WORKing papers in Management Science

WORMS/25/08

Emergence of pluralistic ignorance: An agent-based approach

Katarzyna Sznajd-Weron[1]
Barbara Kamińska[1]

[1] Department of Computational Social Science,
Wrocław University of Science and Technology, Poland

WORMS is a joint initiative of the Management Science departments of the Wrocław University of Science and Technology,
Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland

Emergence of Pluralistic Ignorance: An Agent-Based Approach

Katarzyna Sznajd-Weron and Barbara Kamińska
Wroclaw University of Science and Technology, Poland
Department of Computational Social Science

Abstract

Pluralistic ignorance is a puzzling social psychological phenomenon in which the majority of group members privately reject a norm yet mistakenly believe that most others accept it. Consequently, they publicly comply with the norm. This phenomenon has significant implications for politics, economics, and organizational dynamics because it can mask widespread support for change and hinder collective responses to large-scale societal challenges. The aim of this work is to demonstrate how agent-based modeling, a computational approach well-suited for studying complex social systems, can be applied to investigate pluralistic ignorance. Rather than providing a systematic literature review, we focus on several models, including our own two models based on the psychological Social Response Context Model, as well as two other representative models: one of the first and most influential computational models of self-enforcing norms, and a model of opinion expression based on a silence game. For all of these models, we provide custom `NetLogo` implementations, publicly available at https://barbarakaminska.github.io/NetLogo-Pluralistic-ignorance/, which allow users not only to run their own simulations but also to follow the algorithms step by step. In conclusion, we note that despite differences in assumptions and structures, these models consistently reproduce pluralistic ignorance, suggesting that it may be a robust emergent phenomenon.

1 Introduction

"How is this?" said the Emperor to himself. "I can see nothing! This is indeed a terrible affair! Am I a simpleton, or am I unfit to be an Emperor? That would be the worst thing that could happen—Oh! The cloth is charming," said he, aloud. "I approve of it completely." He smiled most graciously and looked closely at the empty looms. No way would he say that he could not see what two of his advisors had praised so much. Everyone with the Emperor now strained his or her eyes hoping to discover something on the looms, but they could see no more than the others.

The above paragraph is a quote from The Emperor's New Clothes, a classic fairy tale by Hans Christian Andersen (Andersen, 1837), often cited in the context of pluralistic ignorance (Bicchieri, 2005; Bicchieri & Fukui, 1999; Centola et al., 2005; Tang et al., 2022), and with good reason. It vividly illustrates the phenomenon despite being written nearly a century before the concept of pluralistic ignorance was introduced (Katz & Allport, 1931). Privately, each individual disbelieves

Figure 1: Illustration of the scene depicting the Emperor's public fitting of clothes made from a wonderful fabric woven by thieves, from the fairy tale The Emperor's New Clothes by Hans Christian Andersen (Andersen, 1837). Neither the Emperor nor any of the grandees of his court can see the clothes, but they all claim they can and express their appreciation. Each voices an opinion that does not align with their private belief, because each fears being the only one unable to see the clothes. Thus, they all experience pluralistic ignorance and, at the same time, engage in preference falsification. Illustration by Marcin Weron, 2025.

in the existence of the magical cloth. Yet, publicly, they express admiration for it. They believe that others can genuinely perceive the cloth and that their own doubt signals personal inadequacy. The situation escalates when the emperor agrees to wear clothes made from the magical fabric in a parade. During the public fittings, he and the other members of his court claim that the fabric is magnificent, even though they cannot see it, as illustrated in Fig. 1. This discrepancy between private belief and public expression leads everyone to conform outwardly to what they mistakenly think is the group consensus. Thus, the story captures the essence of pluralistic ignorance. At the same time, it is a good example of preference falsification, the act of publicly expressing opinions or preferences that differ from one's true beliefs, usually to avoid social, political, or economic repercussions (Kuran, 1998).

As noted by Taylor (1982), over the years several definitions of pluralistic ignorance, ranging from broader to more narrow, have appeared in the literature. Here, we adopt the one used by Bicchieri and Fukui (1999): "*Pluralistic ignorance, a psychological state characterized by the belief that one's private thoughts, attitudes, and feelings are different from those of others, even though one's public behavior is identical.*" According to Bicchieri and Fukui (1999), the term "ignorance" is somewhat misleading because individuals are not entirely unaware or uninformed. They rather make systematic errors in interpreting the motives, attitudes, and beliefs of others.

These interpretations are guided by what they observe, specifically, the public behavior of others. The central problem is that individuals mistakenly conclude that, unlike themselves, others must genuinely think and feel the way they publicly act. This observation is absolutely fundamental and we will return to it when discussing the assumptions of various computational models.

Pluralistic ignorance, understood as mistaken impressions of how other people feel and think about various matters, is a well-documented phenomenon in public opinion research (Shamir & Shamir, 1997). It affects many aspects of social life and often hinders collective action (for a recent review, see (Miller, 2023)). Beyond its broad applications in political sciences (Kuran, 1998), it has been used in economics (De Langhe, 2014; Wenzel, 2005), and in management (Abraham et al., 2022; Halbesleben & Buckley, 2004; Halbesleben et al., 2007).

However, particularly compelling empirical findings have been recently reported in the context of climate change. For instance, Sparkman et al. (2022) found that although 66–80% of Americans support major climate mitigation policies, they estimate that only 37–43% of others do. Another study (Sparkman et al., 2025) focused on proactive climate narratives that combine personal responsibility with support for policy-level action. It showed that such narratives are widely endorsed, yet both the public and policymakers significantly underestimate their popularity. Furthermore, although most people believe in human-caused climate change, this public consensus is collectively underestimated. A recent study showed that people in 11 countries substantially misperceive others' views on climate change (Geiger et al., 2025). These studies highlight how pluralistic ignorance can conceal widespread support for change, thereby impeding collective responses to large-scale societal challenges. Understanding the dynamics of pluralistic ignorance and the collective behaviors that emerge from it is therefore more important than ever.

One of the methods used to study pluralistic ignorance is agent-based modeling (Centola et al., 2005; Duggins, 2017; Huang & Wen, 2014; Merdes, 2017; F. Seeme et al., 2019). In our view, this is the most suitable computational approach for investigating a wide range of complex systems, including societies and organizations, which becomes evident once you understand what a complex system is and what an agent-based model (ABM) is.

