Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks

Björn Ross, Laura Pilz, Benjamin Cabrera, Florian Brachten, German Neubaum & Stefan Stieglitz

THE OPERATIONAL RESEARCH SOCIETY

Taylor & Francis
Taylor & Francis Group

Check for updates

EMPIRICAL RESEARCH

Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks

Björn Ross [ID], Laura Pilz, Benjamin Cabrera, Florian Brachten, German Neubaum and Stefan Stieglitz [ID]

Department of Computer Science and Applied Cognitive Science, University of Duisburg-Essen, Duisburg, Germany

ABSTRACT

Information systems such as social media strongly influence public opinion formation. Additionally, communication on the internet is shaped by individuals and organisations with various aims. This environment has given rise to phenomena such as manipulated content, fake news, and social bots. To examine the influence of manipulated opinions, we draw on the spiral of silence theory and complex adaptive systems. We translate empirical evidence of individual behaviour into an agent-based model and show that the model results in the emergence of a consensus on the collective level. In contrast to most previous approaches, this model explicitly represents interactions as a network. The most central actor in the network determines the final consensus 60–70% of the time. We then use the model to examine the influence of manipulative actors such as social bots on public opinion formation. The results indicate that, in a highly polarised setting, depending on their network position and the overall network density, bot participation by as little as 2–4% of a communication network can be sufficient to tip over the opinion climate in two out of three cases. These findings demonstrate a mechanism by which bots could shape the norms adopted by social media users.

1. Introduction

Information systems today enable people to interact publicly, ubiquitously and in a self-organised manner on the internet. Social media has played a major role in driving this development (Aral, Dellarocas, & Godes, 2013). Many online interactions are shaped by social influence exerted by peers in social networks (Matook, Brown, & Rolf, 2015). Accordingly, social media have become important in the public's acquisition of information and the formation of consumer opinion (Dewan, Ho, & Ramaprasad, 2017; Fan & Lederman, 2018). However, it is often difficult to determine whether a piece of information or opinion on social media is authentic. This leads to growing worries that malicious actors might spread misinformation online to manipulate the public (Berners-Lee, 2017). Current research suggests that large numbers of automated actors called "social bots" are attempting to influence public opinion (Ferrara, Varol, Davis, Menczer, & Flammini, 2016). Their use has been especially researched in the field of politics, where a successful attempt to influence public opinion will even influence the outcome of votes. Research on the use of social bots during the 2016 U.S. presidential election showed that while bots were used in support of both candidates, many more bots supported Donald Trump than Hilary Clinton (Bessi & Ferrara, 2016). Furthermore, the content produced

by these bots had a much more positive sentiment than the overall debate. The authors concluded that "the fact that bots produce systematically more positive content in support of a candidate can bias the perception of the individuals exposed to it, suggesting that there exists an organic, grassroots support for a given candidate, while in reality it's all artificially generated." It is thought that social bots may have played a role in the outcome of the "Brexit" referendum in 2016 on membership of the United Kingdom in the European Union (Schäfer, Evert, & Heinrich, 2017), and some of the social bots used in the U.S. presidential election were also active before the 2017 presidential election in France (Ferrara, 2017). In the latter case, they participated in an attempt to spread the so-called "MacronLeaks", a dataset containing possibly problematic information on the candidate Emmanuel Macron. While they were successful in the dissemination of the dataset and the story was picked up by major news outlets, the candidate targeted by the social bot accounts did win the election in the end. These cases show that social bots have been used in attempts to influence elections; yet, sufficient findings on why and under what circumstances they can be successful is missing. Consequently, there is a pressing need for research to analyse the extent to which automated actors such as social bots can influence the formation of opinion climates online.
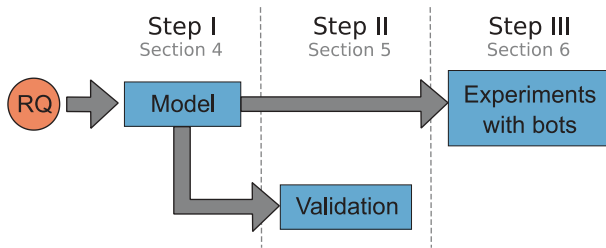
CONTACT Stefan Stieglitz ✉ stefan.stieglitz@uni-due.de

Figure 1. Illustration of the research steps.

According to the spiral of silence theory (Noelle-Neumann, 1974), opinion climates are subject to social influence processes in which a person's public opinion expression behaviour is influenced by his or her environment. If an individual determines that their private opinion differs from that of the perceived majority, the individual becomes less likely to express their opinion publicly. Over time, this mechanism can lead to a "spiral of silence," by which an opinion expressed by a vocal minority may ultimately become a social norm in public life, even if a considerable group of people internally disagree with it. In view of the proliferation of various kinds of social bots and other actors that promote certain opinions in several different social media networks, it is essential to examine the following research question.

RQ: How likely are social bots to influence public opinion formation by triggering a spiral of silence, depending on the characteristics of the bots and the structure of the social network?

Although this theory has been investigated in the context of social media communication (e.g., Hampton et al., 2014), most studies have focused on individual behaviour without considering how individual actions can together lead to social dynamics over time. While cross-sectional surveys and small lab experiments cannot address the dynamic process of the spiral of silence (Matthes, 2015), large-scale field experiments are associated with ethical questions (Agarwal & Dhar, 2014) and high efforts and costs.

In contrast to these previous approaches, we use virtual simulations to study how this social phenomenon emerges over time. To this end, we draw from complex adaptive systems (CAS) theory (Nan, 2011) to develop an agent-based model (ABM) of the spiral of silence. Most previous simulations of this process modelled human interactions in a two-dimensional physical environment (e.g., Sohn & Geidner, 2016). Our approach, however, is to model interactions as a network, which is more suitable to represent human communication behaviour (Wilensky & Rand, 2015). Therefore, the three steps of this research are to (cf. Figure 1):

(a) develop an agent-based network model simulating the spiral of silence from sound theoretical assumptions and empirical findings;

(b) validate the model based on comparisons with previous findings from social media and psychological research; and

(c) apply the model to examine the potential influence of bots.

In particular, we examine how bots could be used to influence the spiral of silence online. Social networks vary in several aspects, especially in the number of connections between users, and the position of an actor in a network determines how influential they are. We therefore examine how the influence of bots depends on the density of the network and their position in it as well as their number. Since bots exhibit varying degrees of sophistication, from simple spam accounts to the intelligent autonomous actors that artificial intelligence research aims to develop, we also take a bot influence factor into account that determines how likely bots are to have an impact of the opinion of a human user. These concepts and the relationships studied are depicted in Figure 2. This research offers first insights into the actual danger posed by social bots in various environments and their interfering effects on the public sphere.

In this paper, we first explain the necessary theoretical background information and summarise related work. The three subsequent sections are each dedicated to addressing one of the above research steps. After a general discussion of implications and limitations of the study in the penultimate section, the final section draws conclusions and outlines possible further research.

2. Theoretical background

2.1. The spiral of silence theory

The spiral of silence theory explains how public opinion is formed (Noelle-Neumann, 1974). Individuals fear isolation from society, which leads them to continuously observe the distribution of opinions in their social environment. Using a "quasi-statistical organ", they estimate the prevalence of agreement and disagreement on contentious issues. If they perceive themselves to be in the minority, they will lose confidence in their own opinion and be more likely to remain silent. Their decision to no longer express their opinion, in turn, influences others in their environment to underestimate the frequency of this opinion. This starts a self-reinforcing process of people being silenced and others further misjudging the opinion climate. Eventually, an opinion is accepted as prevailing even when a large part of the population may silently disagree with it. The spiral of silence theory is therefore used to explain the state of "pluralistic ignorance," where a silent majority wrongly perceives itself to be outnumbered by a much more
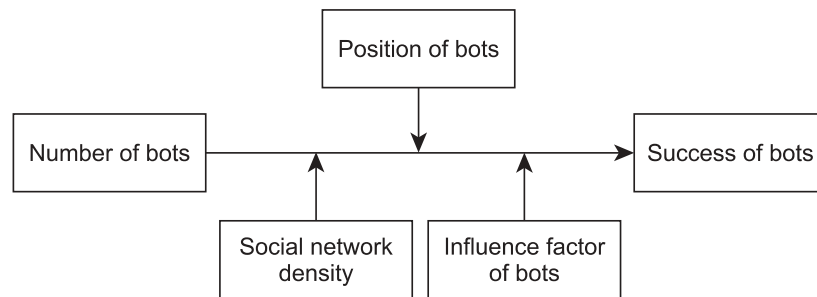
Figure 2. Research model for the third step. The bots are considered successful when their opinion becomes accepted as the majority opinion, while holders of the opposite opinion are silenced.

vocal minority (Noelle-Neumann, 1977, 1993). The mass media and in particular their agenda-setting function play a crucial role in the spiral of silence theory, because individuals rely on them to gauge the climate of opinion (Noelle-Neumann, 1974). However, even in situations in which the public's perception of the opinion climate is generally accurate, the minority opinion is not always completely silenced. In some cases "a minority faction may be reduced to a hard core of persons who are not prepared to conform" (Noelle-Neumann, 1974, p. 48).

The spiral of silence theory relies on observations about individuals made in social psychology but makes predictions about group behaviour relevant to many other disciplines, including information systems. The influence exerted by peers through social networks can play a major role in shaping employees' responses to organisational policies (Bhattacherjee, Davis, Connolly, & Hikmet, 2017), and in forming behavioural intentions such as technology adoption decisions (Sykes, Venkatesh, & Gosain, 2009). Venkatesh and Morris (2000) consider social influence identical with *subjective norm*, one of the most important antecedents of technology acceptance. It is noteworthy that the spiral of silence theory explains normative social influence, meaning that individuals adapt their public behaviour to their environment (e.g., by not expressing their opinion), while their private opinions remain unaffected by the environment.

Empirically, meta-analyses have corroborated the notion of changes in one's confidence based on the perceived opinion climate. Perceiving public support for one's opinion increases people's willingness to express their opinions (Glynn, Hayes, & Shanahan, 1997; Matthes, Knoll, & Von Sikorski, 2018), although there are individual differences: It has been shown that those who feel very certain about their opinion (the "hard core") are less inclined to adapt their opinion expression behaviour to the prevailing opinion climate than those who feel less certain about their opinion being correct (Matthes, Morrison, & Schemer, 2010). Likewise, it has been argued that people have a dispositional tendency or willingness

to self-censor when confronted with opposing majorities, and that this individual disposition influences the extent to which their willingness to voice personal stances in public depends on the opinion climate (Hayes, Uldall, & Glynn, 2010).

