 1 since its impact is not regulated at all. Correspondingly, setting the weight to 0 disables any impact of the dissonance ratio on o, effectively rendering the parameter static again.

This choice required the definition of a function according to which dissonance evolves over encounters that are either belief conflicting or confirming. Unfortunately, not much is known about the exact nature of this process [Mcg17] except for the findings that were already discussed above. Hence, a rolling average calculated over a selected window of an individual agent's contact history was chosen, reflecting both recency and frequency effects while at the same time being simple enough to generalise broadly onto the cognitive processes underlying CD. The contact history contains only 0s and 1s, with 1s reflecting a conflict and 0s reflecting a friendly encounter. Depending on the size of the window, dissonance progresses more gradually (larger windows) or behaves more erratically, switching between extremes quite frequently (smaller window). The figures 2 and 3 show how different window sizes affect dissonance over encounters.

In the implementation discussed here the dissonance threshold parameter is considered to represent yet another "Big Five" personality trait: Neuroticism. More specifically the dissonance threshold parameter represents the inverse of Neuroticism, emotional stability [MC87]. Emotional stability has been shown to be related to individuals' stress tolerance and resilience and therefore suits this parameter well [MC87; Jer20]. Very "emotionally stable" agents will not experience conflicting information as too disrupting but depending on the amount of encounters are still guaranteed to experience too much stress eventually.

Additionally, the updating strategy implemented by Ke, Haiming, Gang, and Yucheng [Ke+20],

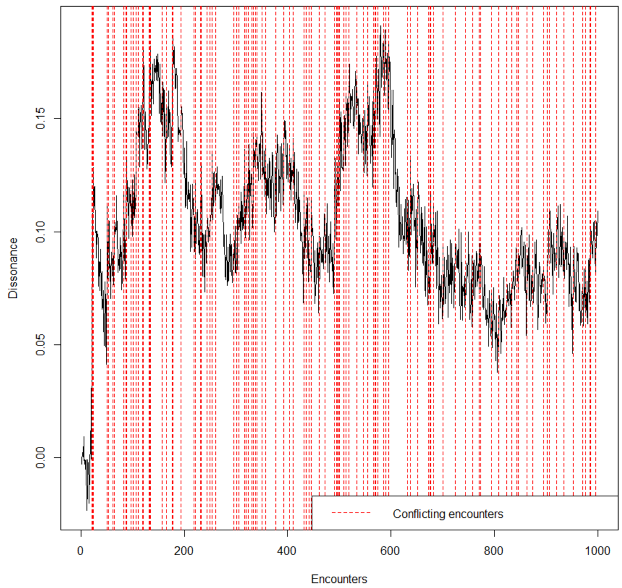


Figure 2: Dissonance over encounters for a window size of 100

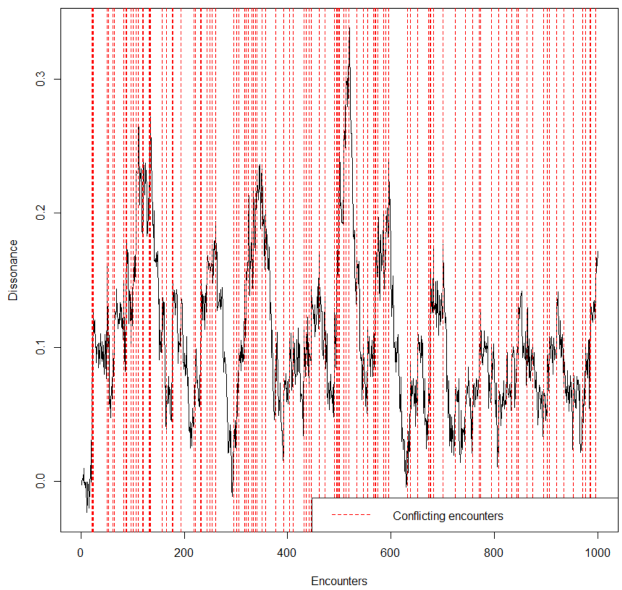


Figure 3: Dissonance over encounters for a window size of 30

allowing agents to form new connections with other agents that match their belief and are therefore more likely to provide access to information confirming with the agent's belief, is also designed to align with the strategy by Festinger [Fes57]. Ke, Haiming, Gang, and Yucheng [Ke+20] implement this by checking whether any agent Z that is connected with an agent Y is considered acceptable with regards to o of agent X that is itself connected to agent Y, reflecting a "a friend of a friend is a friend" approach. Here only minor adjustments were made to this updating rule, focusing mainly on the appropriate connection to the recommendation set introduced earlier. First, an agent W that was part of the recommendation set and that was later added to a confidence set of agent X should attempt to form a connection with X, regardless of whether or not a bridging agent in the middle exists. Additionally, a small beneficial boost is applied to an agent's dissonance level once a new connection is formed, reflecting the positive feeling when receiving confirming information (e.g. confirmation bias [Nic98]). Furthermore, a global maximum connection limit was implemented since it it unlikely that an individual agent will connect to all (confirming) other agents given that this would, in a real network, make it impossible to communicate efficiently. Moreover, this parameter allows to tailor the model to different kinds of networks such as physical social networks which are likely to involve less connections that individuals form on Facebook. Finally, a third and last parameter was introduced, acting mainly as a weight for the maximum connection limit: agents for which this personal connection weight has a value of 1 can form the maximum amount of connections permitted by the global maximum connection parameter. On the other hand, agents for which this personal connection weight is equal to .5 will only be allowed to form half of the amount of connections theoretically permitted by the global connection limit. This personal connection weight parameter is again to be sampled from an asymptotic distribution containing values between 0 and 1 and was inspired by yet another parameter of the "Big Five": Extraversion. Extroverts are claimed to actively seek out social interaction, while introverts on average prefer to spent time on their own [MC87]. Hence, the Extraversion inspired parameter in the model is also used as a weight to the dissonance boost discussed above, which is applied to an agent's dissonance level in case a very disruptive connection is terminated or a new connection is made, reflecting the assumption that a very extroverted agent will be impacted more from changes to the connections than an introverted agent. The final adapted model discussed here therefore contains parameters reflecting 3 out the 5 personality traits included in the "Big Five".

This model makes explicit two strategies: seeking out confirming information and blocking non-confirming information. Additionally, it allows to model individual differences in resilience to cognitive dissonance and how much individuals are willing to diverge from their own opinion when communicating/listening to other agents via the o parameter. The latter effectively corresponds to the original strategy on attitude integration by [Fes57]: if an individual agent accepts input from a broader spectrum, then the agent becomes not only less dogmatic but also, as a consequence, less susceptible to dissonance because less information is seen as non-confirming. In conclusion, this updated model is not only more in line with the idea that cognitive dissonance builds up rather than factoring into every decision and that strategies such as pruning are not constantly utilised but are only deployed if absolutely necessary, but also allows to test specifically the cognitive dissonance strategies discussed by Harmon-Jones and Harmon-Jones [HH08] and Mcgrath [Mcg17].

The computations necessary for every agent at any epoch are summarised in Figure 1 and will now be reviewed briefly. At every epoch, all agents X communicate with all their connections and the agents that were placed in their recommendation set Y, selected via one of the recommendation strategies discussed in the next section. Any agent Y with a belief within the accepted belief range of agent X denoted by o is placed in the confidence set of agent X. For any other agent a conflict is added to agent X's encounter history, which is used to calculate agent X's current dissonance level. Only if agent X's current dissonance level exceeds its dissonance threshold, it will terminate the connection with the agent that caused the conflict. The new belief for this agent is then calculated by equally weighing its own prior belief and all agents' beliefs that are in its confidence set. Finally, agent X will try to connect to any agent Z connected to an agent Y that is connected to X and all agents in the recommendation set, if those agents are within agent X's accepted belief range o.

To validate the choice of including parameters inspired by personality traits we approached Dr. Jeronimus [Jer20], who in his research focuses on personality and individual differences, and asked whether he could comment on the parameter choices and how we utilise them to influence the agent's actions. Based on the work by Garretsen, Stoker, D. Soudisand Martin, and Rentfrow [Gar+18], which showed that the outcome of the Brexit vote can largely be explained by the Openness trait, he considered its addition crucial to model the emergence of ECs in social networks. Furthermore, he considered the use of the emotional stability parameter, to reflect resilience to stress and especially cognitive dissonance, as appropriate. He remarked that individuals over time become more emotionally stable but can commonly be observed to also become more narrow minded, indicating that the results obtained from using a multi-agent model might not apply equally to different age groups. Furthermore, he reasoned that most variance in behavioural choices and psychological processes can be explained by the Extraversion trait. Therefore, in the final version of the model, the parameter is now used to determine how many connections an individual will form, as was discussed above.

3.2 Recommendation strategies

As discussed earlier, to increase the ecological validity even further and to allow insights into a possible tool of interventions, the model provided here makes use of different recommendation strategies to provide each agent with a set of suggestions for connections at every epoch – the recommendation set. Three strategies were considered for this report and are discussed in greater detail in the following sections.