As noted by Ladyman and Wiesner (2020), there is no universally accepted definition of "complexity" or "complex system," and it remains debated whether such a definition is even possible or necessary. In fact, many different definitions have been proposed. However, these definitions have one key, common feature: complex systems are composed of many interacting elements, and their behavior arises not only from the properties of their components, but also from the interactions among them. Typically, a complex system as a whole exhibits behaviors that none of its parts exhibit on their own, a phenomenon known as emergence.

Agent-based models (ABMs), an approach that has been gaining popularity in psychology and cognitive science (Bilewicz & Soral, 2020; Jackson et al., 2017; Madsen et al., 2019; Olsson & Galesic, 2024), represent a system as a collection of autonomous entities, called agents. In our context, these agents represent humans that can interact with one another as well as with external factors. ABMs allow researchers to precisely define the behavior of any number of agents and to study how individual-level interactions can generate complex, system-level outcomes.

The goal of this work is to explain how agent-based modeling can be used to study pluralistic ignorance. Therefore, in Section 2, we describe the concept and the basic building blocks required for an ABM of pluralistic ignorance. In Section 3, we present the development of our own approach, beginning with its theoretical foundation in the four-dimensional psychological model of social response proposed by Nail et al. (2000). Finally, in Section 4, we present examples from two other ABMs of pluralistic ignorance (Centola et al., 2005; Gaisbauer et al., 2020). While many ABMs have been proposed across various disciplines (Dong et al., 2024; Duggins, 2017; Huang & Wen, 2014; F. Seeme et al., 2019; Ye et al., 2019), with a short review available in (Kamińska & Sznajd-Weron, 2025), we chose to focus only of these two models rather than conduct a systematic literature review. This approach aligns with the goal of the paper, which is to explain the methodology so that inexperienced readers can understand it and, in the future, read the literature on ABMs of pluralistic ignorance. To facilitate reader understanding and engagement, for all the models presented in Sections 3 and 4 we provide custom `NetLogo` implementations, which allow users not only to run their own simulations but also to follow the algorithms step by step. All are publicly available at https://barbarakaminska.github.io/NetLogo-Pluralistic-ignorance/.

2 Basic building block of ABM

First, we state that there is no such thing as **the** agent-based model of pluralistic ignorance because many different models have been proposed across various disciplines (Centola et al., 2005; Dong et al., 2024; Duggins, 2017; Gaisbauer et al., 2020; Huang & Wen, 2014; F. Seeme et al., 2019; Ye et al., 2019). As reviewed by Kamińska and Sznajd-Weron (2025), different disciplines use different terminology and sometimes the term "pluralistic ignorance" is not used, even though a given model may still describe it.

Probably the best example of such a situation is one of the first, if not the first, formal models that explicitly distinguishes between privately held and publicly declared preferences and therefore can account for phenomena such as preference falsification and pluralistic ignorance as proposed by Kuran (1987). The aim of his work was to explain why collective decisions are path-dependent, a phenomenon also known as social hysteresis (Sznajd-Weron et al., 2024), rather than to study pluralistic ignorance. In fact, Kuran (1987) does not even explicitly mention pluralistic ignorance, although his model is applicable to research in this area. Moreover, he does not describe his approach as agent-based modeling. In fact, although the approach later known as agent-based modeling was already introduced in the social sciences by Schelling (1971), it was formally established by Epstein and Axtell (1996) and Holland and Miller (1991), which came after Kuran's framework. Nevertheless, Kuran's model can be presented as an ABM, and it will be done so later in this section. But first, we need to explain what we mean by an agent.

The rise of the personal computer era in the late 1970s, followed by major advances in computer-based modeling and machine learning from the 1990s onward, provided the computational power that enabled the widespread adoption of computational methods across many disciplines. As Holland and Miller (1991) emphasized, these advances enabled the creation and study of artificial adaptive agents (AAA) in diverse artificial worlds over extended timescales. Such agents form complex adaptive systems, i.e., networks of interacting entities in which emergent phenomena appear as a result of these interactions. According to Holland and Miller (1991), an agent is adaptive if its actions in an environment can be assigned a value (e.g., payoff, fitness) and it acts to increase that value over time. Several years later, Epstein and Axtell (1996) wrote: "*Agents are people of artificial societies. Each agent has internal states and behavioral rules. Some states are fixed for the agent's life, while others change through interaction with other agents or with the external environment.*" This means that we do not necessarily need to describe agents as utility maximizers, which is the typical approach in economics and was suggested by Holland and Miller

(1991). In fact, in our ABMs (Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025), which will be described later, we do not introduce any utility or fitness.

However, in Kuran's model (Kuran, 1987), agents explicitly maximize a utility function that reflects not only material payoffs but also social and psychological factors. In his framework, the decision to act, such as whether to express a political opinion, depends on the trade-off between these different components of utility. His work is better described as a *framework* rather than a single model, since Kuran (1987) presented several versions, starting from the simplest and progressively adding complexity. It describes a society of N agents indexed by i, each with a private preference $x_i \in [0, 1]$ over two mutually exclusive policies: $P = 0$ (e.g., status quo) and $P = 1$ (e.g., alternative). An agent's publicly expressed preference is denoted $y_i \in [0, 1]$. In the simplest version of the model, the private preference x_i is fixed over time, but in the more advanced versions it is allowed to evolve in response to new information or social influence.

The elements of Kuran's framework described above can be seen as a basic building block for an agent-based model of pluralistic ignorance – namely, an agent characterized by two variables: private and public opinions. The latter is also called an expressed or external opinion in some fields, as reviewed recently by Kamińska and Sznajd-Weron (2025). These opinions can change over time under the influence of economic, social, or psychological factors. Of course, depending on the ABM, agents can be more or less complex. For example, Centola et al. (2005) proposed a model in which, in addition to private belief and compliance (public conformity), agents are described by conviction strength and enforcement. We will describe that model in Section 4. On the other hand, Jedrzejewski et al. (2018) and Kamińska and Sznajd-Weron (2025) proposed ABMs in which agents are described only by their private and expressed (public) binary opinions.

At this point, we want to highlight an important assumption of all ABMs described here (Centola et al., 2005; Gaisbauer et al., 2020; Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025; Kuran, 1987): an agent's perception of others' opinions is based entirely on what those others express. This assumption reflects the observation made by Bicchieri and Fukui (1999), mentioned in the introduction, that people wrongly assume others must truly think and feel the way they publicly act, even though they themselves do not. Finally, we note that throughout the rest of the paper, we will refer to an agent using the pronoun "it" to make clear that the agent is not a real person but a simulated representation of one.