Despite the extensive amount of empirical spiral of silence research, studies that address the reciprocal effects of individual opinion expression behaviour and the formation of opinion climates remain the exception (Matthes, 2015). Scholars have repeatedly called for research that grasps the dynamic nature of the theory and analyses how opinion climates form over time (Matthes & Hayes, 2014; Scheufele, 2007).

2.2. The spiral of silence online

In the age of social media, scholars have begun to examine whether the assumptions of the spiral of silence still hold. Especially with the emergence of social bots, the spiral of silence theory has gained renewed interest. If social bots advocating a certain opinion spread over a network, this could lead to the false impression that the "bot opinion" is shared by more humans that it really is. Consequently, people who agree with this opinion gain the confidence to speak about it publicly, while those who disagree keep silent out of fear of being socially isolated.

Information systems research shows that social influence plays a key role in many online interactions, such as product search (Grange & Benbasat, 2018) or the success of YouTube videos (Susarla, Oh, & Tan, 2012) and crowdfunding decisions (Thies, Wessel, & Benlian, 2016), although it is far from the only factor (Fang, Hu, Li, & Tsai, 2013). There is also evidence that opinions expressed online reflect the opinions of the general population, e.g., in the case of doctor ratings (Gao, Greenwood, Agarwal, & Mccullough, 2015). A representative survey of American adults conducted by the Pew Research Center found that those willing to discuss surveillance programmes in person were unlikely to post their views on social media (Hampton et al., 2014). In addition, only 15% of internet users said they received information about

the Snowden leaks from Facebook. However, the study confirmed that Facebook is used to assess the opinion climate, and influences willingness to speak out both online and offline. Experimental research corroborated that individual opinions expressed in user-generated comments on Facebook can (weakly) shape the way people perceive the opinion climate among the national population (Neubaum & Krämer, 2017). Although people seem to use social media to infer opinion climates, they are not generally willing to voice political opinions in these public environments (Neubaum & Krämer, 2018).

Taken together, these results suggest that when it comes to assessing the opinion distribution, social media play a vital role. This in turn leads to a climate that is especially prone to distortion as social media enable virtually anyone with access to the internet to express their opinion online. These findings become especially relevant when one considers that traffic on online social networks is not only produced by human actors but also by automated accounts programmed to exhibit a certain behaviour and execute different tasks.

2.3. The influence of social bots on the opinion climate online

Automated accounts are often referred to as bot accounts or bots and show diverse patterns of behaviours and different levels of automation and sophistication. Among social media technologies, Twitter is especially popular both for the use of bots and for research on these bots through its open and easily accessible API. Studies found that between 9% and 15% of Twitter accounts are actually bots (Chu, Gianvecchio, Wang, & Jajodia, 2012; Varol, Ferrara, Davis, Menczer, & Flammini, 2017). The behaviour of bots ranges from automatically reposting a human user's content on a different platform, or automatically posting useful information such as the weather forecast, to simulating human behaviour and apparently taking part in communication with human users (Stieglitz, Brachten, Ross, & Jung, 2017). We define *social bots* as bots exhibiting the latter behaviour (Abokhodair, Yoo, & Mcdonald, 2015; Hegelich & Janetzko, 2016), that is automated accounts on social media sites which act similar to humans, post content and are not necessarily recognisable as automated to human users.

Among these social bots are those that convey a particular opinion through their messages and are thus thought to try to influence human users' opinions. Numerous studies have tried to capture this phenomenon and to show the application of social bots in several scenarios, often in political contexts (e.g., Abokhodair et al., 2015; Forelle, Howard, Monroy-Hernandez, & Savage, 2015).

The potential of social bots to influence human users can be deduced from the Media Equation Theory by Reeves and Nass (1996). In their studies, participants responded to computers as if they were other human beings, even though the participants knew they were interacting with inanimate objects. The authors noted that "people were polite to computers. Not only were the computers in these experiments tools for learning new information, they were social actors that people reacted to with the same polite treatment that they would give to another human" (p. 26). In a more recent study, Edwards, Edwards, Spence, and Shelton (2014) transferred these findings to bots on Twitter and showed that, "although [the] scores were generally lower than the condition with a perceived human Twitter feed, the results demonstrate that a Twitterbot is perceived as a credible source of information" (p. 374). Furthermore, it has been shown on an individual level that humans not only respond to artificial dialogue partners like they would to other humans, but that bots actually have the potential to alter the human's behaviour (Munger, 2017). The author showed that automated Twitter messages in response to people posting offensive messages led to a decrease in these messages in the following weeks.

In summary, studies show that, on many occasions, humans interact with computers in general and with bot accounts in particular similarly to how they would interact with other humans—despite being aware of the inanimate character of these bots. This does not mean that they see them as equals, and true social recognition is an arguably much more complex issue (Rowe, 2018). Still, concerning the potential of automated communication to alter a person's behaviour, the above findings suggest that bot accounts can have an influence on a user's actions. All these results show that, on an individual level, there may be the potential for bot accounts to interfere with users. However, it remains unclear whether these findings can carry over to a larger scale and shape a larger part of a conversation on a certain topic online. To see how social bots could alter discussions on topics on a larger scale, different mechanisms need to be examined.

One mechanism by which this influence might take place is astroturfing. The potential here lies in the sheer number of messages that bots are able to post without the need of human action. As Zhang, Zhang, Zhang, and Yan (2013) write, "astroturfers strive to create the falsified impression that the given ideas or opinions are held by a large portion of the population" (p. 2). As there is virtually no limit to the number of social bots that can be deployed, the conversation on a topic can be flooded with automatically generated messages of a certain opinion, thus creating the impression that many people hold that opinion.

The nature and impact of astroturfing on online conversation, as laid out above, clearly bear resemblance to the idea of the spiral of silence. For example, hundreds of bots sharing the same messages or keywords on Twitter could create the false impression that support for this opinion is high among users. Bots could equally spam comments on Facebook discussions with a certain opinion so that people of the opposite opinion feel that they are in the minority, and thus refrain from commenting. In these scenarios, people could act on a distorted perception of the opinion climate, possibly leading to uninhibited or damaging behaviour.

Still, to the best of our knowledge, the spiral of silence theory has not yet been applied in the context of social bots and astroturfing. While there is sufficient evidence for the existence of bots (e.g., Ferrara et al., 2016), their actual impact remains unknown (cf. Cossu, Labatut, & Dugué, 2016). Due to the difficulty of reliably measuring the actual impact of online actors on the opinion climate, the research is largely based on the unproven assumption that social bots can have an impact. This impact is often measured empirically by examining whether humans retweeted bot tweets (e.g., Hegelich & Janetzko, 2016), but the mere fact that a human has retweeted a bot does not prove that the human has been influenced. We approach the problem from the opposite angle, starting with a theory of opinion formation and then reasoning about possible mechanisms by which bots could influence it. This scenario has a vast potential in helping to understand mechanisms of opinion formation and manipulation online. It also has clear implications for events occurring offline, such as elections. Therefore, the need for a valid theoretical model to assess the impact becomes apparent.

3. Related work

3.1. Simulating the spiral of silence

As has been pointed out, the spiral of silence is a widely accepted theory in communication studies on the causes of normative social influence. Previous research has been primarily based on cross-sectional survey data (e.g., Matthes et al., 2010) and relatively small-scale experiments (e.g., Hayes et al., 2010). Both the assumptions and the predictions of the spiral of silence theory have been backed up by this evidence. However, the exact conditions under which it takes place and the mechanisms by which it emerges in large-scale settings are unclear. This lack of research may be due to the fact that such studies have traditionally been difficult to design. Laboratory experiments on conformity have been criticised for their low external validity (Scheufele & Moy, 2000). Large-

scale observational studies on social media would perhaps be an option, but they are fraught with methodological challenges such as bias in the data (Ruths & Pfeffer, 2014; Stieglitz, Mirbabaie, Ross, & Neuberger, 2018). A solution is to conduct a virtual experiment through simulation.

Prominent early work on simulating related opinion formation processes was carried out by Nowak, Szamrej, and Latané (1990). In their model, social interactions occur based on Euclidean distances in a two-dimensional environment. Yet the authors called for communication networks to be considered explicitly in the future:

> Nearly any type of relationship ... may be conceived of as a network.... . This kind of representation would allow one to shape the group structure at will. For example, one could represent an actual network of social contacts as obtained from survey data. Alternatively, one could model an organizational structure including both formal and informal channels of communication.... . A particularly interesting possibility would be to model computer-mediated communication networks of the sort that are increasingly common in business and academia. (Nowak et al., 1990, p. 373)

The call for more network-based research on opinion formation theories has been echoed but rarely answered (Song & Boomgaarden, 2017). In the context of the spiral of silence, researchers have used network models to examine the influence of long-range interactions (Jiang, Hua, Zhu, Wang, & Zhou, 2008), the influence of the chosen network topology (Wu, Du, Li, & Chen, 2015) or the influence of neutrals and hardcore nonconformists (Takeuchi, Tanaka, Fujie, & Suzuki, 2015) on opinion dynamics. However, to our knowledge, there have not been any simulation studies on the deliberate manipulation of public opinion in the context of the spiral of silence.

3.2. The complex adaptive systems perspective and agent-based models

We propose to reason about the spiral of silence theory within the analytical framework of *complex adaptive systems* (CAS). CAS are "systems composed of interacting agents described in terms of rules. The agents adapt by changing their rules as experience accumulates" (Holland, 1995, p. 10). They are characterised by the *emergence* of complex structures and patterns that arise as a result of the interplay of many smaller entities (Wilensky & Rand, 2015). *Agent-based modelling* is a tool to design and examine virtual complex adaptive systems. Because these systems involve a large number of individuals, it is often infeasible to closely monitor the emergence of the phenomenon of interest in detail in the real world, and mathematical models are not analytically

tractable. An ABM can therefore help understand the complex system and can also be valuable in studying its long-term behaviour and in predicting its future behaviour (Wilensky & Rand, 2015).