3.2.1 Reinforcing Strategy

At each epoch the reinforcing strategy provides any agent X with a set of recommended agents W that fall within the range of beliefs deemed acceptable by agent X as reflected by its value for o. This strategy is similar to the mechanism implemented in the study by Madsen, Bailey, and Pilditch [MBP18], where agents seek out confirming information at every epoch by looking for other agents matching their own belief about the world. Moreover, as mentioned previously, this strategy resembles recommendation strategies in existing social networks closely, which also aim to provide access to contacts similar to oneself.

3.2.2 Randomising Strategy

At each epoch, the randomising strategy provides any agent X with a set of recommended agents W that are selected at random from the entire population P of agents in the network. The inclusion of this strategy is warranted by an increasing number of requests toward social media networks to stop utilising personal data to tailor content and contact recommendations to appeal to the individual's desires, and a recommendation algorithm that does not keep a model of a user can *only* make suggestions at random. Hence, applying this strategy does allow for an investigation of the consequences of such a non-adaptive algorithm for the network and the users, in this case agents, that interact within the network.

3.2.3 Neutralising Strategy

At each epoch, the neutralising strategy provides any agent X with a set of recommended agents W that are considered acceptable by agent X with regards to its value for o but are additionally closer to a neutral belief around 0.5. This strategy is therefore similar to the reinforcing strategy in that it also selects agents within the belief range of agent X but it is at the same time different from the reinforcing strategy because it attempts to nudge agents towards a predefined belief state, effectively manipulating all agents in the network. This strategy is particularly intriguing since, if it is successful, it might allow the network itself to control the belief state of the agent population independent of any agent's own belief.

3.3 Implementation details

Figure 4 shows the interface to our implementation. The entire project was written in Java and is available at [VTK20]. The repository in [VTK20] furthermore contains all the necessary scripts to re-create the experimental analyses we conducted. Furthermore, a detailed manual is provided that explains how to interact with the interface and which actions are currently implemented.

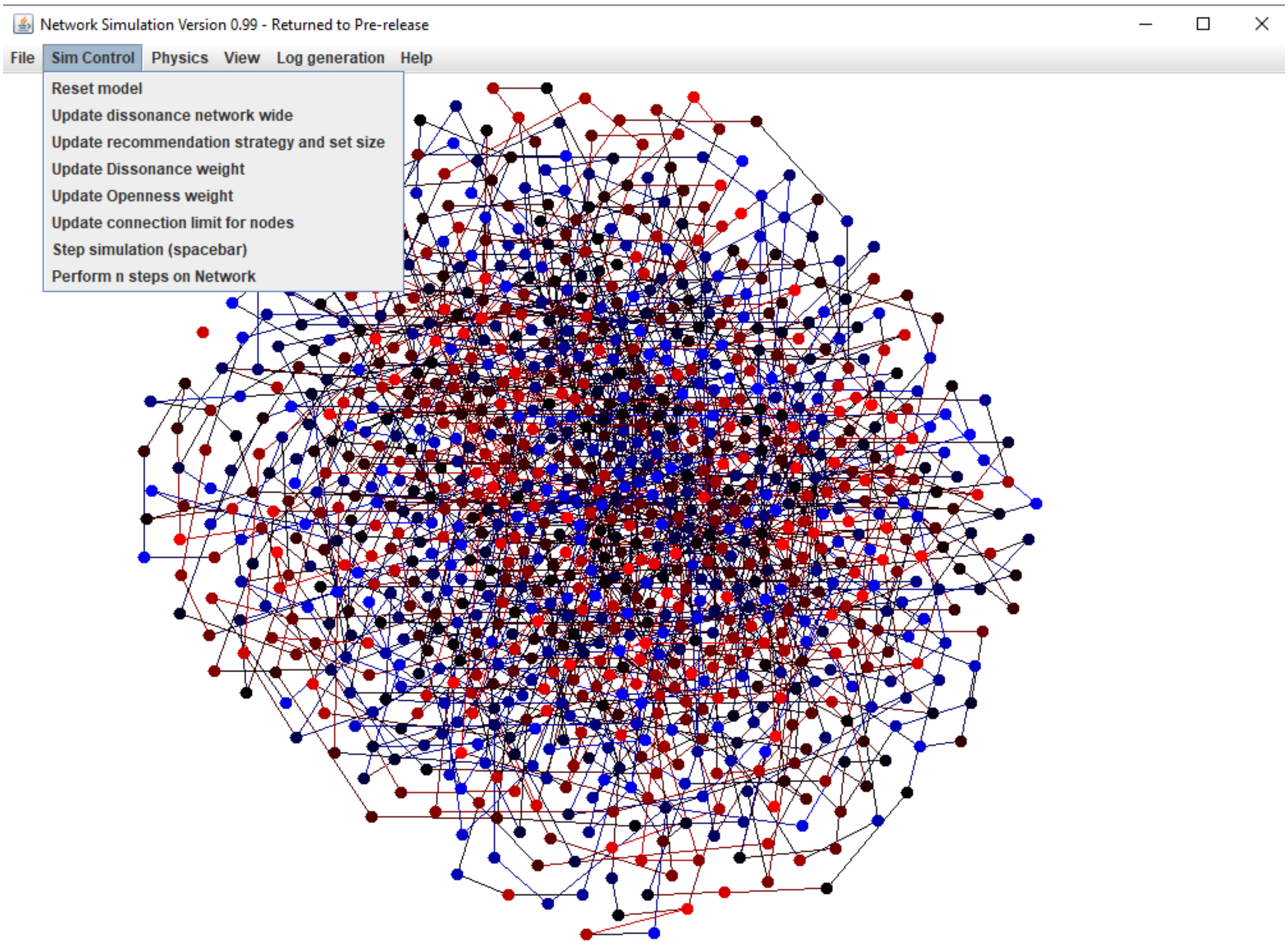


Figure 4: Current Interface to the implementation in Java

3.4 Experimental design

3.4.1 Agent parameter considerations

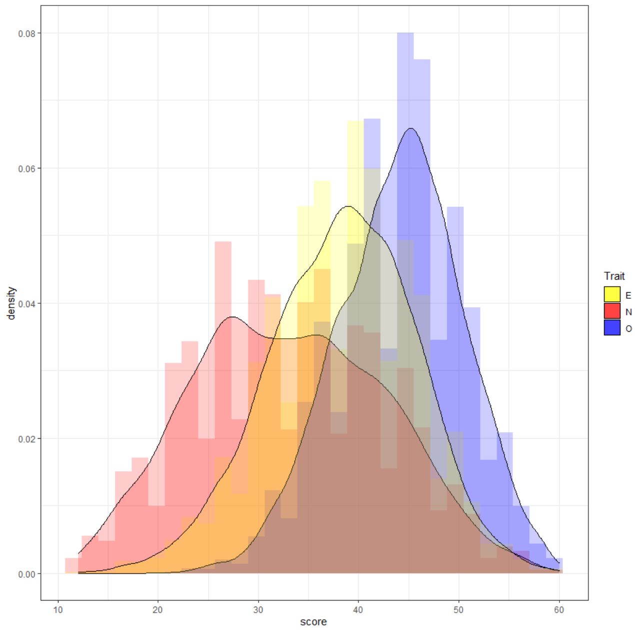
Since our model contains multiple parameters that have an impact on the behaviour of the agents, with some relating to psychological constructs such as personality traits, we decided to turn towards the literature to investigate possible starting values. Table 1 contains an overview of all the parameters related to the rules discussed in the last section. An asterisk next to a parameter name indicates that the parameter was used in a similar form in [Ke+20]. Note that more general parameters such as the size of the network are excluded from this table but will be considered and described more extensively in the section on simulations.

Parameter	Parameter Description	Values	Source
Initial belief*	The initial belief distribution in the entire network.	Uniform distribution between 0 and 1	[Ke+20]
Maximum belief range*	The maximum distance of another agent's belief an agent with an openness score of 1 would still consider.	0-1	[Ke+20]
Dissonance ratio weight	The weight determining how much current dissonance should penalise the openness of an agent.	0-1	
Openness	How open an agent is to information that differ invalue from their own belief. Acts as a weight for the maximum belief range.	Distribution estimated from sample with a mean of 0.67, a standard deviation of 0.13, a skew of -0.20, and a kurtosis of -0.14.	[Cen+17], [Oz15], [GH14], [Kri+15], [Jer20]
Extraversion	Determines how many connections an agent will engage in up to the maximum limit imposed by the maximum connection limit.	Distribution estimated from sample with a mean of 0.55, a standard deviation of 0.15, a skew of -0.19, and a kurtosis of -0.06.	[Cen+17], [Oz15], [GH14], [Kri+15], [Jer20]
Emotional stability	How much conflicting information an agent can tolerate before it becomes unbearable.	1 - value from a distribution estimated from sample with a mean of 0.44, a standard deviation of 0.20, a skew of 0.12, and a kurtosis of -0.62.	[Cen+17], [Oz15], [GH14], [Kri+15], [Jer20]
Dissonance window size	The window size of the moving average function used to calculate the current dissonance for an agent.	10-100 (see simulation model)	
Dissonance boost	The immediate beneficial impact on the current dissonance level of an agent in case a very disruptive connection is terminated or a new connection is made.	Extraversion value of agent * -0.05.	
Maximum connection limit	The maximum number of connections an agent with an extraversion score of 1 can form.	25 - 300	[Stu19], [Smi14]
Recommendation set size	How many recommendations an agent receives at every epoch.	10-100	[Kri+15]