3 From the psychological model of social response to ABM

In this section, we focus on the development of our own ABM (Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025). Although it was not originally designed to study pluralistic ignorance, the model is well suited to this purpose. The inspiration for our model came from a four-dimensional descriptive framework of social response proposed by Nail et al. (2000), later termed the Social Response Context Model (SRCM) (MacDonald et al., 2004). SRCM categorizes individual responses to social influence along four dimensions: public and private opinions, measured both before (pre-exposure) and after (post-exposure) an influence attempt. Each dimension is treated as a binary variable, agreement or disagreement, resulting in 16 distinct combinations that represent different types of social response, as shown in Fig. 2.

We emphasize that the symbolic scheme of social responses depicted in Fig. 2 has not appeared in previous publications, but it was inspired by the diagram provided by Nail et al. (2000). We

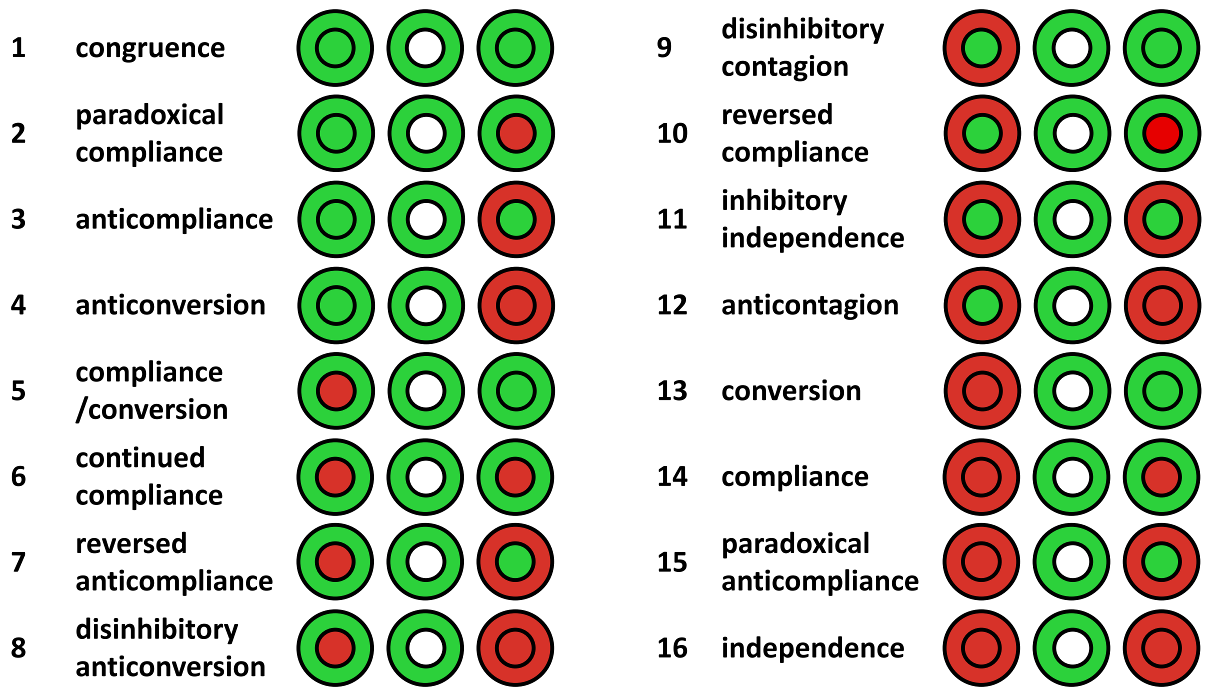

Figure 2: **Possible responses to social influence, as defined by the Social Response Context Model** (MacDonald et al., 2004; Nail et al., 2000). In each triplet, the first circle represents the target's pre-exposure state, the second represents the influence source, and the third represents the target's post-exposure state. The outer circle denotes the public (external/expressed) opinion, while the inner circle denotes the private (internal) opinion. Green (light gray in grayscale) indicates agreement or acceptance; red (dark gray in grayscale) indicates disagreement or opposition; white indicates an unknown opinion (a target cannot see private opinion of the source). In this example, the source of influence displays public acceptance, but analogous picture could be done for the situation in which the source of influence displays public opposition.

created this version to align with the visualizations used in our earlier work (Kamińska & Sznajd-Weron, 2025) and with our NetLogo implementation, available at https://barbarakaminska.github.io/NetLogo-Pluralistic-ignorance/q-voter-EPO/.

Unlike traditional models, as reviewed by Nail et al. (2013), the SRCM captures the complex and sometimes contradictory ways people may agree or disagree with influence sources, publicly or privately. For example, within SRCM it is possible to distinguish between two widely recognized types of conformity: conversion (conformity at both public and private levels) and compliance (public conformity without private acceptance); see responses number 13 and 14 in Fig. 2. The model also accounts for a particularly intriguing type of social response known as disinhibitory contagion (response number 9 in Fig. 2), which, like conformity, involves a change in public behavior toward that of an influence source. However, disinhibitory contagion differs in that it begins with the internal, intra-psychic conflict of the influencee before any exposure to the source of the influence.

To illustrate disinhibitory contagion, the following example was provided by Paul R. Nail: *You and your date are at a dance, but no one is dancing. You and your partner are inhibited from dancing because you would be embarrassed to be the only couple on the floor. This is the 'force' that prevents you from doing what you would like to do. At this point, however, one couple starts dancing (viz., the initiators or triggers), then another, and another. You and your date join in, and the dance floor quickly fills. Note that the influence of the initiators is contagious, as the influence spreads, and the situation progresses from your initial internal conflict over dancing, through influence, to your ultimate internal harmony, happy that you are dancing. In this case, it is only your and the other agents' public behavior that changes, not private opinions* (Jedrzejewski et al., 2018).

Another vivid example of disinhibitory contagion appears at the end of The Emperor's New Clothes, the same fairy tale we quoted at the beginning of this work. A little child suddenly says, "*But the Emperor has nothing at all on!*" This statement then spreads like contagion, freeing the crowd from the internal conflict of not seeing the clothes yet feeling compelled to praise them. Eventually, everyone cries out, "*But he has nothing at all on!*". The above example highlights another crucial difference between disinhibitory contagion and other types of conformity: it spreads like a virus, as a simple contagion, where a single person is enough to influence the target. Compliance, in contrast, is not a simple contagion. According to classical Asch experiments, a unanimous group of several people is required to influence the target, and once unanimity is broken, the likelihood of compliance decreases dramatically (Asch, 1956). Both types of social response are implemented in our ABMs (Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025).