CAS theory has been applied to a wide range of phenomena relating to information systems, either on its own on a purely conceptual level (Nan & Lu, 2014) or in conjunction with agent-based modelling (ABM). It has been used to examine concepts such as IT use (Nan, 2011), agile software development (Vidgen & Wang, 2009), and cohesion in virtual teams (Curșeu, 2006), and several others (Canessa & Riolo, 2006; Havakhor, Soror, & Sabherwal, 2018; Johnson, Faraj, & Kudaravalli, 2014). This wealth of literature illustrates that presently, CAS is a well-established analytical framework in Information Systems. The process of translating a theory into a conceptual CAS model is considered a theory-building exercise (Nan, 2011).

Complex adaptive systems are composed of an environment, agents, and interactions (Nan, 2011). When modelling the system, the problem at hand often immediately determines the nature of the agents. The choice of the environment is closely linked to the topology of interactions within the model. The two most frequently used topologies in ABM are spatial and network-based environments (Wilensky & Rand, 2015). *Spatial environments* are used when interactions between agents are the result of physical proximity. In this case agents are placed in a two-dimensional space. A good example of a spatial model is an ant colony gathering food from different sources (Wilensky & Rand, 2015). Ants only transmit information when close to each other. In one unit of time, they can only move to a neighbouring field. *Networks* composed of nodes and links are the other common topology used in ABM. The nodes are not assigned a physical position. Instead, the neighbourhood of a node is determined by its links to other nodes. This allows for high-dimensional topologies especially useful in modelling social interactions, in particular human communication. Since communities on social media lend themselves easily to network representations (Kane, Labianca, & Sp, 2014), the choice of a network-based environment is common for the application of ABM to this case (e.g., Canessa & Riolo, 2006; Havakhor et al., 2018; Hildebrand, Hofstetter, & Herrmann, 2012; Johnson et al., 2014; Ren & Kraut, 2007).

The strength of ABM lies in the fact that the researcher only needs to specify an environment, agents, and the rules by which agents act and interact in the model. In contrast, "we do not have to define in advance the emergent, global properties; instead we can observe these properties as they arise from a simulation of multiple distributed interacting agents" (Wilensky & Rand, 2015, p. xvii). This property makes ABM a perfect fit for the spiral of silence theory. As mentioned before, assumptions at the individual or local level have been subject to a great deal of scrutiny by psychologists (Glynn et al., 1997; Glynn & Huge, 2014). Our goal is to observe the emergence of the global phenomenon. Here, ABM may bridge this gap between individual and collective behaviour.

Yet, the application of agent-based modelling to the spiral of silence theory is a relatively new field. Sohn and Geidner (2016) examined how the number of communication partners influences the process. When agents communicated with more agents on average, a global spiral of silence was more likely to form and the population was less likely to become polarised. According to the authors, today's social media represent such a densely networked communication environment. However, Sohn and Geidner did not model communication as a network explicitly but used a spatial environment. Actors move on a fixed-size two-dimensional surface. As in Nowak et al. (1990)'s simulations mentioned above, interactions occur based on the Euclidean distance between agents on this surface. Individual agents observe the opinions of other agents and continually adjust their confidence in their own opinion based on the number of supporting and opposing opinions expressed in the vicinity. Likewise, Gawroński, Nawojczyk, and Kułakowski (2015) modelled the spiral of silence in an agent-based environment but assumed that agents perceive the opinions expressed by all other agents. This assumption is justifiable when the goal is to model dynamics in small groups, but not online on global social media.

4. Model

This section addresses the first research step, developing the model. A key strength of agent-based models is that they force the researcher to make assumptions about individual behaviour and the environment explicit, and much more so than in textual descriptions of theory (Wilensky & Rand, 2015). Consequently, when developing agent-based models, it is essential to explicitly justify the choices made. Therefore, in each modelling step, we first describe the assumptions made, which are grounded in previous research. We used NetLogo (Wilensky, 1999) as a simulation tool, a programming environment that offers an API which makes the simulation of agent-based models straightforward, along with the package RNetLogo (Thiele, 2014) for the statistical computing environment R.

4.1. Agent-based modelling world

For the reasons laid out above, we consider a network topology the most appropriate environment to model the spiral of silence. In order to reflect previous findings on the topology of social interactions, this network model should satisfy the following requirements and assumptions.

(1) Real-world social networks often exhibit *power-law tailed degree distributions* (Barabási & Albert, 1999) and are *small-world* networks (Watts & Strogatz, 1998). Since we intend to model a social network, these properties are desirable.

(2) The network remains static during the opinion formation process, because opinion formation is assumed to take place faster than changes in the underlying real-world social networks. The network is therefore generated once for each simulation run. While this simplification might not fully mirror real-world networks, it is assumed that networks do not change as quickly as an issue can become a widely discussed topic. Recent evidence, for example, has shown that only a few users are willing to make changes to their network as a result of political disagreements (Bode, 2016; John & Dvir-Gvirsman, 2015).

(3) The network model should make it possible to study the influence of different network topologies, especially density and the positions of actors, on the spiral of silence by varying parameters.

The Barabási-Albert model (Barabási & Albert, 1999) generates power-law degree distributions and small-world networks (Dommers, Van der Hofstad, & Hooghiemstra, 2010). The generalised preferential attachment model described by Dorogovtsev, Mendes, and Samukhin (2000) additionally satisfies the third requirement. This version allows for variations in the number of links and the scaling parameter of the degree distribution. As in the original model of preferential attachment, a graph is built by adding one node at a time. Whenever a new node is added, it is connected to m existing nodes chosen with a probability proportional to an attractiveness score. The attractiveness of node i is defined as

$$A_i = m(\gamma - 2) + q_i,$$

where the first term is its initial attractiveness and is the number of edges that have previously been added to it. This means that nodes with many connections, so-called hubs, are more likely to be connected to a new node than nodes with few links. Such a rich-get-richer principle has also been observed empirically in social networks (Newman, 2001).

In the following section we use common notation from graph theory to describe the network. The network is denoted by $G = (V, E)$, where $V = \{1, \ldots, n\}$ is the set of all (enumerated) agents and $E = \{\{i, j\} : i, j \in V\}$ the set of undirected connections between them. Let $N(v)$ denote the set of neighbouring agents of an agent v, that is

$$N(i) = \{j \in V : \{i, j\} \in E\}.$$

The decision to model relationships as undirected instead of directed is based on a few considerations. First, there are real-world examples of both types of edges. Networking sites that extend the classical notion of friendships (e.g. Facebook, LinkedIn) use bidirectional edges while other services (e.g. Twitter, Instagram) have directional follower/followee relationships. However, we argue that even the pages and feeds of large influencers on directional sites such as Twitter contain the opinions of far more people than just the few accounts they follow. That is because these sites usually allow for the followers to interact by leaving comments and voting on posts. The reader of such a page will likely read the comments and be influenced by many people. Likewise, the portrayal of public opinion on news media is often shaped by nationwide polls. In conclusion, influence spreads both ways, and undirected edges are a plausible simplification. Apart from that, undirected networks lead to more generalisable results. There are many more parameters to choose from for directed scale-free networks than for undirected ones, narrowing the context in which a particular model can be applied.

4.2. Agent behaviour

Apart from modelling the environment, we also model agent behaviour, in our case related to the spiral of silence. As far as agent behaviour is concerned, our model is based on the approach described by Sohn and Geidner (2016). In this section, we recap their model and highlight the changes we introduced in our variant of the model.

As mentioned above, the spiral of silence explains changes in the willingness to express one's beliefs (or opinions), not changes in one's actual beliefs. As a result, we model an agent's *opinion* as a static value. Also, for the sake of simplicity, we only consider two possible opinions, *positive* and *negative*. Every agent is therefore equipped with a fixed dichotomous value o_i in $\{+, -\}$ modelling the preference or dislike of a certain topic.

In contrast to the fixed opinion, the question of whether the agents publicly express their beliefs

depends on their surroundings. However, since the willingness to self-censor varies from individual to individual (Hayes, Glynn, & Shanahan, 2005), some people are generally more likely to express their opinions than others regardless of the topic. This tendency of an individual i to speak out is modelled as the *willingness to self-censor* Φ_i (Hayes et al., 2005, 2010), a continuous value in $[0; 1]$. Like the opinion, an agent's willingness to self-censor does not change over time.

In contrast, an agent's *confidence* $c_i(t)$ is influenced by the majority opinion the agents observe in their neighbourhoods (cf. Sohn & Geidner, 2016), as proposed by the spiral of silence theory. The confidence of an agent is a continuous value in the range $[0; 1]$. At each step of the process, individuals will compare their confidences with their willingness to self-censor in order to determine whether or not to speak out. A confidence higher than the disposition to self-censor causes an individual to share his or her opinion while a lower value causes the individual to remain silent. Of course, the opinion and confidence of an individual depend on the topic. However, in our simulation we model the spiral of silence in the context of a single topic only. Consequently, every agent has only one value at a time for each of the variables.

The *opinion climate* $\delta_i(t)$ observed by agent i at time t is the proportional difference between the perceived opinions of the neighbours:

$$\delta_i(t) = \begin{cases} \frac{n_s(i,t) - n_o(i,t)}{n_s(i,t) + n_o(i,t)}, & \text{if } n_s(i,t) + n_o(i,t) > 0 \\ 0, & \text{if } n_s(i,t) + n_o(i,t) = 0 \end{cases}, \text{where}$$

$$n_s(i,\ t) = \left| j \in N(i) : o_j = o_i,\ \Phi_j > c_j(t) \right|,$$
$$n_o(i,\ t) = \left| j \in N(i) : o_j \neq o_i,\ \Phi_j > c_j(t) \right|.$$

Thus, $n_s(t,\ i)$ and $n_o(t,\ i)$ denote the number of individuals in the neighbourhood of i with identical and opposite opinions compared to i's opinion o_i at time t, respectively. $\delta_i(t)$ is a value in $[-1;\ 1]$, where the extreme value -1 is approached if the neighbourhood of agent i only consists of opinions opposing i's opinion for a prolonged period of time and $+1$ is approached when the neighbourhood completely agrees with i. Note that the numbers $n_s(t,\ i)$ and $n_o(t,\ i)$ only take agents into account who express their opinions. Silent agents are not considered.

The variable *opinion climate* captures the idea that "individuals form a picture of the distribution of opinion in their social environment" (Noelle-Neumann, 1974, p. 45). This social environment, in our model, consists of their neighbours in the social network. It is important to note that the perceived opinion climate depends on the relative distribution of opinions among one's neighbours. Actors with many neighbours are influenced by many, but each

of these neighbours only has a small amount of influence on them. Actors with many neighbours also, in turn, influence many other actors. It is therefore advantageous for those who seek to influence public opinion bots, for example to acquire many connections to others.