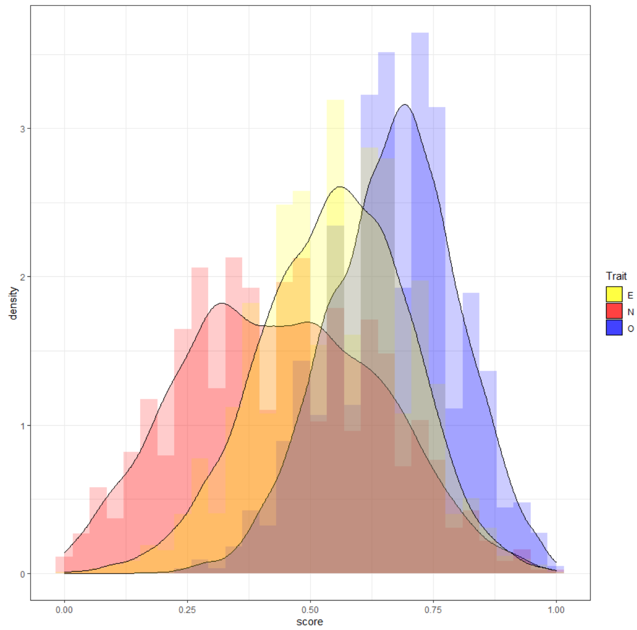
Table 1: Parameter considerations for the CD reduction strategies based on an initial literature research.

Most of the parameters in Table 1 are sampled from a network wide distribution: individual values for the agents are sampled from the same distribution. If no distribution is mentioned then the respective parameter applies globally: it is the same for all agents in the network. For the initial belief distribution we rely on the same initialisation used in [Ke+20]. For the distribution of openness, emotional stability, and extraversion we initially turned towards the psychological literature to identify possible starting point values but more importantly to search for evidence for a possible relationship between the constructs. Unfortunately, as is often the case when studying human subjects, the evidence is partially conflicting: while some studies indicate a strong negative relationship between openness and emotional stability [GH14], others find no such relationship [Oz15]. The distribution of the parameters itself has also been discussed before [Cen+17]. However, their sample size was extremely small, consisting of only 93 individuals for which personality traits were measured. Based on the inconclusiveness of our findings we decided to again approach [Jer20] who has also been involved extensively in the crowd-sourcing study "HoeGekIsNL" ("HowNutsAreTheDutch - HGNL", [Kri+15]) in which more than 10.000 participants that live within the Netherlands have participated so far. One of the cross-sectional questionnaires included in the study is the NEO-FFI-3 [HOF07; Kri+15], designed to measure the "Big Five" personality traits. We inquired whether it would be possible to use the available data from the NEO-FFI-3 [HOF07] to inform our parameter choices and after requesting formal approval we obtained a sample, consisting of 5117 individuals that had completed the NEO-FFI-3 [HOF07]. Hence, we decided to estimate the multivariate distribution that would best describe the data obtained [20]. This estimation of the distribution allows us to re-sample parameter triplets for any network size, ensuring that the co-variance structure of the personality traits remains intact in the entire network and that individual agents also receive parameters that reflect a realistic standing on the three personality traits [20].

The original scale of the NEO-FFI-3 [HOF07] ranges from 0-60 was transformed to a scale ranging from 0 to 1, enabling us to use the variables as weights. Figure 5 shows the density estimates for the three parameters from the original sample (5 a) and the re-scaled version (5 b). Table 2 depicts the mean and standard deviation values for the re-scaled trait variables. The Openness mean score is relatively high, indicating a relatively open-minded and liberal population from which the sample was obtained.



(a) Original sample



(b) Rescaled variables

Figure 5: Density estimates for the original sample (a) and the re-scaled sample fitting the limits imposed by our agent world (b) for the parameters extraversion (yellow), neuroticism (red), and openness (blue)

The data from the HGNL study also allowed for an investigation of the correlations between the different traits. Table 3 depicts the correlations between the different traits in the sample. A strong negative correlation between Neuroticism and Extraversion can be observed, indicating that high Extraversion scores are on average associated with lower Neuroticism scores. Additionally, a weak to moderate positive correlation between Extraversion and Openness indicates that high Extraversion scores are associated with high Openness scores at an elevated rate in the sample. Finally, the strong correlation between Openness and Neuroticism as observed by Gohary and Hanzaee [GH14] cannot be observed in the sample we obtained here. However, as pointed out by Dr. Jeronimus [Jer20], the relationship between the traits is likely to vary with age. The sample from the HGNL study consists of a wide range of age-groups, while the sample in [GH14] was much smaller and more restricted with regard to age, which might explain the differences in results.

Trait	Mean	Standard deviation
Neuroticism	0.44	0.20
Extraversion	0.55	0.15
Openness	0.67	0.13

Table 2: Descriptive statistics for the three different personality traits (after re-scaling)

Parameter	Neuroticism	Extraversion	Openness
Neuroticism	1.0	*	*
Extraversion	-0.46	1.0	*
Openness	-0.01	0.19	1.0

Table 3: Correlations between the three different personality traits

Finally, for the maximum connection limit the literature suggested a broad range of maximum limits that are enforced regardless of the extraversion parameter: The lower limit of the range in Table 1 is inspired by research on physical social networks and corresponds to the maximum size of connections an individual maintains in such a physical social network [Stu19]. The upper limit corresponds to the average number of friends individuals have on Facebook [Smi14]. As was already discussed earlier, not much is known about the exact processes underlying dissonance [Mcg17]. Therefore the numerical choice for the dissonance boost (-0.05), that will be applied for an agent with a perfect openness score, was chosen to be very small to have only a minor impact.

3.5 Simulations

3.5.1 Simulation study to investigate the formation of echo chambers

As discussed earlier, we wanted to investigate whether ECs will emerge in our model and how varying both the maximum belief range and the dissonance ratio weight would impact the formation of ECs. To that end, 10 networks with 1000 agents each were simulated for 1000 iterations. For all simulations the window size for the dissonance parameter was set to 30. Similarly, the maximum connection limit was set to 50 for all simulations, since the network to be simulated was relatively small in all cases. The reinforcing recommendation strategy was deployed in every simulation since it basically resembles the strategy used by individual agents in [MBP18], where agents reached out to other agents providing access to conforming information. 100 agents were recommended at every iteration to increase the speed at which agents would engage in interactions, allowing to draw conclusions even within the short time-frame of 1000 simulated iterations. The remaining parameters were initialised according to the values in Table 1.

In the first five simulations the weight for the dissonance ratio was fixed at 0 - removing its impact on the network. Additionally, for each of the first five simulations, a different maximum belief range was assigned, ranging from 0.05 to 0.25. For the remaining five simulations the maximum belief range was fixed at 0.25, while the dissonance ratio weight varied from 1 to 0.8 for every simulation.

Those 10 simulations were used to determine the impact of both hyper-parameters named above while controlling for the respective other one, allowing for an investigation of their individual contributions to the formation of ECs and polarisation in the network. Subsequently, the results obtained from those initial simulations were used to investigate the impact of three different recommendation strategies on the formation of ECs.

3.5.2 Simulation study to investigate the impact of different recommendation strategies on the formation of echo chambers

In earlier sections a strong argument in favour of a more complex but ecologically more valid model was that such a model would permit generalising results, obtained from investigating interventions, more easily to real social networks. In this report we decided to investigate the implications of the recommendation strategies, outlined in an earlier sections, for the network and individual agents.

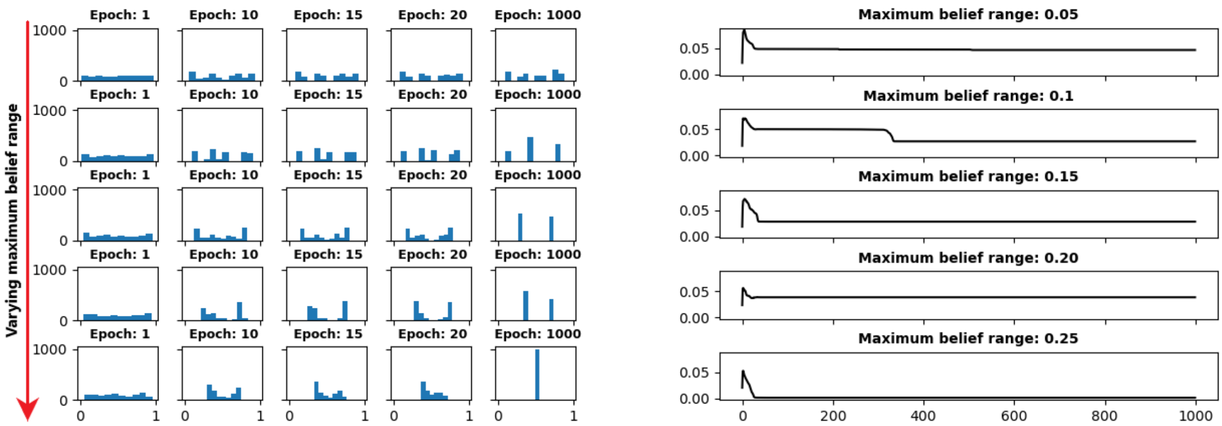
To investigate this, three additional networks were simulated with 1000 agents for 1000 iterations. In each of the three simulations, a different recommendation strategy was used, while all simulations were equal with regards to the remaining parameters. The maximum belief range for these three simulations was set to 0.25 and the dissonance ratio weight was set to 0.85, reflecting the most intriguing parameter combination observed in the earlier simulations with ECs forming despite a large maximum belief range. The remaining parameters were initialised to match the choices made for the first 10 simulations described above.