Our ABMs describe a population of N agents, indexed by $i = 1, \ldots, N$, placed on the vertices of a given graph representing a social network. Agents can interact only if they are directly connected; such connected agents are referred to as *neighbors*. Different network structures can be considered. For example, Jedrzejewski et al. (2018) studied the original version of the model only on a complete graph, which corresponds to all-to-all interactions. In contrast, Kamińska and Sznajd-Weron (2025) examined a modified version on both a complete graph and real social networks from organizations. In our custom `NetLogo` implementation, we use a Watts–Strogatz graph, as shown in Fig. 3, one of the most popular models of social networks (Barabási & Pósfai, 2017). In the limiting case, it reduces to a complete graph, which allows us to reproduce the results of Jedrzejewski et al. (2018) and Kamińska and Sznajd-Weron (2025).

Each agent is characterized by a private opinion and an expressed opinion, both binary variables

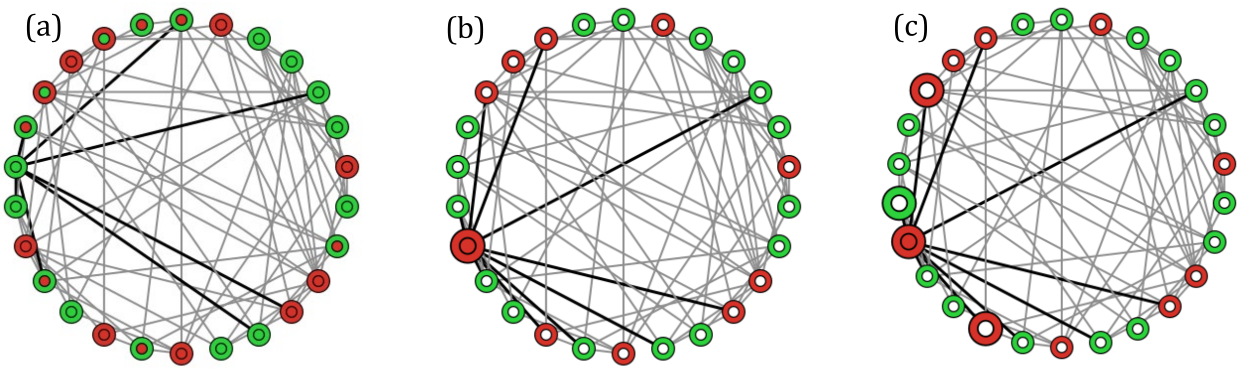

Figure 3: **Basic elements of our agent-based model.** (a) Each circle represents an agent, connected to others by social ties (gray lines). Each agent has a private opinion (inner circle) and an expressed opinion (outer circle), shown as either "agree" (green/lighter gray) or "disagree" (red/darker gray). (b) At each step, one target agent (larger circle) is selected and can only observe the expressed opinions of its direct contacts (white = unknown). (c) Only a subset of q contacts (here $q = 3$; larger circles) is visible to the target at any one time, and these may influence its expressed opinion. The subset changes over time, so agents never see all their contacts' opinions simultaneously.

(e.g., agree/disagree, acceptance/opposition, green/red, +/−), as in the SRCM, shown in Fig. 2. Both private and expressed opinions can change due to social influence from neighbors. The model dynamics proceed as follows: at each update step, one agent is randomly selected from the entire system to reconsider its opinion (the *target agent*), shown as a larger circle in panel (b) of Fig. 3. The target can only observe the expressed opinions of its direct connections. Therefore, in panels (b) and (c) of Fig. 3, the private opinions of other agents are shown in white, indicating they are unknown to the target. After a target is selected, a subset of q agents is randomly chosen from the target's neighbors, as shown in Fig. 3(c). These neighbors, also marked with larger circles, can influence the target's opinion. The target updates both its private and expressed opinions. Specifically, the expressed opinion is updated as follows:

1. With probability p, the target agent decides to act independently of others and express its private opinion, i.e., update the expressed opinion to match the private one.

2. With complementary probability $1 - p$, the target decides to look at q neighbors:

   - If the target is initially in *harmony* (private and expressed opinions are the same), it is less susceptible to social influence. Following the results of Asch's experiment (Asch, 1956), all q neighbors must express the same opinion to influence the target. If they are not unanimous, the target does not change its expressed opinion, as shown in Fig. 4. If they are unanimous, the target *complies* with them, i.e., updates its expressed opinion to match that of the q neighbors, as shown in Fig. 5.

   - If the agent is in *dissonance* (private and expressed opinions differ), it is motivated to resolve this internal conflict. In this case, it is sufficient that at least one of the q neighbors shares the same opinion as the target's private opinion to encourage it to

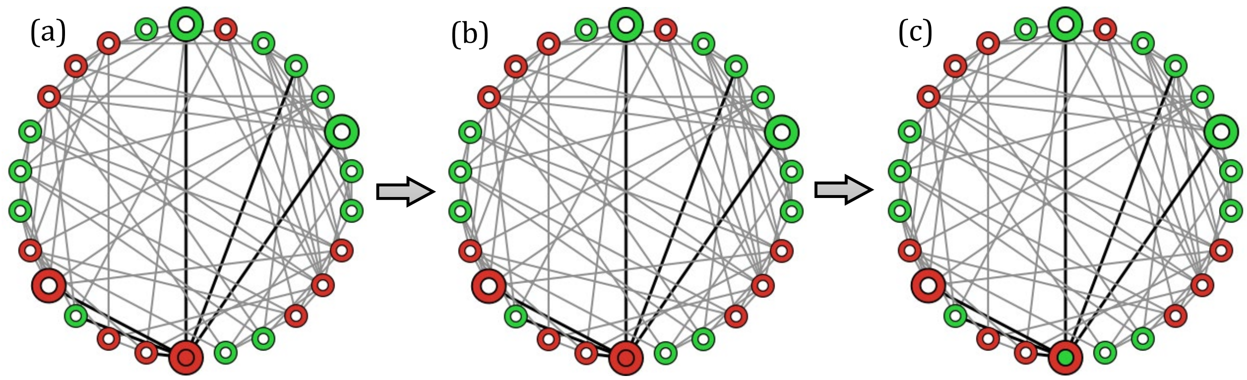

Figure 4: **Visualization of the lack of sufficient influence for compliance and private independence.** (a) A target (larger circle, publicly and privately disagreeing with a given idea) and a subset of $q = 3$ contacts (larger circles white inside = unknown private opinions) are selected. (b) The target updates its expressed opinion. Because it is not experiencing internal dissonance, it can change its expressed opinion only through compliance. The chosen neighbors are not unanimous, so they cannot influence the target. (c) The target updates its private opinion. Because the neighbors are not unanimous, it can change its private opinion only through independence, which occurs with probability p; in this example, it does.