A crucial question is how to model changes in one's confidence based on the observed opinion climate. As Noelle-Neumann writes of someone observing his [*sic*] surroundings, "he may discover that he agrees with the prevailing (or winning) view, which boosts his self-confidence and enables him to express himself with an untroubled mind and without any danger of isolation." Since this is the only agent variable changing over time, the long-term behaviour of our ABM is closely linked to how confidence is modelled. The confidence function should fulfil the following axioms.

(1) *Boundedness*: The confidence can never be lower than zero and should always be less than one. Because we limit the willingness to self-censor to the interval $[0;\ 1]$ we also have to limit the confidence, which is directly compared to it.

(2) *Strictly increasing*: Although bounded from above by 1, the confidence will always increase if the perceived opinion climate supports an agent's opinion.

(3) *Diminishing marginal effects*: Agents with a low confidence are more easily influenced by the opinion climate than already confident ones. Likewise, it requires more time for the opinion climate to influence those who are more confident (and consequently more outspoken). This pattern is in line with previous spiral of silence survey research which showed that an individual's susceptibility to normative social influence depends on how certain they feel about their opinion (Matthes et al., 2010).

(4) *Symmetry*: If it takes x years of a supportive opinion climate to consolidate an opinion in an actor and make them confident about it, then it should take the same time under an equally opposing opinion climate to erode this confidence again. This is based on the notion that human beings tend to avoid information that contradicts their own opinion, and strive instead for cognitive consonance (Festinger, 1957). Drawing on the ideas of cognitive dissonance and selective exposure (Knobloch-Westerwick, 2014), it is argued that it is more difficult to change an individual's opinion the longer they have held it.

The following composition of two functions satisfies the above axioms. Every agent has an internal

value $\hat{c}_i(t)$ hat additively grows or decreases based on the observed opinion climate in the neighbourhood of i. However, because there is no such thing as negative confidence (axiom 1), we bound the value from below by zero:

$$\hat{c}(t) = \max\{\hat{c}(t-1) + \delta(t); 0\}$$

Note that $\hat{c}(t) \in [0; \infty[$. The values can be transformed to the interval $[0; 1]$ by applying a sigmoid function (cf. Sohn & Geidner, 2016) to obtain the *confidence*:

$$c(t) = 2 * \left(1 + e^{-\hat{c}(t)}\right)^{-1}$$

The confidence is used in the comparison with the willingness to self-censor. To initialise the model, $\hat{c}_i$ and the willingness to self-censor of each agent i are drawn from a uniform distribution, i.e. $\hat{c}_i(0) \sim U(0, 1)$, $\Phi_i \sim U(0, 1)$. This assumption is later relaxed as part of a sensitivity analysis.

Figure 3 shows an example of how confidence updates work in practice for an individual node. Consider the individual at the centre. At time t, there is $\hat{c}(t) = 1.5$. Since three of the agent's four neighbours are of the opposite opinion, the agent's internal value $\hat{c}$ at time decreases to $\hat{c}_i(t+1) = 1.5 + \frac{1-3}{1+3} = 1$. This change causes the confidence to drop below the agent's willingness to self-censor, preventing the agent from further expressing their opinion in public.

Confidence updates are executed repeatedly until the model reaches a stable state, that is none of the agents changed their confidence by more than between two units of time (or ticks). Figure 4 shows an example network (larger than the one in Figure 3) and its subsequent states after various numbers of ticks.

4.3. Modelling social bots

To assess the influence of automated agents such as social bots in the process, bots are also modelled as agents. In contrast to the "human" agents, they always assert their programmed "opinion" regardless of the perceived opinion climate. As a result of the theoretical considerations laid out earlier, we assume, for the time being, that bots influence the perceived opinion climate the same way humans do. This assumption is relaxed later as part of a sensitivity analysis.

Bots are only added once all human users have been added. The human actors are inserted using generalised preferential attachment, as described above. Every newly added bot is linked to m existing nodes, just like the other agents. Since bots are added later, their average degree will be lower than that of the normal agents, as they have fewer chances to be connected to new nodes. This phenomenon is consistent with previous research. Ferrara et al. (2016) assume that bots have fewer
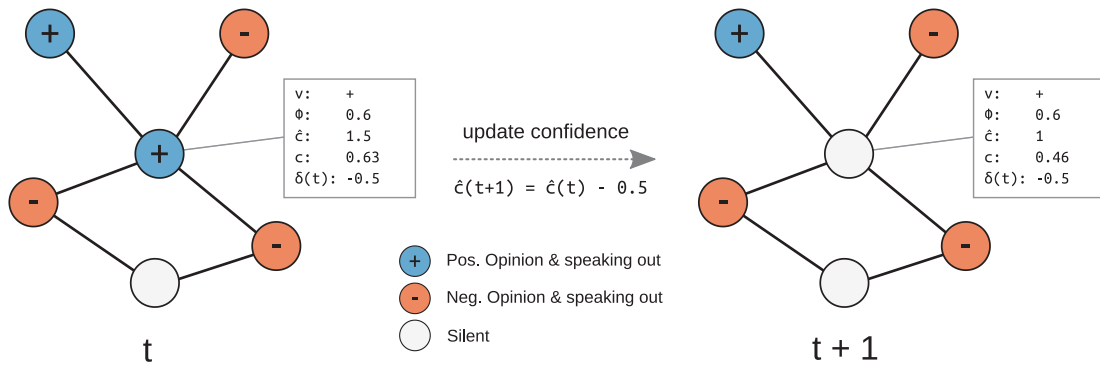

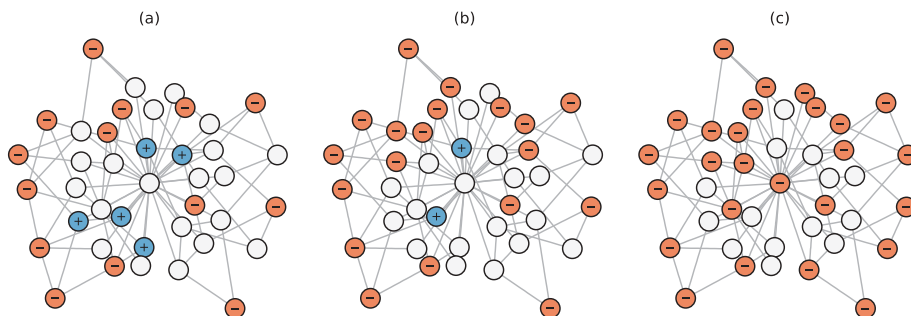**Figure 3.** Example confidence update.

Figure 4. Example of the spiral of silence process (a) at the beginning, (b) after 10 ticks, and (c) at the stable state ($m = 3$, $\gamma = 2.5$). In this case, the minority is silenced completely and a social norm is definitely established.

connections to real users than real users have among themselves.

The position of bots in the network may strongly affect their influence on opinion formation as, like Aiello, Deplano, Schifanella, and Ruffo (2012) report, their influence depends in part on their position in the social graph. To further analyse the effect of the position of bots in the communication network, we consider three mechanisms for creating links to existing nodes:

- *(Generalised) Preferential attachment*: The probability that a bot is connected to an existing node is proportional to the attractiveness of that node. Connecting bots to hubs could be a fruitful strategy for bot creators because successfully influencing them might have a large cascading effect on the rest of the network.
- *Inverse preferential attachment*: The probability that a bot is connected to an existing node is inversely proportional to that node's attractiveness. Connecting to poorly connected agents might also be a viable strategy for bots as peripheral users estimate the opinion climate based on a small number of users. They are therefore more easily influenced by a single bot.
- *Random attachment*: Bots are linked to random agents. This is a simple strategy, and it is useful in the evaluation for comparison purposes.

Experiments were run for each of the three attachment models and results were compared to identify potential differences in bot strategies.

5. Validation

The second research step is to ensure that the modelled process is indeed the spiral of silence. Beese, Haki, and Aier (2015) stress the need for a "thorough and well-documented validation procedure" in simulation-based Information Systems research. Several validation techniques were therefore used together to ensure that the model is accurate (Sargent, 2005).

5.1. Correctness of implementation

During model development, randomly chosen actors were observed over time, and their internal state and behaviour were traced and compared to calculations by hand to ensure that the desired mechanisms were implemented correctly (cf. Figure 3). Animations of the running model were inspected visually to ensure that the model behaves as intended (cf. Figure 4). Parameters were varied with a focus on extreme values to ensure that the model behaves as expected in extreme conditions, such as when all actors are of

the same opinion. To ensure internal validity, the amount of stochastic variability was measured and is reported in the following sections and figures where appropriate. These tests ensure that the NetLogo code is a correct implementation of the conceptual model outlined above.

5.2. Validity of the model

To ensure that this model accurately reflects the theoretical concept of the spiral of silence, operational graphics were created, some of which are reproduced in this article (Figures 5–8). These results were compared to past empirical findings (Noelle-Neumann, 1974) and to the results of other simulation models (Nowak et al., 1990; Sohn & Geidner, 2016). In particular, in the following, we consider three distinct observations that are central in the spiral of silence theory, as discussed earlier. Each of these observations is first presented, then discussed.