4 Results

4.1 Experimental findings

4.1.1 Simulation study to investigate the formation of echo chambers

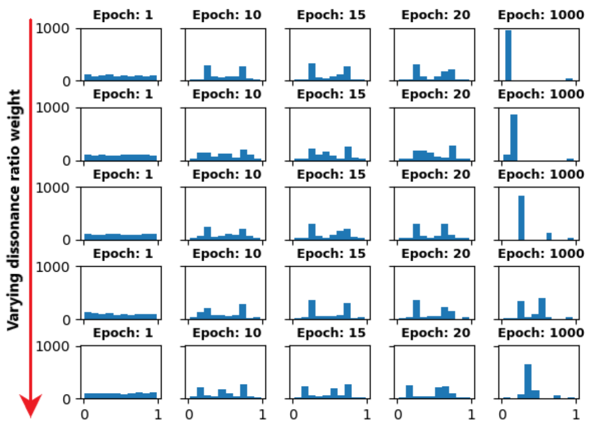
Figure 6a shows the distribution of existing belief values for five networks with differing maximum belief ranges considered acceptable by an agent with a perfect openness score (rows) for a selection of epochs (columns). Figure 6b depicts the average dissonance value at each epoch for all maximum belief ranges that were considered. The graphs included in this latter figure suggest that with increases in the maximum belief range considered by an agent with an openness score of 1, the average dissonance across the entire time-span of 1000 epochs decreases. Additionally, it can be observed that irrespective of the maximum belief range, the average dissonance value settles after approximately 50 epochs and remains stationary for the remaining iterations. The simulation with a maximum belief range of 0.1 forms an exception to this pattern with a visible decrease in the average dissonance value at approximately epoch 370. In all simulations except the last one, with a maximum belief range of 0.25 (1/4 of the possible belief range in the agent world), multiple belief state clusters (ECs) have formed at epoch 1000 (Figure 6a, last column).



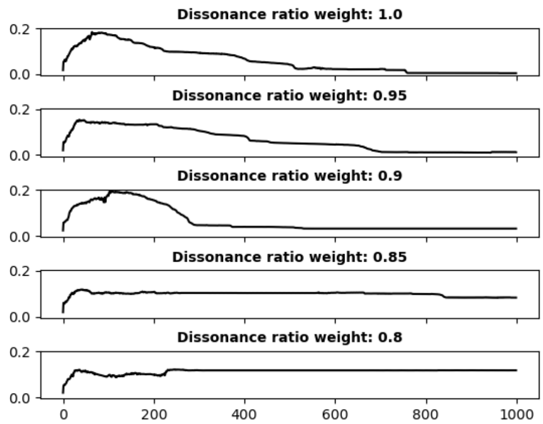
(a) Belief distribution for varying maximum belief ranges (b) Average dissonance for varying maximum belief ranges

Figure 6: Results from first 5 simulations. Each row corresponds to a different maximum belief range. In (a) each column further shows a different epoch.

Figure 7a visualises the same information as Figure 6a, except that for the simulations depicted here the maximum belief range was fixed at a value of 0.25 while the dissonance ratio weights varied for each situation. Figure 7b again depicts the trajectories of the average dissonance value in the network for the different simulations. Evidently, belief state clustering at epoch 100 can be observed again (Figure 7a, last column) with the shape and number of clusters varying drastically for different weights. The average dissonance values remain higher in those last five simulations when compared to the first five simulations, while the patterns of the individual trajectories also differ drastically for different weights: for weights larger than 0.85, the average dissonance first increases rapidly before slowly starting to decrease towards 0. For the remaining weights the dissonance value remains almost constant across the entire time-span simulated.



(a) Belief distribution for varying dissonance ratio weights

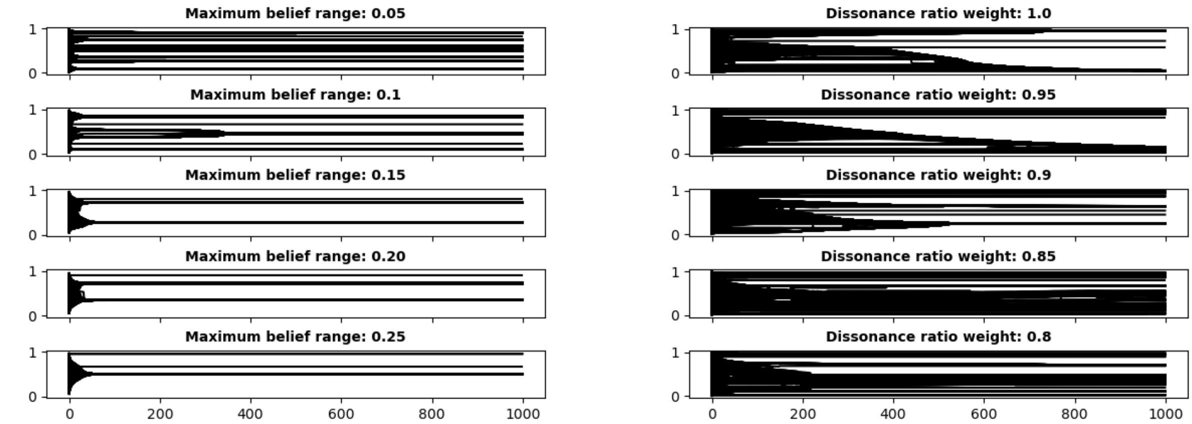


(b) Average dissonance for varying dissonance ratio weights

Figure 7: Results from last 5 simulations in which the maximum belief range was kept fixed at 0.25. Each row corresponds to a different dissonance ratio weight. In (a) each column further shows a different epoch.

Figure 8 shows the individual belief trajectories for the first (a) and last five (b) simulations over time. The individual belief trajectories reveals that the histogram plots obscured persistent individual differences in beliefs for the simulation in which the dissonance ratio weight was fixed at 0 and the maximum belief range to be considered by an agent with an openness score of 1 was 0.25 (8a, final graph). However, compared to the other extreme in which the maximum belief range was set to 0.005, where almost as many different beliefs appear to exist as there are agents in the network, the network can still be said to have converged drastically, especially since a visual inspection of the network reveals that that the network has formed one giant cluster coloured in red which indicates that all beliefs are slightly larger than 0.5.

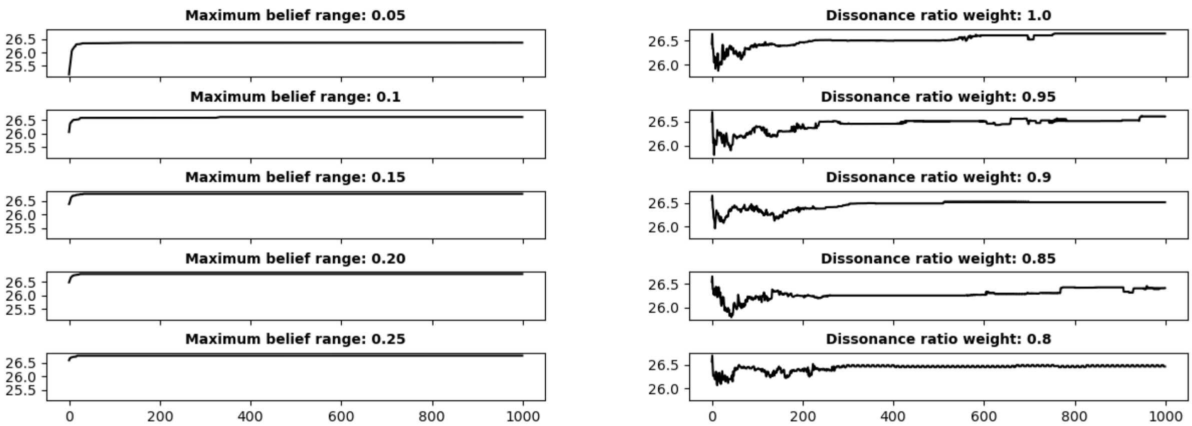
The trajectories for the simulations in which the dissonance ratio weight varied (Figure 8b) show much more fluctuation in the belief states of the individual agents compared to the trajectories for the simulations in which the weight remained fixed at 0 (Figure 8a).



(a) Individual beliefs for varying maximum belief ranges (b) Individual beliefs for varying dissonance ratio weights

Figure 8: Results from all 10 simulations. (a) shows the individual belief trajectories over epochs for varying maximum belief ranges. (b) shows the individual belief trajectories over epochs for a fixed maximum belief range of 0.25 and varying dissonance ratio weights.

Figure 9 shows the average number of neighbours for every epoch for the first five (a) and last (b) simulations. The patterns observed for the first five simulations are similar to what could be observed in the dissonance trajectories: the networks appear to reach a point of homeostasis at which point the trajectories become stationary. The same cannot be said about the patterns revealed by the plots for the last five simulations (Figure 9) in which the dissonance ratio weight differed from zero: the trajectories are more noisy and show much more fluctuation, especially during the first 200 iterations.