"speak up," i.e., set its expressed opinion equal to its private opinion. This type of response corresponds to *disinhibitory contagion* and is illustrated in Fig. 6.

The private opinion can also change due to social influence, specifically:

1. With probability p, the target agent decides to change its private opinion independently of others. In this case, the target changes its opinion to the opposite with probability f, as shown in Fig. 5, or keeps its current opinion with probability $1 - f$, where usually $f = 1/2$ (Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025).

2. With complementary probability $1 - p$, the target decides to look at q neighbors. At this point, the specific rule depends on the version of the model:

   - Jedrzejewski et al. (2018) proposed the following: If all q neighbors express the same opinion, the target sets its private opinion to match theirs, as shown in Fig. 6. This can lead to a situation in which the agent ends up in dissonance: even though it privately agrees with the group, it continues to express an opposing opinion. Both conversion and paradoxical anticompliance are possible.
   - Kamińska and Sznajd-Weron (2025) introduced a mechanism to avoid cognitive dissonance (Festinger, 1957): The target adjusts its private opinion to match that of a unanimous q-panel only if it is consistent with the target's expressed opinion (only conversion is possible), as shown in Fig. 6.

To answer the question of whether there is a discrepancy between expressed and private opinions on the level of society, which we take as a sign of pluralistic ignorance, we measure the fraction of

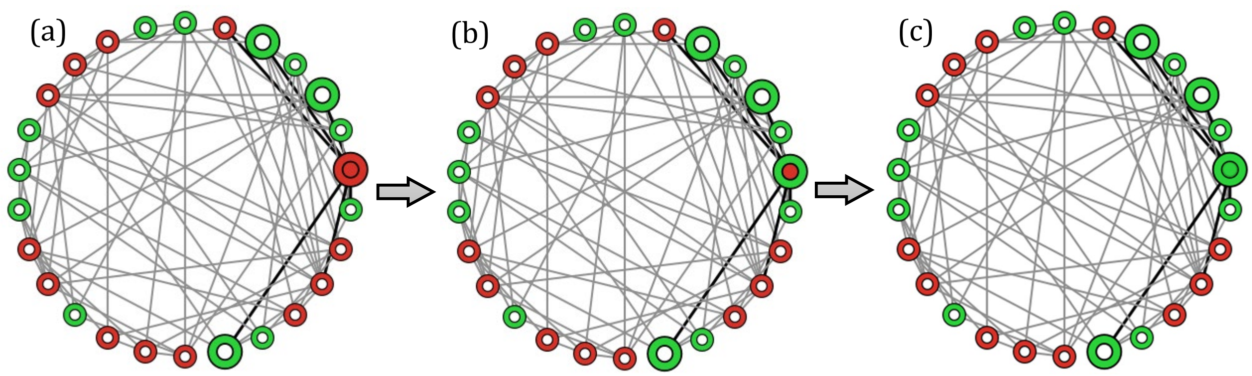

Figure 5: **Visualization of compliance and conversion in our EPO models.** (a) A agent (larger circle, publicly and privately disagreeing with a given idea) and a subset of $q = 3$ contacts (larger circles white inside = unknown private opinions) are selected. (b) The target updates its expressed opinion. Because it is not experiencing internal dissonance, it can change its expressed opinion only due to compliance. Because all q chosen neighbors express the same opinion (agree), so the target changes its expressed opinion to agree. (c) The target updates its private opinion. This can occur either independently or due to the unanimous opinion of the q neighbors. In this example, the target changes its private opinion due to the influence of neighbors, resulting in a full conversion during the update.

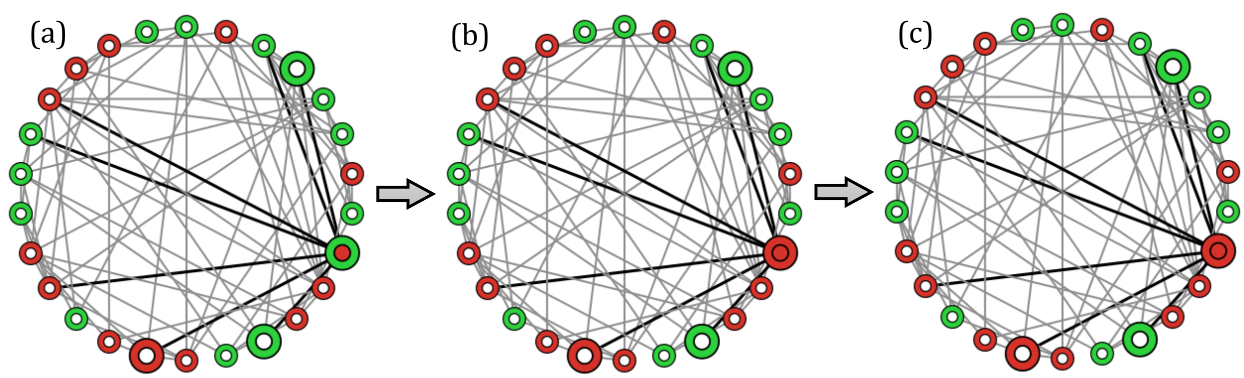

Figure 6: **Visualization of disinhibitory contagion in our EPO models.** (a) A target (larger circle, publicly agreeing and privately disagreeing with a given idea) and a subset of $q = 3$ contacts (larger circles white inside = unknown private opinions) are selected. (b) The target updates its expressed opinion. Because it is experiencing internal dissonance, it can change its expressed opinion either independently, with probability p, or due to disinhibitory contagion, with probability $1 - p$, if at least one of the q neighbors expresses the same opinion as the target's private opinion. (c) The target updates its private opinion. Because the opinions of the q neighbors are not unanimous, it can change its private opinion only due to independence, which occurs with probability p.