5.2.1. The effect of network density

Social media connect more people than ever before. This increasing density of communication networks has been proposed to affect opinion perception with regard to the spiral of silence (Sohn & Geidner, 2016). We therefore varied the number of edges m with which each new user is linked to the existing graph. An increase in m also increases the density of the network. We analysed values of m from 1 to 10. For each value of m, 1000 simulations were run and the results averaged. All other parameters remained constant. The graph contained 1000 nodes. Each node was assigned a random opinion ("+" or "−" with a probability of 50% each). The parameter that controls the degree distribution of the network was set to 2.5, a value common for observed "real-world" scale-free networks, whose exponent usually lies between 2 and 3 (Dorogovtsev, Mendes, & Samukhin, 2001). For this first analysis bots were excluded.
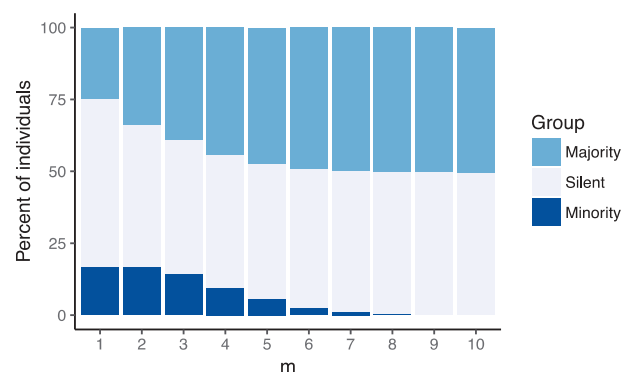

Figure 5. Effects of density on dominance of majority over minority opinion and the proportion of silenced users.
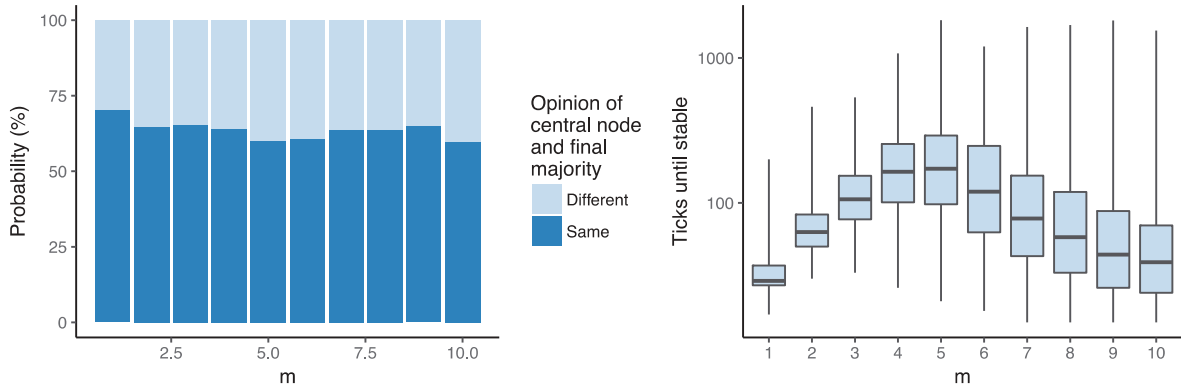
Figure 6. (left) Influence of the most central node on the final majority opinion; (right) Average number of ticks needed until the simulations reached a stable state.
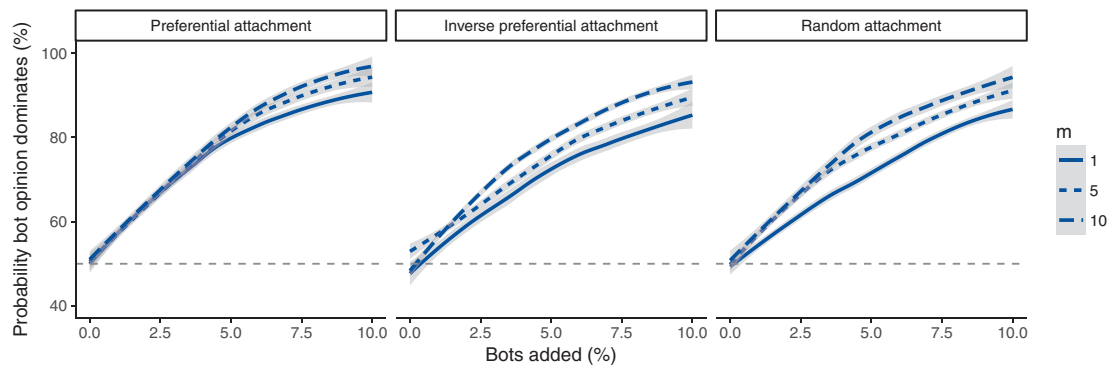

Figure 7. Influence of bots added at different positions in the network. Curves are Loess smoothers (span = 0.75) fit to the empirical probabilities to highlight the trend (Cleveland, 1979). Grey areas around the curves are 95% confidence intervals. Straight dashed lines indicate the 50% baseline probability.
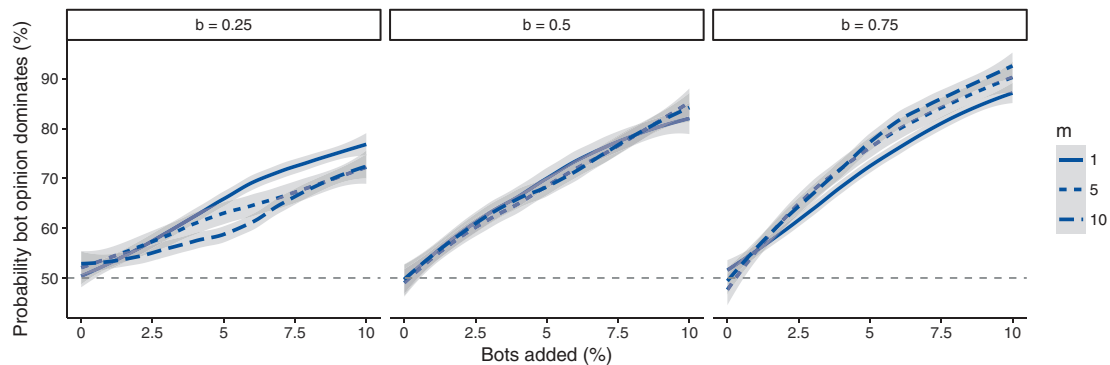

Figure 8. Influence of bots with varying bot influence factor b (using preferential attachment). Curves are Loess smoothers (span = 0.75) fit to the empirical probabilities to highlight the trend. Grey areas around the curves are 95% confidence intervals. Straight dashed lines indicate the 50% baseline probability.

With increasing network density, the majority group gets larger and the minority smaller. Evidently, barring ties, one of the two opinions is always held by more actors than the other opinion at the end of the run. We use the term "majority" to refer to the opinion expressed by the majority of actors that express their opinions at the end of the run, while "minority" refers to the actors of the other opinion confident enough to speak. Silenced actors are ignored. As expected, in half of the simulation runs the final majority is of the positive opinion, while in the other half of the cases the majority is of the negative opinion. However, the difference between the majority opinion and the minority increases with larger m (Figure 5). For $m = 1$, the minority opinion is expressed by 16.8% of actors at the end of the simulation while at $m = 10$, it is only expressed by 0.1%. At the same time, the size of the majority approaches 50% for large m, as virtually all agents that hold it are confident enough to express it.

As a result, the proportion of silent users stays roughly the same. With higher m, more holders of the majority opinion dare to speak, but roughly as many holders of the minority opinion are silenced.

In other words, when individuals' ego-networks are small, subgroups within the network are able to retain confidence in their minority views and discuss them openly. This observation is in accordance with the finding by Sohn and Geidner (2016) that several local "spirals of silence" form, whereas in a setting with large ego-networks, a global spiral of silence emerges. It is also consistent with the findings by Nowak et al. (1990).

5.2.2. The role of mass media

The network model also allows us to observe phenomena that are more difficult to observe in spatial environments. Some users are naturally more central than others, and the findings indicate that the position of the actors in the network can strongly influence the outcome. Specifically, the initial opinion of the most central node is equal to the final opinion of the majority in 70.4% of cases for $m = 1$, and 59.8% for $m = 10$ of cases (Figure 6, left). There is a weak but distinct downward trend as m is increased. To put this result into context, recall that the opinion of a user is assigned randomly in our model. Thus, if the central node did not influence the final results at all, its opinion would be expected to coincide with that of the final majority about 50% of the time. As can be seen, the initial opinion of the most central node coincides with the majority much more often, both for small and for large values of m.

The central nodes in our model represent opinion leaders such as the mass media which are able to shape people's views but are also influenced by many individuals in how they portray public opinion (e.g. through opinion polling or extensive friendship networks). Our results indicate that these actors strongly influence the outcome of the spiral of silence, although their influence is diminished slightly in a more densely networked society.

5.2.3. The speed of consensus formation

Finally, the agent-based model does not only make it possible to examine the outcome of the process, but also the process itself. The number of ticks needed until a stable state was reached exhibits an interesting behaviour (Figure 6, right). First, it has a large variance. For $m = 5$, the 90% confidence interval is [42; 566.3]. Secondly, it increases with larger m, from a mean of 36.0 for $m = 1$ to a mean of 210.1 for $m = 5$. For an even larger m, it decreases again, down to 63.9 for $m = 10$. As mentioned before, the stable state was determined as the point in time at which no agents changed their $\hat{c}(t)$ value by more than $\varepsilon = 0.001$. We observed that for both, very small

and very large m, a stable state was reached quickly. For values of m in between, it took much longer for a consensus to be reached (Figure 6, right).

The number of ticks it takes for a stable state to be reached can be interpreted as the time until a societal consensus is formed, or a norm established. The large variance in the number of ticks means that our approach can model cases in which public opinion forms quickly as well as those where this process takes years or even decades. In each of these cases, we have shown that the time taken depends on the number of connections between actors. A possible explanation for this observation is the superposition of two phenomena: The increase in the magnitude of occurring "spirals of silence" and the increase in the number of ways through which information can be diffused. On the one hand, as Sohn and Geidner (2016) found, the absence of many connections between the users leads to the emergence of local spirals of silence, i.e. small groups form in which one opinion dominates. Only in more densely connected communities, a global spiral of silence occurs. As expected, the more global the spiral is (and the higher m), the longer it should take for it to reach a stable state. On the other hand, information can diffuse faster through a more densely connected graph. Thus, the time it takes for a global opinion climate to form should decrease with an increase in m. Both effects combined may lead to the graph seen in the right part of Figure 6. A simulation with a higher m might resemble what has been labelled as "world public opinion" (Rusciano, 2014), referring to topics of worldwide relevance that are discussed by a global public.

Overall, the present evaluation supports the notion of the spiral of silence theory (Noelle-Neumann, 1974) and offers initial evidence on the circumstances under which a spiral process can come to a paralysis (Matthes & Hayes, 2014).

5.3. Sensitivity analysis

The above results are based on the modelling assumption that the value $\hat{c}(t)$ and *willingness to self-censor* are initialised by drawing values in [0, 1] uniformly. In order to judge the influence of this initialisation procedure on the results, we conducted a sensitivity analysis. The model was tested using different distributions to initialise the variables. More precisely, we additionally initialised both variables with normal distributions ($\mu = 0$ and 0.5, respectively, $\sigma = 0.5$) truncated at 0 to prevent negative values. We also examined an initialisation with a more extreme exponential distribution ($\mu = 1$). There is no difference in the final majority/minority distribution for any of these variants compared to the uniform initialisation. It seems that the stable final

state is not influenced by the choices for the initial *confidence* and *willingness to self-censor*. The number of ticks needed for the model to converge is also not affected that much. We observe that the effect of the initialisation wears off a few steps into the simulation.