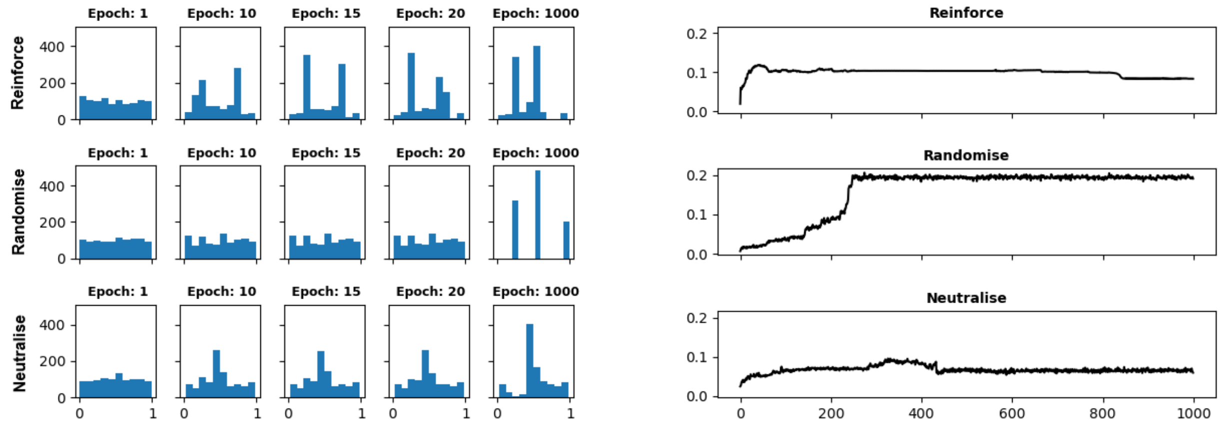


(a) Neighbour count for varying maximum belief ranges (b) Neighbour count for varying dissonance ratio weights

Figure 9: Results from all 10 simulations. (a) shows the average number of neighbours over epochs for varying maximum belief ranges. (b) shows the average number of neighbours over epochs for a fixed maximum belief range of 0.25 and varying dissonance ratio weights.

4.1.2 Simulation study to investigate the impact of different recommendation strategies on the formation of echo chambers

Figure 10a shows the histogram of the belief distribution at different epochs for the three networks in which different recommendation strategies were deployed. Figure 10b shows the average dissonance values at every epoch for those three networks. The first row in both Figures depicts the results for the reinforcing strategy. The histograms reveal that the network moved closer towards two belief state clusters (ECs) which became narrower with time. The average dissonance value remains consistently low. The second row shows the results for the network in which the randomising strategy was deployed. The histograms reveal clear segregation, while the dissonance graph indicates a steep increase in the average dissonance value until the value becomes stationary at a stable but high level. The final row shows the results corresponding to the network where the neutralising strategy was applied. The histograms indicate that the network largely converged towards a single belief state around 0.5. Additionally, the dissonance graph reveals that the average dissonance value in the network remained fairly low and even reached a level below the average value in the network where the reinforcing strategy was applied after around 450 epochs.

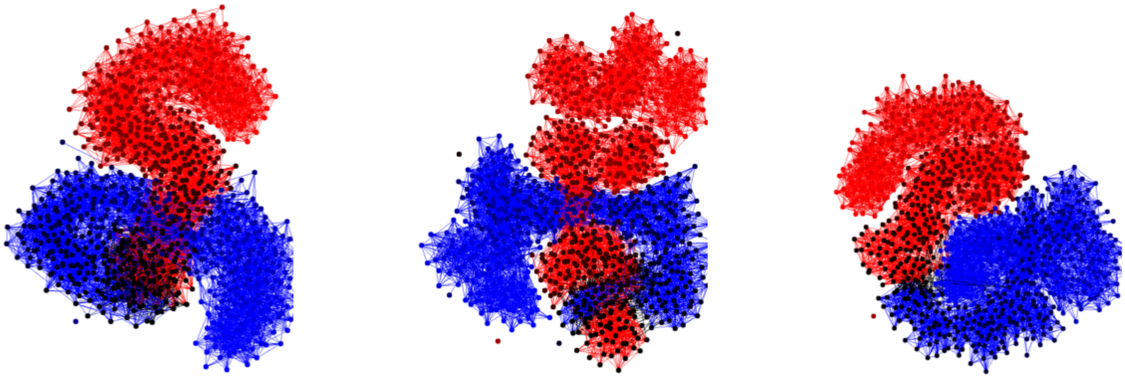


(a) Belief distribution for different strategies

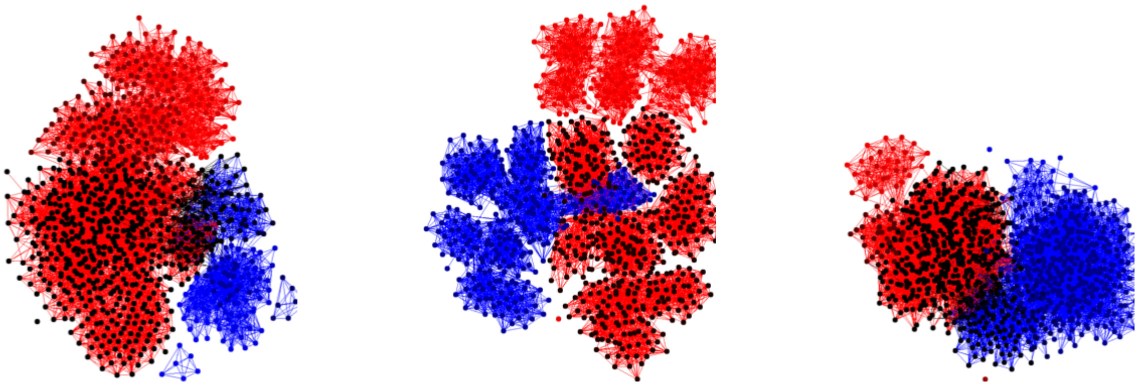
(b) Average dissonance for different strategies

Figure 10: Results from all 3 simulations. Each row corresponds to a different strategy (reinforcing, randomising, neutralising). In (a) each column further shows a different epoch.

Figure 10 shows the states of the network after a single epoch (a) and 1000 epochs (b). Video recordings of the simulations, showing the changes to the network over time, can be found in [KVT20]. The comparison across time-points suggests that the network in which the neutralising strategy was applied has changed the most over time (left-most network). The network in which the randomising strategy was applied (network in the middle) already showed signs of segregation after a single epoch and appears to have become even more segregated with time. Finally, the network in which the reinforcing strategy was applied (right-most network) indicates two large belief clusters, with more extreme beliefs existing in the agents at the poles of the clusters.



(a) Three strategies after epoch 1



(b) Three strategies after epoch 1000

Figure 11: Results from all 3 simulations. (a) shows the three networks to which the neutralising, randomising, and reinforcing recommendation strategy was applied after the first epoch. (b) shows the three networks after 1000 epochs.

4.2 Interpretation of findings

4.2.1 Simulation study to investigate the formation of echo chambers

The first five simulations that were part of this simulation study were designed to generate sufficient data to investigate the first research question. Hence, the maximum belief range was increased until a network settled at a single visible cluster at epoch 1000. The remaining five simulations were conducted to investigate whether EC formation would progress differently in case agents would become more narrow-minded when experiencing increasing levels of distress, which reflects the second research question. Figure 6a indicates that ECs do form reliably as long as the maximum belief range does not exceed 1/5th of the possible belief range in the agent world. The formation of ECs can already be observed after as few as 10 iterations in which agents went through the entire cycle of exchanging information, updating their belief state, and updating their set of connections: For the simulation in which the maximum belief range was set to 0.20, two belief "camps" already become visible after epoch 10 (row 3 in Figure 6a). Over time those "camps" then become narrower and more extreme until, after 1000 iterations, only two extreme belief states can be observed. The individual belief trajectories nicely highlight this convergence over time as well, as becomes visible in Figure 8a. Additionally, these individual trajectories suggest that, while the belief distribution suggests convergence, there actually still exist differences in the beliefs between agents for the simulation in which the maximum belief distance was set to 0.25. This general pattern aligns well with the original description of ECs according to which the belief of each individual becomes more extreme and narrower over time due to the constant reinforcement of information acquired from other individuals in the chamber that share a similar belief [JC09] and yields compelling evidence that, in our ecologically more valid model, ECs do form consistently until the maximum belief range exceeds 20% of the possible belief range in the agent world, providing first insights into the first research question and permitting the conclusion that our model replicates the findings by both Madsen, Bailey, and Pilditch [MBP18] and Ke, Haiming, Gang, and Yucheng [Ke+20].