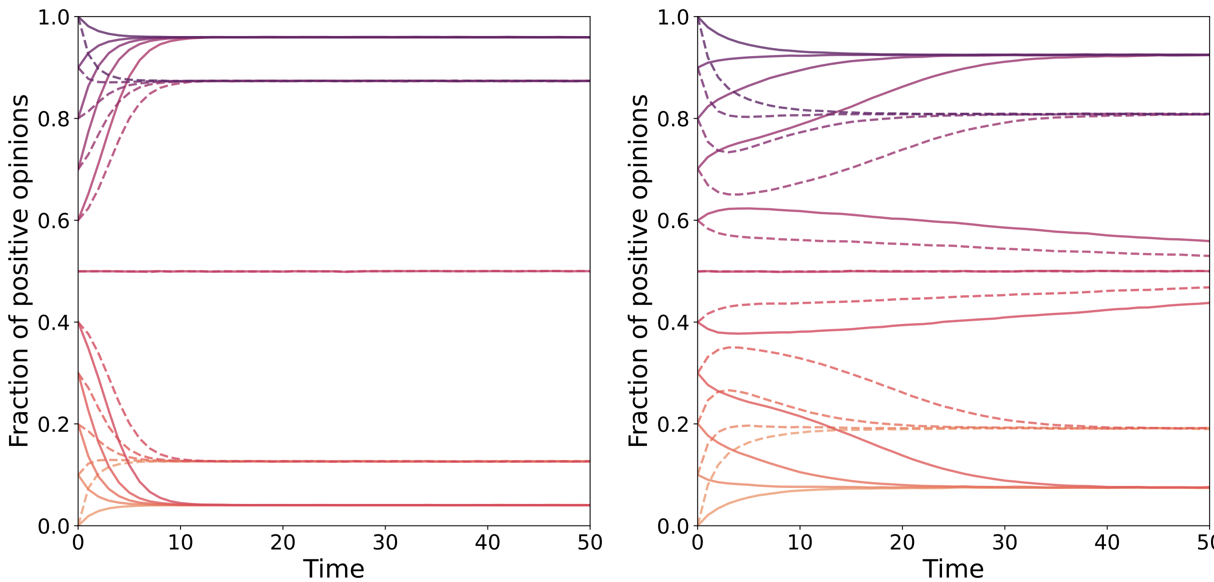

Figure 7: **Change in the fraction of positive expressed (solid lines) and private (dashed lines) opinions over time**: the left panel shows the model proposed by Jedrzejewski et al. (2018), the right panel the model proposed by Kamińska and Sznajd-Weron (2025). Results from 11 different initial conditions are shown; the darker the line, the higher the initial fraction of positive opinions. In this plot, agents are placed on a complete graph (all-to-all interactions), with the size of the influence group $q = 3$ and the probability of independence $p = 0.23$.

agents with a positive (agreeing) expressed opinion and a positive private opinion. Results for both versions of the model are shown in Fig. 7. This plot presents the results of a single set of model parameters ($p = 0.23, q = 3$) and several initial conditions. These specific values of parameters are provided as an example. The full range of parameters was studied and reported in (Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025). Initially, all agents have the same private (dashed lines) and expressed (solid lines) opinions, so we see pairs of dashed and solid lines overlapping at time=0. Over time, these lines separate, and a clear discrepancy between private and public opinions appear, which we interpret as the emergence of pluralistic ignorance.

As seen in Fig. 7, both the fraction of agents with a positive private opinion and the fraction with a positive expressed opinion change over time due to social influence. After some time, these fractions stabilize at certain values that, for each model, depend on the probability of independence p, the size of the influence group q, and, to some extent, the initial conditions. For a more detailed analysis, see the original papers (Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025). For the values of parameters shown in Fig. 7, the discrepancy between private and expressed opinions is larger in the model with dissonance reduction. We also see that in the model without the additional dissonance-reduction mechanism, as introduced by Jedrzejewski et al. (2018) the initial conditions play a smaller role than in the model with such a mechanism (Kamińska & Sznajd-Weron, 2025). This means that the second version of the model is more path-dependent, i.e., it exhibits *social hysteresis*, a particularly intriguing collective phenomenon (Sznajd-Weron et al., 2024).

All results presented in Fig. 7 come from the setting in which, for each chosen target, we first update the expressed opinion and then the private opinion. This naturally raises the question of which order is more appropriate, that is, whether the private or expressed opinion should be updated first. Since social psychology does not provide a clear answer, we tested both scenarios for each version of the model (Jedrzejewski et al., 2018; Kamińska & Sznajd-Weron, 2025). In the first scenario, called *Act then think (AT)*, the expressed opinion was updated first, followed by the private opinion. In the second scenario, called *Think then act (TA)*, the private opinion was updated first, followed by the expressed opinion. Before updating the opinion at each level, private or expressed, the source of influence is chosen independently.

For both versions of the model, as reported by Jedrzejewski et al. (2018) and Kamińska and Sznajd-Weron (2025) , the two scenarios produced identical results at the aggregate level, at least on the complete graph (all-to-all interactions). Consequently, Fig. 7 is identical for both scenarios. On more complex networks, the results differ but only slightly, as shown by Kamińska and Sznajd-Weron (2025). However, the results depend strongly on the model parameters. Interestingly, Kamińska and Sznajd-Weron (2025) showed that when the source of influence is a single person ($q = 1$), there is no discrepancy between the fraction of positive opinions expressed publicly and the agents' private opinions; in other words, pluralistic ignorance does not occur. As already mentioned, if one person is sufficient to spread an idea, this corresponds to simple contagion, in contrast to complex contagion (Centola & Macy, 2007), where influence requires more people. The model shows that under simple contagion, pluralistic ignorance does not arise.

Most importantly, pluralistic ignorance also depends on the version of the model. Notably, in (Kamińska & Sznajd-Weron, 2025), the additional mechanism of dissonance reduction causes people to begin thinking in alignment with their expressed opinions. This means that pluralistic ignorance may increase when individuals attempt to reduce dissonance by rationalizing their expressed opinions. This effect may help explain the large-scale pluralistic ignorance observed in climate-related issues, as discussed in the Introduction. We encourage the reader to experiment with these parameters using our `NetLogo` implementation, as illustrated in Fig. 8.

4 Two other ABMs of pluralistic ignorance

In this section, we present two additional ABMs that can be used to study pluralistic ignorance (Centola et al., 2005; Gaisbauer et al., 2020). Although many other ABMs have been proposed in the literature, some explicitly designed to address this phenomenon (Anderson & Ye, 2019; Brown et al., 2022; Duggins, 2017; Hoffmann et al., 2024; Huang & Wen, 2014; Merdes, 2017; Topuz & Yücel, 2024), we deliberately focus on just two. Notably, one of them was not originally intended to model pluralistic ignorance directly (Gaisbauer et al., 2020). However, under the assumptions adopted in this paper, particularly that agents infer others' beliefs solely from publicly expressed opinions, it can be interpreted as capturing the dynamics of pluralistic ignorance, even though the authors do not use that term.