To conclude the validation, we showed earlier that (1) the individual-level assumptions of our model are backed up by empirical evidence, and have shown that (2) the group-level predictions of our model are equally consistent with prior findings, where such findings are available. Since our model is consistent with empirical evidence, it seems plausible that it is an accurate representation of the spiral of silence process. We can now turn to using it to examine the influence of social bots on opinion formation.

6. Influence of social bots

6.1. Network structure and network positions of bots

The final step is to analyse the influence of social bots on the perceived opinion climate. We model a scenario in which bots that are programmed to spread opinions are inserted into the social network. If they succeed in triggering a spiral of silence, their views are eventually accepted as public opinion. We therefore ran simulations on a network of human agents and bots, and calculated the probability that the opinion spread by the bots ultimately prevails. We created 1000 "human" agents and added x% of bots to the graph, where x varied from 0 to 10. The bots were added according to the three different attachment strategies described earlier, while humans were always added using preferential attachment. Bots added with preferential attachment attempt to sway hubs or "influencers", while inverse preferential attachment models bots that try to influence peripheral and potentially more vulnerable nodes. Random attachment models a bot that picks nodes uniformly at random. The total resulting number of agents in the graph was 1000 + 10x. For each parameter combination 1000 simulations were run. Again, each human was either assigned a positive or negative opinion with a probability of 50% each. In contrast, all bots were assigned the "–" opinion. We then analysed the number of simulations in which the "–" opinion supported by the bots was the perceived majority opinion at the end. Without the influence of bots, each opinion should dominate in 50% of the simulation runs on average. In contrast, the "–" opinion supported by bots should dominate in 100% of runs if bots determine the results alone. To also assess the effect of network density, we carried out this procedure for multiple values of m. Figure 7 shows the results. We formulate several propositions based on our findings (P1–P6).

Table 1. Influence of bots added at different positions in networks of varying density (m).

| m | % Bots | % Bot Opinion Dominates | | |
		Preferential attachment	Inverse preferential attachment	Random attachment
1	0	48.1	44.3	50.7
	5	70.6	73.2	80.4
	10	90.8	86.1	86.2
5	0	50.3	54.7	50.7
	5	82.4	75.2	77.3
	10	94.3	90.3	91.3
10	0	52.7	49.2	52.1
	5	83.2	78.6	80.7
	10	97.3	92.9	94.5

After running the simulations with these parameters, we found that (P1) a small number of bots of one opinion is already sufficient to tip over the global opinion climate in favour of that opinion and (P2) the addition of bots always has an influence on their chances of success, regardless of the attachment method, network density or percentage of bots added. The influence of bots added with preferential attachment, inverse preferential attachment as well as random attachment were close with differences mainly in regard to the network densities' contribution. However, preferential attachment results in a slightly higher influence of bots than either of the other attachment methods. In other words (P3), the influence of these human-like bots is slightly higher when they attach to central actors in the network than when they attach to peripheral actors or to random actors. In case of preferentially attached bots, when a mere 2.0–2.5% (depending on m) of actors are bots, the opinion climate tips over in their favour in two out of three cases. The addition of 10% of bots already ensures that the " – " opinion is perceived as the dominant opinion in 90.8%–97.3% of simulation runs although it is only supported by 54.5% of actors (100 bots and around 500 humans). The result that a relatively small number of bots could effectively tip over the global opinion climate remained relatively stable regardless of the network density. However (P4), the influence of human-like bots is slightly higher in scenarios with denser networks in contrast to less connected networks. For example, with random attachment and $m = 10$, only 2.5% of bots were required to tip the opinion climate two out of three times, and 10% of bots swayed the opinion climate in 94.5% of cases. In comparison, in case of $m = 1$, 4.0% of bots were required for a 66% success chance, and 10% for a 86.2% chance. Table 1 presents more detailed results.

6.2. Human likeness of bots

The above results assume that bots influence human users in the same way other humans do. They differed in their positions in the network, but they had

equal influence on the confidence of their neighbours. However, in reality, bots might be perceived differently from humans. A human might doubt the authenticity of a message and be influenced by it less in their confidence than they would if the message had been spread by a human. To model this possibility, we also ran the simulations with an additional bot influence factor b in $[0; 1]$. When an agent counts its neighbours of the same (n_s) or opposite opinion (n_o), we now add b if this neighbour is a bot and, as before, 1 if it is a human. In other words, when $b = 0.5$, a human is influenced half as much by a bot as by a human. Again, we ran 1000 simulations for varying values of b (0.25, 0.5, 0.75) and m ($m = 1, \ldots, 10$). Bots were added using the preferential attachment method.

As expected, bots in scenarios with a lower bot influence factor at the individual level also have less influence on the final group-level outcome (see Figure 8). However, the effect of a few bots is still quite large: if 10% of agents are bots, the bots' opinion dominates in 75% of cases even when the influence of a bot is only one fourth of that of a human. What is even more surprising is the observation that the density of the network has different effects depending on the bot influence factor. For a low value of b ($b = 0.25$), sparse networks ($m = 1$) are more susceptible to bot influence. However, for a high bot influence factor ($b = 0.75$), dense networks ($m = 10$) are more vulnerable to bots. In other words, (P5) simplistic bots are more successful in sparse networks, while realistic, human-like bots benefit from highly connected networks. Finally, by judging the relative importance of all four studied variables, we can conclude that (P6) the influence of bots depends much more on their number and their degree of sophistication than on the structure of the network or their position in it.

Overall, the results show the theoretical potential for the application of automated bot accounts in the opinion formation process online. Through the addition of bots, the balance of the opinion expression tendency was influenced by the bots' opinion tendency in every scenario, regardless of the mechanism with which the bot accounts were linked to existing nodes. The results were affected by the degree to which each individual bot affected its neighbouring human users, but even when they were perceived as much less credible than humans, they still had a considerable influence on the outcome. This finding underlines the notion that in practice bots can sway public opinion—or the expression thereof—in one way or another. The conjunction of the mechanisms of astroturfing and the spiral of silence, which makes perfect sense in theory, is also apparent when simulating the alleged mechanism through an ABM.

While the findings do not show that social bots have an influence on the actual opinion, they may have a normative social influence on the tendency to express one's opinion. As Munger (2017) showed, messages sent by social bots can influence the willingness to speak about a topic or use certain words. While in his study the bots directly addressed individual users, it is easy to see this effect carrying over to a broader audience in a social network as simulated in the current paper. In a scenario where a discussion takes place e.g. on Facebook or Twitter and social bots take one side of an argument, actors with the opposite opinion may be less likely to express their own because they perceive an opinion climate that has been distorted by the artificial opinions of the social bots and is therefore contrary to their own.

7. Implications

The findings discussed above have several important implications for research in Information Systems. In carrying out the first two research steps of developing and validating the model, this paper contributes to the study of human communication, particularly online. It extends the literature on the spiral of silence theory with an explicit model of the process, while building on and refining previous attempts to do so. This model makes it possible to observe the bridge from individual to group behaviour and observe the "spiral" in the spiral of silence (see Matthes, 2015). The information systems discipline is uniquely suited to explore this bridge especially in the context of social media networks. As Lee (2001) put it, "research in the information systems field examines more than just the technological system, or just the social system, or even the two side by side; in addition, it investigates the phenomena that emerge when the two interact" (p. iii). The broader research question to ask in IS is, therefore: How does information technology-mediated communication affect the emergence of the spiral of silence?

In terms of theoretical implications for spiral of silence research, the present work not only corroborates the propositions on the spiralling process from the original theory but also offers insights into how the density of a network accelerates or slows down the spiralling process. Employing a network-based environment, the present work underlines the importance of network characteristics when it comes to observing the dynamic process within the spiral of silence theory.

By addressing the main question of how likely social bots are to influence public opinion formation, this paper further contributes to the growing field of research into the relationship between information systems and social bots and, more generally, the

manipulation of online opinion formation. It describes a plausible mechanism of manipulation on the basis of an established theory of opinion formation. While most of the current research on this field has either focused on testing various techniques to identify social bots (e.g., Subrahmanian et al., 2016) or on describing the behaviour of these bots (e.g., Hegelich & Janetzko, 2016), our findings demonstrate the eligibility of these approaches and link their mode of action to mechanisms described by the spiral of silence theory. Moreover, IS research studies the effects of the internet on political campaigns and elections (Wattal, Schuff, Mandviwalla, & Williams, 2010). Given the potential impact of social bots demonstrated by the present simulation, it becomes clear that their existence can considerably undermine democratic processes (e.g., inhibit the free exchange of opinions). IS research needs to consider how technology helps defend freedom of expression or endangers it (Rowe, 2018), and we hope that our work can contribute to this debate.

In terms of implications for future research, we conceptualised human interactions as a network and our findings partially reproduced results from related models of the spiral of silence that did not use a network model, further supporting those results. However, in the network model, there are central actors which may greatly influence the opinion climate, and thereby the final consensus adopted. This distinguishes our model from previous spatial models of the spiral of silence. In the analysis of bots, the network model allowed us to conclude that the influence of the bots depends on their position in the network and the density of the network. Bots were slightly more powerful when attached to popular users, and usually more powerful in dense networks. In the future, this network approach to agent-based modelling can be used to examine other opinion formation processes in an IS context by modelling different actors than social bots. For example, this approach could be used to study the effects of organisational structure on opinion formation.

The spiral of silence theory is new in information systems research. It explains normative social influence, one of the major causes of subjective peer pressures. In the future, this theory can be applied in many other IS-related contexts influenced by social norms, including technology adoption decisions. Finally, in future research the simulation approach could be combined with Big Data Analytics (Müller, Junglas, Vom Brocke, & Debortoli, 2016) by simulating the spiral of silence in several real-world scenarios gathered from social media. The model could also be extended by incorporating other theories on public opinion formation, such as those pertaining to the influence of exogenous events.