Figure 6b depicts the average levels of dissonance per epoch for the first five simulations. Evidently, a negative relationship exists between the maximum belief range and the average levels of dissonance: as the maximum belief range increases, the average dissonance levels per epoch, aggregated over time, decrease. This relationship aligns with the claim outlined in earlier sections that, as agents become more open-minded and thus have broader belief states, they more easily integrate conflicting information into their own world-view reducing their experience of distress associated with cognitive dissonance [Fes57; Mcg17]. However, the average dissonance levels also remain relatively low for the smallest maximum belief range considered (0.05), indicating that the maximum belief range itself is not the main factor influencing the stress levels in the network. Indeed, we hypothesised that dissonance levels would remain low on average in a network utilising the reinforcing recommendation strategy since this strategy should ensure consistent access to information in line with an individual's belief state [Mcg17; Fes57]. The findings observed here suggest that the reinforcing strategy does in fact impact average dissonance levels in the network and generate initial evidence for this hypothesis.

The results from the last five simulations included in this simulation study allow for both a more refined investigation of those findings, and an investigation of the second research question reviewed above. Figure 7a again depicts the belief distribution over time for the last five simulations in which the dissonance ratio weights were varied while the maximum belief range was kept fixed at a value of 0.25. Apparently, allowing agents to become more narrow-minded with increasing levels of distress complicates the interpretation of the relationship between the maximum belief range and the average dissonance levels in the network. For all the different dissonance ratio weights (1 - 0.8) ECs do form, albeit to a different extent, and persist even after 1000 iterations, suggesting that if agents do become more narrow-minded when they experience elevated levels of distress the formation of ECs can be observed for maximum belief ranges larger than 0.25.

As indicated by the graphs in Figure 7b, the average dissonance levels vary more drastically over time in the last five simulations than they did in the first ones. For the three highest dissonance ratio weights (1 - 0.9) the average dissonance in the network increases steeply during the first 100 epochs before slowly starting to decrease until eventually reaching 0. This pattern suggests that during the first 10 epochs, where the agents start to form "camps" matching their beliefs, they still are confronted with information outside of their acceptable belief range quite frequently. Due to the high dissonance ratio weights the acceptable belief range of those agents will therefore be reduced drastically, further enhancing the likelihood that they will be confronted with conflicting information during the next iteration. This initially leads to a vicious cycle, explaining the steep increase in the average dissonance levels, where agents are confronted with conflicting information, which narrows their acceptable belief range and therefore leads to even more confrontations with conflicting information. This is further supported by the graphs describing the average number of neighbours over time for the 10 simulations included in this study depicted in Figure 9: The graphs belonging to the last five simulations show more drastic fluctuations early on when compared to the graphs for the first five simulations, indicating that the agents early on more frequently terminate and form new connections in the last five simulations reflecting the effect of the vicious cycle they find themselves in, as was outlined above. The fact that the average dissonance levels then begin to decline after 150 iterations can be attributed to the reinforcing recommendation strategy, adding further support for the hypothesis that this strategy does have an alleviating impact on the average dissonance levels in a network over time: the reinforcing strategy, providing access to conforming information via the recommendation set, reduces the access to conflicting information, providing agents experiencing distress with a set of like minded individuals with which they will readily form connections to increase their access to information within their own belief range. This again becomes visible in the average number of neighbours over time, which later stabilises for the last five simulations and then progresses similar to what could be observed in the graphs from the first five simulations (Figure 9). Therefore, the formation of ECs persists in this case even though the maximum belief range is set to 0.25. Additionally, the ECs that formed in these simulations are much more extreme which is exemplified by the fact that they are located at the extremes of the agent world belief range. Hence, while the reinforcing strategy does have a beneficial impact on the network wide dissonance, it also promotes polarisation in the network.

For the remaining two dissonance ratio weights (0.85 and 0.8) dissonance levels first increase and then remain relatively stable around approximately 0.15. At the same time the ECs that formed during the 1000 iterations are not as extremely polarised as was the case for the other dissonance ratio weights (1 - 0.9) which suggests that the self-reinforcing effects leading to the vicious cycle observed earlier require quite high dissonance ratio weights to become clearly visible. However, they are likely still at play: since for the entire time-span the average dissonance levels remained elevated, agents continue to have access to conflicting information which, in combination with the impact of the reinforcing recommendation strategy discussed earlier, should lead to further polarisation that could not be observed in the restricted time-span considered here.

4.2.2 Simulation study to investigate the impact of different recommendation strategies on the formation of echo chambers

To answer the last two research questions an investigation of both the visualisations of the networks over time as well as descriptive statistics including the belief distribution and average dissonance levels per epoch was conducted. Figure 11a depicts three networks utilising the reinforcing, randomising, and neutralising strategy, respectively, after a single iteration. Already after a single epoch,clear differences in the shapes of the three network plots exist. The randomised network already hints at increased levels of polarisation with a group of extreme belief clusters forming at the top of the network (in red). Additionally, in the randomised network only a relatively small number of agents, located at the lower end of the network, appears to have formed connections with other agents falling outside of their belief range when compared to the other two networks.

The other two networks on the other hand show larger "borders" where the two large chambers are connected via agents that form connections with agents from both chambers. It is worth to point out that the emergence of such an agent type, acting as an interface between both chambers, would not have been possible in the work by Ke, Haiming, Gang, and Yucheng [Ke+20] and Madsen, Bailey, and Pilditch [MBP18] where agents always terminated connections with agents that were outside of their acceptable belief range. As already discussed earlier, such a radical behaviour ignores the fact that terminating a relationship likely comes at a cost and whether or not an agent is willing to pay that price will depend on multiple factors including the salience of the incoming information that clashes with the agents own belief [Mcg17; Fes57]. The visualisations here add support for the claim that, should an agent have sufficient access to belief conforming information, such conflicting connections can and will be maintained. Such an agent will still experience a negative encounter and exclude the agents too far away in their belief from it's confidence set. However, should an agent's belief shift closer to the information provided by those conflict generating agents the agent will begin to integrate the information into the own belief state, at which point these "sleeper" connections will contribute to a rapid belief change.

Figure 11b shows the three networks after 1000 iterations. Evidently, the differences in the shapes between epoch 1 and epoch 1000 differ for the recommendation strategies. In the reinforced network the two chambers can still be observed. The gradient in colours is now easier to interpret: the clustering of almost black nodes around the "border" suggest that the belief states of agents close to this border have settled at a similar belief around 0.5, while more extreme beliefs exist at the poles of the network. This pattern intuitively makes sense, since the agents around the "border" had access to a broader range information than those at the poles.

The randomised network also shows only minor changes in the overall shape. However, as was already indicated after a single epoch, this network now depicts extreme segregation into multiple sub-graphs after 1000 epochs.

Finally, the neutralised network shows the most drastic changes: a large imbalance in colours can be observed, with a majority of the nodes having formed connections rendered as red, indicating that the agents belonging two this connection both have a belief larger than 0.5. Additionally, most nodes are actually rendered with black colours. Those two findings suggest that in this network convergence towards a more neutral belief around 0.5 can be observed. At the same time the poles of the networks are again populated by agents with more extreme beliefs. In fact, the extreme belief group at the bottom of the network, rendered in blue, only maintains few connections with the large "neutral" cluster, indicating that this group might actually separate from the large cluster later.

Figure 10 shows the belief distribution for a selection of epochs and the average dissonance level per epoch for the three strategies investigated in this simulation study. The first row depicts the results corresponding to the reinforced network. The belief distribution, in Figure 10a, supports the observations made based on the visualisation of the network itself after 1000 iterations: two distinct belief chambers appear to have formed that are however relatively close to each other. The belief distribution also indicates the presence of more extreme beliefs, that were populating the poles in the network. One of the research questions of interest focused on the time necessary for a network to settle on a belief states (or on multiple ones). Evidently, for the reinforced network the formation of the chambers can again already be observed at epoch 10. Over the remaining 990 epochs those chambers become narrower and more extreme. As was already discussed before, the average levels of dissonance in the network remain fairly constant and low over time as indicated by the first graph in Figure 10b.

The second row in Figure 10 shows the results for the randomised network. Over the first 20 epochs the belief distribution does not indicate drastic changes. However, after 1000 iterations the belief distribution shows extremely narrow belief chambers around 0.25, 0.6, and 0.9. This again supports the visualisation of this network after 1000 iterations, where clear segregation into multiple sub-graphs was visible. Additionally, the graph depicting the development of the average dissonance over time diverges strongly from the graph corresponding to the reinforced network: the average dissonance levels show a non-linear, almost exponential, increase until reaching an abrupt halt after approximately 230 epochs. The initial increase in dissonance levels can be attributed to the randomising recommendation strategy, which constantly confronts agents with information outside of their acceptable belief range. Hence, the agents in the network will become increasingly more narrow-minded as a result of their increasing stress levels. To combat this the agents begin forming connections with like-minded individuals. However, agents will only engage in connections with other agents that are within their, at this point, extremely narrow range of acceptable beliefs which ultimately results in the significant amount of segregation visible after 1000 iterations. The abrupt halt in the increasing stress levels reflects the point where the agents have formed a sufficient amount of connections with other agents matching their narrow belief states. Henceforth, while still being confronted with conflicting information via the recommendation set, each agents has sufficient access to conforming information preventing the dissonance levels from rising further. This comforting characteristic of the chambers, while allowing the agents to minimise distress, prevents the agents from ever attempting to engage with other agents that might have different beliefs, ultimately resulting in persistent polarisation. These results here generate initial evidence suggesting that recommendation strategies which promote a broader belief state lead to higher levels of cognitive dissonance (research question 3b). Finally, the results here contradict the third hypothesis that a randomising strategy should reduce the polarisation in a network because each agent would have access to a broad range of different beliefs via the recommendation set. If agents become more narrow-minded as they experience increasing levels of distress the opposite effect can be observed: the strategy promotes segregation as described above. However, if the dissonance ratio weight is fixed at zero this strategy does eventually lead to convergence, similar to the neutralising strategy providing additional insights into research question 3c, strengthening the claim that if individuals become more narrow-minded with increasing levels of distress the impact of the recommendation strategies will differ. In that case however, the dissonance levels remain much higher in the entire network across time, with almost all agents experiencing distress exceeding their dissonance threshold at which point individuals would likely engage in the most radical actions to reduce the amount of distress they experience [Mcg17; Fes57], even leaving the social network entirely, which the model currently does not permit for.