Our goal is not to provide a systematic review, but rather to offer a deeper exploration of selected models, accompanied by interactive implementations that allow readers to experiment and gain hands-on understanding. The models we chose are relatively simple, making them easier to explain, implement, and modify. The first, proposed by Centola et al. (2005), is one of the earliest formal models addressing preference falsification and norm enforcement, and its influence on the field is substantial. The second, developed by Gaisbauer et al. (2020), was originally intended to model the spiral of silence, but as noted above, it can be reinterpreted as a model of pluralistic

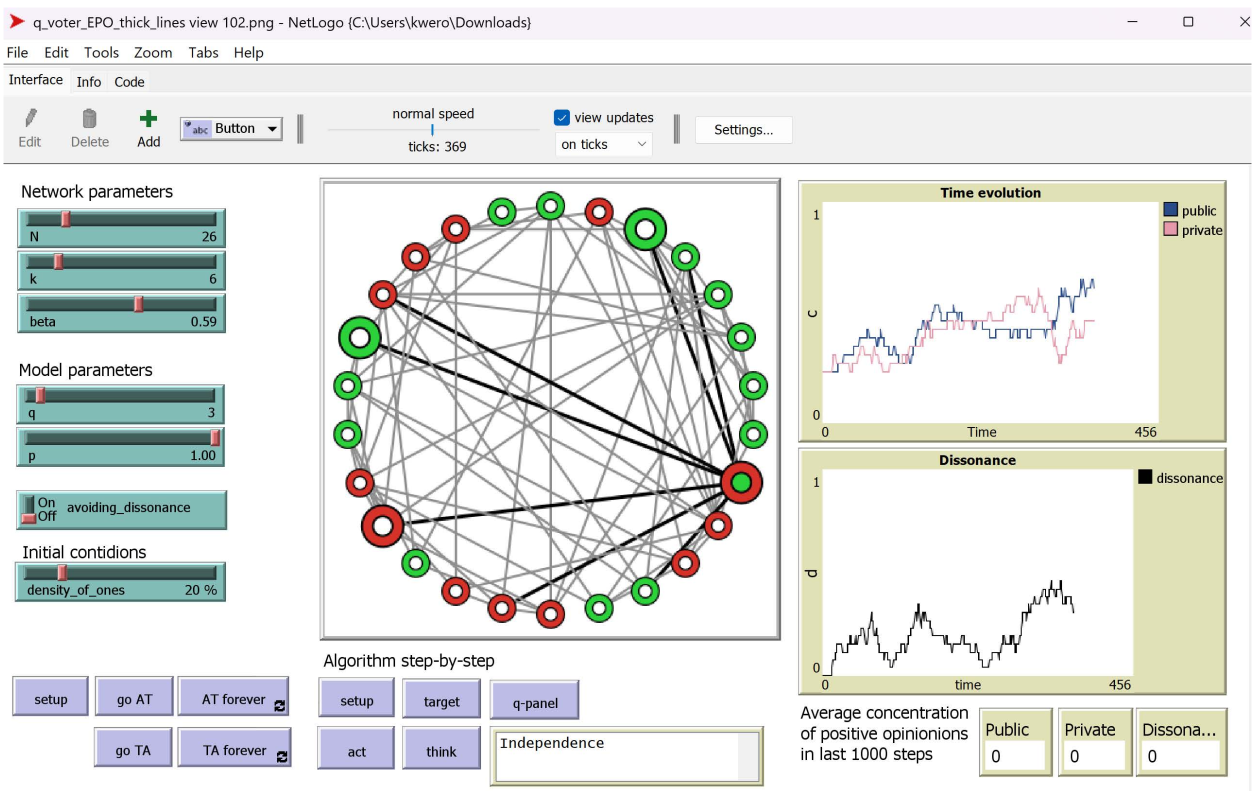

Figure 8: **NetLogo model view**, in which we can switch between version in which avoiding dissonance is on, as proposed by Kamińska and Sznajd-Weron (2025), or off, as proposed by Jedrzejewski et al. (2018). The model is publicly available at https://barbarakaminska.github.io/NetLogo-Pluralistic-ignorance/q-voter-EPO/. There is possibility to run the model online, or download offline version shown in this figure.

ignorance.

4.1 Agent-based model of the "emperor's dilemma"

In the model proposed by Centola et al. (2005) each agent i is described by the following variables:

- Private Belief $B_i \in \{-1, +1\}$. It indicates whether the agent is a disbeliever ($B_i = -1$) or a true believer ($B_i = +1$). In the basic model it is fixed in time, but in the extended models, agents may change belief from -1 to $+1$ if false enforcement persists and conviction is low.
- Conviction Strength $S_i \in [0, 1]$ represents the strength of the agent's belief and is fixed in time. For true believers ($B_i = 1$) $S = 1$, for disbelievers S is randomly drawn from uniform distribution $[0, 0.38]$, hence with mean value $\bar{S} = 0.19$.
- Compliance $C_i \in \{-1, +1\}$ indicates whether the agent complies with the norm ($C_i = +1$) or deviates ($C_i = -1$). It is dynamic variable, updated at each iteration based on local social pressure and conviction.
- Enforcement $E_i \in \{-1, 0, +1\}$ indicates whether the agent enforces deviance ($E_i = -1$), does not enforce ($E_i = 0$), or enforces compliance ($E_i = +1$). It is dynamic variable, updated at each iteration.

Agents update their compliance C_i and enforcement E_i by observing the behavior of their neighbors, that is those directly connected in the social network. They monitor how many neighbors comply with the norm versus deviate from it, and how many actively encourage others to comply or deviate. The compliance decision is made based on enforcement from neighbors and the strength of one's convictions. The decision to enforce is more complex and depends on whether the agent's decision to comply aligns with his beliefs. If so, the agent enforces if his convictions and perceived need for enforcement are stronger than the cost. If compliance opposes his beliefs, the agent enforces under the influence of neighbors if their influence exceeds the strength of his convictions and the cost. Otherwise, he does not enforce anything. This mechanism models how individuals respond to social influence in environments where norms may be privately rejected but publicly enforced. Equations describing exactly how C_i and E_i are updated are given in (Centola et al., 2005). We do not repeat them here to avoid any equations.