In terms of practical implications, organisations will be interested in these findings because people rely on social media to form decisions, for instance, to purchase certain products, to consult particular physicians or to vote for specific political candidates (Goh, Heng, & Lin, 2013; Matook et al., 2015; Metzger, Flanagin, & Medders, 2010). If online endorsements are created by bots, and the human users do not recognise the bots, this decision-making process is impaired. Similarly, data gathered from social media can be used to predict variables such as firm equity and the results of votes (Luo, Zhang, & Duan, 2013; Stieglitz, Meske, Ross, & Mirbabaie, 2018), which means that manipulations of online opinion could have serious negative consequences for businesses. Not all bots are malicious, however, and social media users might accept the danger of their opinion being influenced if the bot offers useful features such as reducing information overload by summarising information. Ethical guidelines will need to be developed to address the trade-off between these conflicting goals. If we accept that bots will become part of the social media landscape, how can we ensure that their influence is transparent, and that users know who controls them? Whose responsibility is it to ensure this, and to detect bots? Who should have the right to distinguish between a helpful bot that facilitates civic education and a harmful one that spreads propaganda (Stieglitz et al., 2017), and to decide which bots should be deleted? Which political actors should be allowed to advertise on social media? Our findings should further alert the field of internet policy and governance as well as the managers of social media platforms to develop accurate ways to detect and quickly delete social bots. The enforcement of guidelines, such as requiring bot profiles to be clearly labelled as such, also poses technical challenges. Methods from natural language processing, social network analysis, and other fields will need to be combined to build effective systems that minimise the danger of the manipulation of public opinion without censoring legitimate political voices. At the same time, there also needs to be an open discussion about whether this solution is at all promising. It might equally be argued that privately owned businesses such as Facebook, or even the state, should not intervene, and that it is more important for citizens, voters and consumers to develop the media literacy that allows them to decide for themselves who to trust.

8. Limitations

Of course, the limitations of the spiral of silence theory apply equally to our model. The spiral of silence is limited to modelling normative social

influence. Accordingly, the actors in our agent-based model do not change their underlying opinions but only their willingness to express these opinions.

As with any model, the present approach makes simplifying assumptions on the structure of the network. Ties in the network represent local "real-life" relationships based on physical proximity as well as far-reaching connections that are common in social media. Two agents are connected if they are expected to be able to influence each other's decision-making somehow. Future research could consider how different types of connections influence the spiral of silence, e.g., directed edges that represent unidirectional influences.

In addition, our model makes the assumption that the reference groups that influence opinion formation stay roughly the same for the duration of the opinion formation process. Given that some political debates last decades or longer, this assumption could be relaxed in future research by modelling dynamic graphs where edges and nodes are deleted and added over time. In reality, people seek out new connections to like-minded individuals (Mcpherson, Smith-Lovin, & Cook, 2001), and occasionally sever existing ties, for example if their opinions differ (John & Dvir-Gvirsman, 2015; Noel & Nyhan, 2011). Future research could consider how these dynamics affect the spiralling process, and how, together with the spiral of silence, they may lead to the development of "filter bubbles" or "echo chambers", in which individuals are exposed to online network environments that merely reinforce but do not challenge their viewpoints on issues (Flaxman & Rao, 2016).

The results on social bots likewise depend on several assumptions. Importantly, we only consider bots that attempt to influence public opinion and who are not immediately recognisable to humans as non-human agents. However, users of social media face an overload of information (Maier, Laumer, Eckhardt, & Weitzel, 2015), and resort to heuristics to evaluate the credibility of information (Metzger et al., 2010). Previous research suggested that misinformation (as given by manipulated entities such as social bots) can influence people's cognitions even when the information or its source has been discredited or debunked (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). Thus, the effects of social bots can be detrimental, even when they are identified as artificial actors. Additionally, we considered the case where exactly half the population is of one opinion, while the other half is of the opposite opinion. This is an adequate model for many real-world scenarios, especially elections with two major candidates and polarising decisions such as the Brexit referendum, but there are other situations in which a much smaller proportion of people supports one of the two sides. Future research should analyse these cases.

Empirical studies estimate that 9–15% of Twitter accounts are bots (Chu et al., 2012; Varol et al., 2017), and according to our simulations, it only takes 2–4% of human-like social bots to sway the opinion climate in two out of three cases. While these numbers pointing to the theoretical efficacy of social bots might appear alarming, it has to be considered that not all bots act this similarly to humans on social media, at least not yet. The more human-like they act, the better they can persuade human users or at least have an impact on perceptions of the public opinion. Against this backdrop, it is possible and even likely that bots will be able to influence opinion formation online if they continue to improve and succeed in imitating humans.

9. Conclusion

We conceptualised the spiral of silence as a complex adaptive system and modelled it using agent-based modelling. The results indicate that the increase in connectedness between individuals caused by modern information technology, including social media, may further contribute to the marginalisation of minority opinions. These results are consistent with previous theoretical considerations and empirical findings, which lends support to the validity of our model.

Using this model, we examined the influence of social bots from a theoretical perspective. A relatively small number of bots was sufficient to sway the opinion climate in the direction of the opinion supported by the bots, triggering a spiral of silence process that ultimately led to the bot opinion becoming accepted as the perceived majority opinion. Overall, this work successfully demonstrated the applicability of agent-based models to the spiral of silence and its relevance in conjunction with the current topic of social bots, effectively expanding existing research on all three relevant fields.

ORCID

Björn Ross 🔟 http://orcid.org/0000-0003-2717-3705
Stefan Stieglitz 🔟 http://orcid.org/0000-0002-4366-1840

References

Abokhodair, N., Yoo, D., & Mcdonald, D. W. (2015). Dissecting a social botnet. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing – CSCW '15,* Vancouver, BC, Canada (pp. 839–851). New York, NY: ACM.

Agarwal, R., & Dhar, V. (2014). Editorial — Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research, 25*(3), 443–448.

Aiello, L. M., Deplano, M., Schifanella, R., & Ruffo, G. (2012). People are strange when you're a stranger: Impact and influence of bots on social networks. In: *Proceedings of the Sixth International Conference on Weblogs and Social Media (ICWSM-2012),* Dublin, Ireland (pp. 10–17). Palo Alto CA: The AAAI Press.

Aral, S., Dellarocas, C., & Godes, D. (2013). Social media and business transformation: A framework for research. *Information Systems Research, 24*(1), 3–13.

Lee, A. S. (2001). Editor's comments: MIS quarterly's editorial policies and practices. *MIS Quarterly, 25*(1), iii–vii.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*(5439), 509–512.

Beese, J., Haki, M. K., & Aier, S. (2015). On the conceptualization of information systems as socio-technical phenomena in simulation-based research. In: *Proceedings of the International Conference on Information Systems,* Fort Worth, TX (pp. 1–19).

Berners-Lee, T. (2017). three challenges for the web, according to its inventor. [Online] Retrieved 1 June 2018, from http://webfoundation.org/2017/03/web-turns-28-letter/

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday, 21*, 11.

Bhattacherjee, A., Davis, C. J., Connolly, A. J., & Hikmet, N. (2017). User response to mandatory IT use: A Coping Theory perspective. *European Journal of Information Systems.* https://link.springer.com/journal/41303/onlineFirst/page/1

Bode, L. (2016). Pruning the news feed: Unfriending and unfollowing political content on social media. *Research & Politics, 3*(3), 1–8.

Canessa, E., & Riolo, R. L. (2006). An agent-based model of the impact of computer-mediated communication on organizational culture and performance: An example of the application of complex systems analysis tools to the study of CIS. *Journal of Information Technology, 21*(4), 272–283.

Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2012). Detecting automation of Twitter accounts: Are you a human, bot, or cyborg? *IEEE Transactions on Dependable and Secure Computing, 9*(6), 811–824.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association, 74*(368), 829–836.

Cossu, J.-V., Labatut, V., & Dugué, N. (2016). A review of features for the discrimination of twitter users: Application to the prediction of offline influence. *Social Network Analysis and Mining, 6*, 1.

Curşeu, P. L. (2006). Emergent states in virtual teams: A complex adaptive systems perspective. *Journal of Information Technology, 21*(4), 249–261.

Dewan, S., Ho, Y.-J. (Ian), & Ramaprasad, J. (2017). Popularity or proximity: Characterizing the nature of social influence in an online music community. *Information Systems Research, 28*(1), 117–136.

Dommers, S., Van der Hofstad, R., & Hooghiemstra, G. (2010). Diameters in preferential attachment models. *Journal of Statistical Physics, 139*(1), 72–107.

Dorogovtsev, S. N., Mendes, J. F., & Samukhin, A. N. (2001). Size-dependent degree distribution of a scale-free growing network. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics, 63*(6 Pt. 1), 062101.

Dorogovtsev, S. N., Mendes, J. F. F., & Samukhin, A. N. (2000). Structure of growing networks with preferential linking. *Physical Review Letters, 85*(21), 4633–4636.

Edwards, C., Edwards, A., Spence, P. R., & Shelton, A. K. (2014). Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior, 33*, 372–376.

Fan, H., & Lederman, R. (2018). Online health communities: How do community members build the trust required to adopt information and form close relationships? *European Journal of Information Systems, 27*(1), 62–89.

Fang, X., Hu, P. J. H., Li, Z. L., & Tsai, W. (2013). Predicting adoption probabilities in social networks. *Information Systems Research, 24*(1), 128–145.

Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday, 22*, 8.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Flaxman, S. R., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly, 80*(Special Issue), 298–320.

Forelle, M. C., Howard, P. N., Monroy-Hernandez, A., & Savage, S. (2015). Political bots and the manipulation of public opinion in venezuela. *SSRN Electronic Journal.*

Gao, G. (Gordon), Greenwood, B. N., Agarwal, R., & Mccullough, J. S. (2015). Vocal minority and silent majority: How do online ratings reflect population perceptions of quality. *MIS Quarterly, 39*(3), 565–590.

Gawroński, P., Nawojczyk, M., & Kułakowski, K. (2015). Opinion formation in an open system and the spiral of silence. *Polish Academy of Sciences Institute of Physics, 127*(3-A). http://przyrbwn.icm.edu.pl/APP/PDF/127/a127z3ap07.pdf

Glynn, C. J., Hayes, A. F., & Shanahan, J. (1997). Perceived support for one's opinions and willingness to speak out: A meta-analysis of survey studies on the 'spiral of silence'. *The Public Opinion Quarterly, 61*(3), 452–463.

Glynn, C. J., & Huge, M. E. (2014). Speaking in spirals: An updated meta-analysis of the spiral of silence. In W. Donsbach, C. T. Salmon, & Y. Tsfati (Eds.), *The spiral of silence: New perspectives on communication and public opinion* (pp. 65–72). New York, NY: Routledge.

Goh, K.-Y., Heng, C.-S., & Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content. *Information Systems Research, 24*(1), 88–107.

Grange, C., & Benbasat, I. (2018). Opinion seeking in a social network-enabled product review website: A study of word-of-mouth in the era of digital social networks. *European Journal of Information Systems*, 27 (6), 629–653.