The final row in Figure 10 depicts the results for the neutralised network. For this network the belief distribution after epoch 10 is again already indicative of the belief distribution after 1000 iterations, as was the case for the reinforced network. Over time a majority of the agents in the network shift their belief towards a relatively neutral belief around 0.5. However, even after 1000 epochs there still exist a minority of agents with an extreme belief as was also visible in the network visualisation itself. The average dissonance levels per epoch constantly remain low in this network when compared to both the randomised and reinforced network. The latter one can be explained by the fact that, because the majority of agents shifts closer to a neutral belief, any agent is less likely to get into contact with conflicting information when reaching out to the agents connected to it's current set of connections, reflecting the "a friend of a friend is a friend" approach [Ke+20]. The results here suggest that the network itself, via a recommendation strategy, can drastically influence the belief distribution in a network. Since the recommended agents lie within an agent's acceptable belief range this strategy prevents segregation, as was observed in the randomised network. Additionally, since this recommendation strategy actually also promotes a broader belief state the answer to research question 3b needs to be adapted to reflect that the actual relationship between the dissonance levels in a network and broadening recommendation strategies is more complex and that the impact of the strategy will depend on its exact implementation. Finally, the results presented here provide convincing evidence that specific recommendation strategies can reduce the polarisation in a network, which provides an answer to research question 4.

Combining the results obtained through this simulation study suggests that different recommendation strategies do in fact impact the formation of ECs differently (research question 3). Additionally, the time it takes the network to converge to different belief states (research question 3a) also depends on the actual recommendation strategy implemented in the network: In the reinforcing network and the neutralising network the convergence to specific belief states was already visible after 10 iterations and only become more extreme with time. The visualisation of the randomised network on the other hand hinted at segregation already after epoch 1 while the belief distribution remained relatively uniform during the first 20 iterations. Finally, since even in the neutralising network EC formation was visible, albeit to a drastically reduced degree, the results here generate support for our second hypothesis and add weight to the claims made by Madsen, Bailey, and Pilditch [MBP18] that EC formation is a robust phenomenon that can be observed reliably in social networks.

5 Conclusion

5.1 Discussion

5.1.1 Simulation study to investigate the formation of echo chambers

The results from this simulation study complimented the findings by Madsen, Bailey, and Pilditch [MBP18] by showing that ECs constantly form in networks for a variety of maximum belief ranges. Additionally, the results suggested that, should individuals become more narrow-minded over time, ECs would form for even larger maximum belief ranges, exceeding 1/4th of the agent world belief range. Given that in our model a belief of either 0 or 1 reflects extreme opposites in any real world debate and 0.5 corresponds to a more or less neutral position, it seems unlikely that in a real social network any participant would ever consider the real world equivalent of a maximum belief range of 0.25. The magnitude of the range for which ECs are still observable, paired with the fact that the openness distribution of our Dutch sample resembles a very open-minded population, suggest that ECs are likely to emerge even in very liberal networks. This observation does not suggest that the formation of ECs will be restricted to specific parts of a population and rather suggest that ECs are likely to form across broad and different populations. This supports, and extends, the findings by Madsen, Bailey, and Pilditch [MBP18] but does not yet answer the question why ECs form so reliably.

Based on theoretical considerations, ECs are so attractive because they provide access to belief confirming information, which reduces distress [Fes57; Mcg17]. Since the model discussed in this report uniquely allows to manipulate an agent's exposure to different information and explicitly takes into account the experience of distress that arises when an individual agent is confronted with conflicting information, it allows for valuable insights into the intertwined relationship between the experience of cognitive dissonance and the formation of ECs. Gaining these insights from studies involving human subjects is difficult since in those cases the control about the access to different information and beliefs is limited, which is why this multi-agent model discussed here is optimally suited to extend traditional research techniques used across the social sciences.

The findings from the first simulation study additionally suggest that the reinforcing strategy, promoting access to belief confirming information and thereby promoting the formation of ECs, reduces the stress levels across the entire network independently of whether or not agents become more narrow-minded with increasing levels of distress. These findings are well in line with the idea that humans actively seek out confirming information and might therefore explain why the recommendation strategies used in social media networks are so effective in keeping the user actively involved on their platform: the platform provides sufficient access to belief confirming information, which is more satisfying than being confronted with conflicting information and since, outside of the social media network, an individual will constantly be exposed to information clashing with their own beliefs, the individual will always return to the platform to ensure sufficient access to belief confirming information.

5.1.2 Simulation study to investigate the impact of different recommendation strategies on the formation of echo chambers

This simulation study indicated that ECs did form reliably across a wide range of recommendation strategies, further supporting the claim above that ECs do emerge because they offer individuals access to confirming information preventing them from experiencing too much distress. In that sense EC formation could be claimed to be *momentarily* adaptive since their constellation reduces the individual's experience of negative arousal and distress. However, in the long run ECs are likely to promote phenomena associated with negative outcomes for the individual such as for example groupthink and deindividuation [CG04; AWA10; FT16].

The results from the second simulation study additionally provided insights into the emerging consequences for a network should a non-adaptive recommendation algorithm be deployed that does not keep track of user preferences or beliefs: the network showed an extreme degree of polarisation, resulting in actual segregation of the entire graph into multiple sub-graphs. At the same time the average levels of dissonance per epoch were strongly elevated in this network when compared to the other two. These findings suggest that utilising a non-adaptive recommendation algorithm might not necessarily result in reduced polarisation. Rather the increasing number of confrontations could motivate users *even more* to seek out information that aligns with their own beliefs. An investigation of the impact of this strategy in case agents do not become more narrow-minded as they experience increasing levels of distress suggested that under those conditions the network would eventually converge. However, as mentioned before, the average dissonance levels in such a network were extremely high to the point at which users might actively decide to leave the network. In any case, the results of our study suggest that an implementation of a non-adaptive random recommendation system might have grave consequences for the users of a social network which is why it should be investigated carefully whether such an algorithm would offer any real benefit when compared to current implementations.

In contrast, the results from the neutralised network showed the potential consequences for a network should a recommendation algorithm be utilised that does keep track of a users belief and serves a predefined goal. The goal, or target belief, was selected to be 0.5 in this study to show that the network itself can in fact reduce polarisation across its population. The results confirmed this assumption and showed that, if the algorithm maintains a reliable estimate of an agent's belief and acceptable belief range and therefore is capable of making biased suggestions that are not too extreme to cause distress, it can influence, and drastically alter, the belief distribution across the entire network. While these results show that a carefully crafted algorithm can contribute to depolarisation in a network, they also highlight the extreme potential for abuse of such an algorithm: Instead of aiming at depolarisation such an algorithm could easily enforce the opposite and even shift the belief distribution across the network towards any arbitrary target value selected by the instance controlling the algorithm.

In conclusion, this simulation study suggested that the network itself, via an appropriate recommendation strategy can influence not only the degree of polarisation in a network but also the belief distribution and stress levels across the population within the network.

5.1.3 General Discussion

Despite offering valuable insights into the interplay between cognitive dissonance and the formation of ECs as well as allowing for an investigation of the impact different recommendation strategies would have on a social network, there are some limitations to the utilisation of the model as discussed here. For example, the current implementation does only allow for a single network graph initialisation algorithm that connected nodes before the first iteration. The current implementation connects agents proportionally: if an agent already has multiple connections any new agent is more likely to be connected to this agent rather than another agent with less connections. Furthermore, this algorithm connects agents irrespective of their own belief states. Hence, an agent with a belief of 0 is as likely to be connected to an agent with a belief of 0 and an agent with a belief of 1 if both agents have the same number of connections. While this is unlikely to represent the formation of real social networks were users will be very selective when joining, this choice was made to reduce any potential bias that could be introduced by pre-connecting nodes taking into account their belief ranges. Correspondingly it could be expected that such an initialisation algorithm, taking into account belief ranges, would actually increase the polarisation even further, which should be investigated more closely in future research.

Additionally, while we reached out to an expert on personality research and based the actions available to agents on psychological theories discussed extensively in the literature, it remains non-testable whether our utilisation of the personality traits is reflective of their true implications. To investigate this, future research should investigate the polarisation in existing social networks and relate the findings to estimates of the personality trait distribution that is likely underlying the population in the network. For example, as recommended by Dr. Jeronimus [Jer20], it should be investigated whether the impact of the recommendation strategies will differ for different age groups that are likely show a different distribution of personality traits. Furthermore, our sample was restricted to Dutch citizens. While this ensures that the results are more likely to generalise to social networks in the Netherlands it remains to be seen whether the same results would emerge within other nations. It would be especially interesting to investigate how the results of our model would differ when a less open-minded population is used, since the Dutch individuals in our sample consistently scored high on the openness scale in the NEO-FFI-3 [HOF07; Kri+15].

Moreover, future research should investigate how the results presented here will differ in case individual agents will be allowed to leave the network. For example, a rule could be implemented that would automatically remove agents from the network once they terminate their last connection or exceed their dissonance threshold for too long. Additionally, future research should investigate whether or not the results presented here will differ in case the belief updating procedure utilises weighted connections, suggesting that some beliefs contribute more to the updated belief than others. This could for example be based on the length for which a connection already exists, reflecting the assumption that longer relationships result in greater trust.

Finally, since the model is capable to simulate both, large scale networks with more than 10000 nodes and smaller networks, it should be investigated how the conclusions drawn here generalise to different network sizes. Small scale simulations could also be used to investigate interaction in physical social networks [Stu19].

5.2 Relevance

The multi-agent model presented here offers a powerful tool that can compliment existing, more traditional, investigation techniques of social networks. Specifically, the model allows for a level of precision and control about the access to information that cannot be achieved in experiments relying on human subjects. Due to the reliance of real data from human subjects the model ensures, to the degree possible, that the results are more likely to still generalise to real social networks. Moreover, the model here attempts to approximate the process underlying dissonance, thereby offering an initial formalisation of a process that so far has been subject to exclusively theoretical considerations [Mcg17]. In conclusion, this model exemplifies the contributions multi-agent modelling in general has to offer for phenomena that so far have been predominantly investigated using more traditional approaches.

The results obtained from the second simulation study are of particular relevance not only for researchers but also for policy makers involved in the regularisation of social network platforms. Our model can generate important insights into the impacts associated with different recommendation strategies and could be used to inform guidelines for designs of recommendation strategies in social networks. Our results revealed the extreme potential for abuse offered by different recommendation algorithms but also indicates that, if used appropriately, the algorithms can actively contribute to depolarisation and facilitate a broader belief state in the entire network. The ethical implications that arise from the use of such recommendation strategies were also discussed in a recent debate during which officials of the Dutch government highlighted the need for an investigation into the implication of different recommendation strategies [Oll20]. Such developments further highlight the relevance of this study since the model presented here allows to investigate the impact of the different recommendation strategies without risking harm to human subjects.

5.3 Team Work

Max spearheaded the program development, implemented most of the functionality of the programme, and wrote the extensive manual and ReadMe document, while Joshua both researched and embedded experimental findings in the programme to warrant and increase the ecological validity of our model. Joshua and Tom were responsible for research on the theoretical background of echo chambers and Joshua generated experimental findings and discussed them in this report.

References

[20] *Kernel density estimation in Scikit-learn.* Software implementation. 2020. `https://github.com/scikit-learn/scikit-learn/blob/ef5cb84a/sklearn/neighbors/kde.py#L177`.

[And07] J. Anderson. *How Can the Human Mind Occur in the Physical Universe?* Jan. 2007. `https://doi.org/10.1093/acprof:oso/9780195324259.001.0001`.

[AWA10] E. Aronson, T. D. Wilson, and R. M. Akert. *Social psychology.* Upper Saddle River, NJ: Prentice Hall, 2010. ISBN: 0138144788 9780138144784.

[Bre00] Marilynn Brewer. *Handbook of research methods in social and personality psychology. Research design and issues of validity.* Cambridge University Press, 2000. ISBN: 978-0521551281.

[Cen+17] S. Centellegher, E. López, J. Saramäki, and B. Lepri. "Personality traits and ego-network dynamics". In: *PLOS ONE* 12.3 (Mar. 2017), pages 1–17. `https://doi.org/10.1371/journal.pone.0173110`.

[CG04] R. Cialdini and N. Goldstein. "Social Influence: Compliance and Conformity". In: *Annual review of psychology* 55 (Feb. 2004), pages 591–621. `https://doi.org/10.1146/annurev.psych.55.090902.142015`.

[Fes57] L. Festinger. *A theory of cognitive dissonance.* Peterson and Company, 1957.

[FT16] S.T. Fiske and S.E. Taylor. *Social Cognition: From brains to culture.* SAGE Publications, 2016. ISBN: 9781473988262. `https://books.google.nl/books?id=7qPUDAAAQBAJ`.

[Gar+18] H. Garretsen, J. Stoker, R. L. Martin D. Soudisand Martin, and P. J. Rentfrow. "Brexit and the relevance of regional personality traits: More psychological Openness could have swung the regional vote". In: *Cambridge Journal of Regions Economy and Society* 11 (Feb. 2018). `https://doi.org/10.1093/cjres/rsx031`.

[GH14] A. Gohary and H. K. Hanzaee. “Personality Traits as Predictors of Shopping Motivations and Behaviors: A Canonical Correlation Analysis”. In: *Arab Economic and Business Journal* 10 (Oct. 2014), pages 166–174. https://doi.org/10.1016/j.aebj.2014.10.001.

[HH08] E. Harmon-Jones and C. Harmon-Jones. “Action-Based Model of Dissonance: A Review of Behavioral, Anterior Cingulate, and Prefrontal Cortical Mechanisms”. In: *Social and Personality Psychology Compass* 2.3 (2008), pages 1518–1538. https://doi.org/10.1111/j.1751-9004.2008.00110.x.

[HK68] J.A. Hardyck and M. Kardush. “A modest modish model for dissonance reduction.” In: Rand-McNally, 1968, pages 684–692.

[HOF07] H. Hoekstra, J. Ormel, and F. Fruyt. *Handleiding NEO-PI-R en NEO-FFI persoonlijkheidsvragenlijsten.* Dutch. Other. 2007.

[JC09] K. H. Jamieson and J. N. Capella. *Echo chamber: Rush Limbaugh and the conservative media establishment.* Oxford University Press, 2009. ISBN: 978-0195366822.

[Jer20] B. F. Jeronimus. *Real data informing multi agent simulation.* Personal Correspondence with J. Krause. 2020.

[Ke+20] L. Ke, L. Haiming, K. Gang, and D. Yucheng. “Opinion dynamics model based on the cognitive dissonance: An agent-based simulation”. In: *Information Fusion* 56 (July 2020), pages 1–14. https://doi.org/https://doi.org/10.1016/j.inffus.2019.09.006.

[Kri+15] L. Krieke et al. “HowNutsAreTheDutch (HoeGekIsNL): A crowdsourcing study of mental symptoms and strengths”. In: *International Journal of Methods in Psychiatric Research* 25 (Sept. 2015), pages 123–44. https://doi.org/10.1002/mpr.1495.

[KVT20] J. Krause, M. V. Valk, and T. Theuer. *Impact of Strategies on EC formation.* video recordings of simulation. 2020. https://1drv.ms/u/s!AreayAheoFblwfFna0ws4iT7H2indg?e=pyg5h6.

[MBP18] J. K. Madsen, R. M. Bailey, and T. D. Pilditch. “Large networks of rational agents form persistent echo chambers”. In: *Scientific Reports* 8.1 (Aug. 2018), page 12391. ISSN: 2045-2322. https://doi.org/10.1038/s41598-018-25558-7.

[MC87] R. McCrae and P. Costa. “Validation of the five factor model of personality across instruments and observers”. In: *Journal of personality and social psychology* 52 (Feb. 1987), pages 81–90. https://doi.org/10.1037/0022-3514.52.1.81.

[Mcg17] A. Mcgrath. “Dealing with dissonance: A review of cognitive dissonance reduction”. In: *Social and Personality Psychology Compass* 11.12 (2017), e12362. https://doi.org/10.1111/spc3.12362.

[Nic98] R. Nickerson. “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises”. In: *Review of General Psychology* 2 (June 1998), pages 175–220. https://doi.org/10.1037/1089-2680.2.2.175.

[Oll20] K. H. Ollongren. *Desinformatie en digitale inmenging.* 2020. https://www.tweedekamer.nl/debat_en_vergadering/commissievergaderingen/details?id=2020A02405.

[Oz15] H. Oz. “Personality Traits and ideal L2 self as predictors of academic achievement among prospective English teachers”. In: *ICERI2015 Proceedings.* 8th International Conference of Education, Research and Innovation. Seville, Spain: IATED, 18-20 November, 2015 2015, pages 5833–5841. ISBN: 978-84-608-2657-6.

[SGB95] L. Simon, J. L Greenberg, and J. Brehm. “Trivialization: the forgotten mode of dissonance reduction.” In: *Journal of Personality and Social Psychology* 68 (1995), pages 247–260. https://doi.org/10.1177/0146167208325613.

[Smi14] A. Smith. Survey report. 2014. `https://www.pewresearch.org/fact-tank/2014/02/03/what-people-like-dislike-about-facebook/`.

[Stu19] G. Stulp. *Connections in physical social networks*. Personal Correspondence with J. Krause. 2019.

[VTK20] M. V. Valk, T. Theuer, and J. Krause. *misinformation-spread-using-cogdis*. Software implementation. 2020. `https://github.com/MaxVinValk/b-14-misinformation-spread-using-cogdis`.