Hampton, K., Raine, L., Lu, W., Dwyer, M., Shin, I., & Purcell, K. (2014) *Social* media and the 'spiral of silence.' [Online] Retrieved 1 June 2018, from http://www.pewinternet.org/2014/08/26/social-media-and-the-spiral-of-silence/

Havakhor, T., Soror, A. A., & Sabherwal, R. (2018). Diffusion of knowledge in social media networks: Effects of reputation mechanisms and distribution of knowledge roles. *Information Systems Journal*, 28(1), 104–141.

Hayes, A. F., Glynn, C. J., & Shanahan, J. (2005). Willingness to self-censor: A construct and measurement tool for public opinion research. *International Journal of Public Opinion Research*, 17(3), 298–323.

Hayes, A. F., Uldall, B. R., & Glynn, C. J. (2010). Validating the willingness to self-censor scale II: Inhibition of opinion expression in a conversational setting. *Communication Methods and Measures*, 4(3), 256–272.

Hegelich, S., & Janetzko, D. (2016). Are social bots on Twitter political actors? Empirical evidence from a Ukrainian social botnet. In: *Proceedings of the Tenth International Conference on Weblogs and Social Media (ICWSM-2016)*, Cologne, Germany (pp. 579–582). Palo Alto, CA: The AAAI Press.

Hildebrand, C., Hofstetter, R., & Herrmann, A. (2012). Modeling viral marketing dynamics in social networks – Findings from computational experiments with agent-based simulation models. In: *Proceedings of the International Conference on Information Systems*, Orlando, FL, USA.

Holland, J. H. (1995). *Hidden order: How adaptation builds complexity*. Boston, MA: Addison-Wesley.

Jiang, L. L., Hua, D. Y., Zhu, J. F., Wang, B. H., & Zhou, T. (2008). Opinion dynamics on directed small-world networks. *European Physical Journal B*, 65(2), 251–255.

John, N. A., & Dvir-Gvirsman, S. (2015). I don't like you any more: Facebook unfriending by israelis during the Israel-Gaza conflict of 2014. *Journal of Communication*, 65(6), 953–974.

Johnson, S. L., Faraj, S., & Kudaravalli, S. (2014). Emergence of power laws in online communities: The role of social mechanisms and preferential attachment. *MIS Quarterly*, 38(3), 795–808.

Kane, G., Labianca, G., & Sp, B. (2014). What's different about social media networks? A framework and research agenda. *MIS Quarterly*, 38(1), 274–304.

Knobloch-Westerwick, S. (2014). *Choice and preference in media use: Advances in selective exposure theory and research*. New York, NY: Routledge.

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131.

Luo, X., Zhang, J., & Duan, W. (2013). Social media and firm equity value. *Information Systems Research*, 24(1), 146–163.

Maier, C., Laumer, S., Eckhardt, A., & Weitzel, T. (2015). Giving too much social support: Social overload on social networking sites. *European Journal of Information Systems*, 24(5), 447–464.

Matook, S., Brown, S. A., & Rolf, J. (2015). Forming an intention to act on recommendations given via online social networks. *European Journal of Information Systems*, 24(1), 76–92.

Matthes, J. (2015). Observing the 'spiral' in the spiral of silence. *International Journal of Public Opinion Research*, 27(2), 155–176.

Matthes, J., & Hayes, A. F. (2014). Methodological conundrums in spiral of silence research. In W. Donsbach, C. Salmon, & Y. Tsfati (Eds.), *The spiral of silence: New perspectives on communication and public opinion* (pp. 54–63). New York, NY: Routledge.

Matthes, J., Knoll, J., & Von Sikorski, C. (2018). The "Spiral of Silence" revisited: A meta-analysis on the relationship between perceptions of opinion support and political opinion expression. *Communication Research*, 45(1), 3–33.

Matthes, J., Morrison, K. R., & Schemer, C. (2010). A spiral of silence for some: Attitude certainty and the expression of political minority opinions. *Communication Research*, 37(6), 774–800.

Mcpherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413–439.

Müller, O., Junglas, I., Vom Brocke, J., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal of Information Systems*, 25(4), 289–302.

Munger, K. (2017). Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior*, 39(3), 629–649.

Nan, N. (2011). Capturing bottom-up information technology use processes: A complex adaptive systems model. *MIS Quarterly*, 35(2), 505–532.

Nan, N., & Lu, Y. (2014). Harnessing the power of self-organization in an online community during organizational crisis. *MIS Quarterly*, 38(4), 1135–1157.

Neubaum, G., & Krämer, N. C. (2017). Monitoring the opinion of the crowd: Psychological mechanisms underlying public opinion perceptions on social media. *Media Psychology*, 20(3), 502–531.

Neubaum, G., & Krämer, N. C. (2018). What do we fear? Expected sanctions for expressing minority opinions in offline and online communication. *Communication Research*, 45(2), 139–164.

Newman, M. E. J. (2001). Clustering and preferential attachment in growing networks. *Physical Review E*, 64 (2), 025102.

Noel, H., & Nyhan, B. (2011). The 'unfriending' problem: The consequences of homophily in friendship retention for causal estimates of social influence. *Social Networks*, 33(3), 211–218.

Noelle-Neumann, E. (1974). The spiral of silence: A theory of public opinion. *Journal of Communication*, 24(2), 43–51.

Noelle-Neumann, E. (1977). Turbulences in the climate of opinion: Methodological applications of the spiral of silence theory. *Public Opinion Quarterly*, 41(2), 143–158.

Noelle-Neumann, E. (1993). *The spiral of silence: Public opinion, our social skin* (2nd ed.). Chicago, IL: University of Chicago Press.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, 97(3), 362–376.

Reeves, B., & Nass, C. (1996). *How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.

Ren, Y., & Kraut, R. E. (2007). An agent-based model to understand tradeoffs in online community design. In: *Proceedings of the International Conference on Information Systems*, Montreal, Quebec, Canada.

Rowe, F. (2018). Being critical is good, but better with philosophy! From digital transformation and values to the future of IS research. *European Journal of Information Systems*, 27(3), 380–393.

Rusciano, F. L. (2014). World public opinion. In W. Donsbach, C. T. Salmon, & Y. Tsfati (Eds.), *The spiral of silence: New perspectives on communication and public opinion* (pp. 179–186). New York, NY: Routledge.

Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213), 1063–1064.

Sargent, R. G. (2005). Verification and validation of simulation models. In: *Proceedings of the 37th Winter Simulation Conference (WSC'05)*, Orlando, FL, USA (pp. 130–143).

Schäfer, F., Evert, S., & Heinrich, P. (2017). Japan's 2014 general election: Political bots, right-wing internet activism, and prime minister Shinzō Abe's hidden nationalist agenda. *Big Data*, 5, 4.

Scheufele, D. (2007). Spiral of silence theory. In W. Donsbach & M. W. Traugott (Eds.), *The SAGE handbook of public opinion research* (pp. 173–183). London, UK: SAGE.

Scheufele, D. A., & Moy, P. (2000). Twenty-five years of the spiral of silence: A conceptual review and empirical outlook. *International Journal of Public Opinion Research*, 12(1), 3–28.

Sohn, D., & Geidner, N. (2016). Collective dynamics of the spiral of silence: The role of ego-network size. *International Journal of Public Opinion Research*, 28(1), 25–45.

Song, H., & Boomgaarden, H. G. (2017). Dynamic spirals put to test: An agent-based model of reinforcing spirals between selective exposure, interpersonal networks, and attitude polarization. *Journal of Communication*, 67(2), 256–281.

Stieglitz, S., Brachten, F., Ross, B., & Jung, A.-K. (2017). Do social bots dream of electric sheep? A categorisation of social media bot accounts. In: *Proceedings of the Australasian Conference on Information Systems*, Hobart, Tasmania, Australia.

Stieglitz, S., Meske, C., Ross, B., & Mirbabaie, M. (2018). Going back in time to predict the future - the complex role of the data collection period in social media analytics. *Information Systems Frontiers*.

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, 39, 156–168.

Subrahmanian, V. S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., ... Menczer, F. (2016). The DARPA twitter bot challenge. *Computer*, 49(6), 38–46.

Susarla, A., Oh, J. H., & Tan, Y. (2012). Social Networks and the diffusion of user-generated content: Evidence from youtube. *Information Systems Research*, 23(1), 23–41.

Sykes, T. A., Venkatesh, V., & Gosain, S. (2009). Model of acceptance with peer support: A social network perspective to understand employees' system use. *MIS Quarterly*, 33(2), 371–393.

Takeuchi, D., Tanaka, G., Fujie, R., & Suzuki, H. (2015). Public opinion formation with the spiral. *Nonlinear Theory and Its Applications*, 6(1), 15–25.

Thiele, J. C. (2014). R marries netlogo: Introduction to the RNetLogo package. *Journal of Statistical Software*, 58(2), 1–41.

Thies, F., Wessel, M., & Benlian, A. (2016). Effects of social interaction dynamics on platforms. *Journal of Management Information Systems*, 33(3), 843–873.

Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. In: *Proceedings of the Eleventh International Conference on Web and Social Media (ICWSM-2017)*, Montréal, Québec, Canada (pp. 280–289). Palo Alto, CA:The AAAI Press.

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for direction? Gender, social influence and their role in technology acceptance and usage behaviour. *MIS Quarterly*, 24(1), 115–137.

Vidgen, R., & Wang, X. (2009). Coevolving systems and the organization of agile software development. *Information Systems Research*, 20(3), 355–376.

Wattal, S., Schuff, D., Mandviwalla, M., & Williams, C. B. (2010). Web 2.0 and politics: The 2008 US presidential election and an E-politice research agenda. *MIS Quarterly*, 34(4), 669–688.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small world' networks. *Nature*, 393(6684), 440–442.

Wilensky, U. (1999). NetLogo. *Center for connected learning and computer-based modeling, northwestern university*.

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modeling: Modeling natural, social, and engineered complex systems with netlogo*. Cambridge, MA: MIT Press.

Wu, Y., Du, Y. J., Li, X. Y., & Chen, X. L. (2015). Exploring the spiral of silence in adjustable social networks. *International Journal of Modern Physics C*, 26(11), 1550125.

Zhang, J., Zhang, R., Zhang, Y., & Yan, G. (2013). On the impact of social botnets for spam distribution and digital-influence manipulation. In: *2013 IEEE Conference on Communications and Network Security (CNS)*, National Harbor, MD, USA (pp. 46–54).