Centola et al. (2005) distinguish four types of agents: true believers (who publicly and privately support the norm), true disbelievers (who publicly and privately oppose it), and false believers and false disbelievers (whose public behavior contradicts their private beliefs due to social pressure). They also identify three types of enforcement: by true believers, by true disbelievers, and by false disbelievers (who privately oppose but publicly support the norm). They show that unpopular norms spread through cascades of false enforcement following false compliance. Even though some people disagree with the norm, they still comply with it (preference falsification). Furthermore, under strong enough social pressure, they will even enforce behavior that goes against their beliefs. However, these cascades of false enforcement require a locally clustered group of true believers who initiate the enforcement as they change the perceived majority opinion in their surroundings.

The readers who would like to know all details of the model can check the original paper (Centola et al., 2005) and our custom NetLogo implementation of the basic version of the model, in which private opinions are fixed, is available at https://barbarakaminska.github.io/NetLogo-Pluralistic-ignorance/emperors-dilemma/.

4.2 A silence game and dynamics of opinion expression

Gaisbauer et al. (2020) proposed an ABM based on a silent game in which agents, indexed as usual by $i = 1, \ldots, N$, are described by two binary variables: opinion o_i and action a_i. Opinions are static variables, meaning they do not change over time. If agent i holds opinion $o_i = 1$, it belongs to group G_1, while if agent i holds opinion $o_i = 2$, it belongs to group G_2. Each agent i can take one of two actions: publicly express its opinion $a_i = 1$ (denoted additionally by e) or remain silent $a_i = 0$ (denoted additionally by s). An agent decides which action to choose based on the state of its neighbors. If it believes it holds the opinion of the minority, it becomes silent; if it believes it holds the opinion of the majority, it expresses it. This judgment is made using only the information about the opinions of others that are publicly expressed. The authors additionally assume that expressing one's opinion carries a constant cost c.

The update rule is as follows. Agent i calculates two values: (V_e) the expected number of its neighbors who publicly agree with its opinion, and (V_s) the expected number of publicly disagreeing neighbors plus the cost of expressing an opinion. If $V_e > V_s$, the agent prefers to express its opinion, which means it updates its action to $a_i = 1$. If $V_e < V_s$, the agent prefers to be silent, which means it updates its action to $a_i = 0$. If $V_e = V_s$, the agent is indifferent and does not change its action a_i. We note that Gaisbauer et al. (2020) do not explicitly introduce the values V_e and V_s but instead provide equations that allow these values to be calculated depending on the opinion held by agent i. We use this notation here to avoid introducing equations. The latter part of the paper also proposes a more stochastic approach. The probability that an agent will express his opinion depends on his willingness to express Q. This is updated using reinforcement learning based on feedback from a randomly selected neighbor.

The model of Gaisbauer et al. (2020) demonstrates the possibility of a minority opinion achieving public dominance. However, the authors do not explicitly refer to pluralistic ignorance, but rather to the spiral of silence. As highlighted by Taylor (1982), both the spiral of silence and the pluralistic ignorance theories examine how social perception shapes the formation of public opinion. The spiral of silence framework assumes that people's perceptions of public opinion are fixed and investigates how these perceptions influence behavior. In contrast, research on pluralistic ignorance focuses on the accuracy of such perceptions, frequently finding that individuals' beliefs about the majority opinion are often mistaken. Under the assumption outlined in the last paragraph of Section 2, namely, that agents' perceptions of others' opinions are entirely based on what those others express, we interpret the result obtained by Gaisbauer et al. (2020) as an example of pluralistic ignorance, which is why we present this model here.

To enable the reader to better understand this model, or simply play with the model and conduct their own experiments, we have prepared our own special implementation of this model in NetLogo, publicly available at https://barbarakaminska.github.io/NetLogo-Pluralistic-ignorance/opinion-expression/.

5 A cautionary conclusion

We borrowed the title of this last section from Centola et al. (2005), who used it for their own concluding section in the paper on ABM of the "emperor's dilemma". In that section, they frankly acknowledge that their study uses a highly specific model of social influence and that it is not the only possible specification. They suggested to explore alternative models of the influence process. Since their seminal work, many other ABMs have been proposed to address the puzzle of pluralistic ignorance (Anderson & Ye, 2019; Brown et al., 2022; Duggins, 2017; Huang & Wen, 2014; F. B. Seeme & Green, 2016; Topuz & Yücel, 2024). Each of these ABMs not only uses different models of social influence, but also addresses pluralistic ignorance in other contexts. For example, (Hoffmann et al., 2024; Kamińska & Sznajd-Weron, 2025) examine the problem of pro-environmental change, noting that although many people have pro-environmental convictions, they often fail to translate these beliefs into actual behavior.

Although the assumptions and rules of these models differ substantially, all are able to account for the emergence of pluralistic ignorance. This convergence of outcomes across diverse ABMs suggests that pluralistic ignorance may be a robust emergent phenomenon, not overly sensitive to the specific assumptions or mechanisms of a given model. In other words, even when models differ in how they represent agents, social influence, or decision-making, they can still produce similar macro-level patterns, such as the persistence of unpopular norms or the discrepancy between private beliefs and public behavior.

Looking at various ABMs, it seems that the key ingredients for the emergence of pluralistic ignorance are (1) the existence of two levels of opinion, private and expressed (public), which do not necessarily align, (2) social influence, which shapes public expression based on perceived norms, and (3) a desire to reduce cognitive dissonance. However, the way these opinions are represented (binary, continuous, or even distributions instead of scalars, as proposed by Brown et al. (2022)), as well as the specific forms of social influence and cognitive dissonance reduction, seems to be less important. This observation leads us to reflect on the concept of universality and the role of ABMs in social modeling.

When thinking about universality, the first thought that comes to our mind is universality in the physics of critical phenomena. One of the great discoveries of twentieth-century physics was the observation that at criticality, various physical systems and mathematical models may display the same behavior on the macroscopic (system-level) scale, independently of microscopic details, that is, on the scale of particles or, in our case, agents. This critical behavior is described by so-called critical exponents, and it was shown that various systems can have exactly the same values of these exponents, and if so they belong to the same universality class. For a more detailed description, but still accessible to an interdisciplinary audience, we recommend the textbook Introduction to the Theory of Complex Systems by (Thurner et al., 2018). The association with critical phenomena is not accidental, because critical phenomena are among the most surprising emergent, collective phenomena that can appear in complex systems.

Of course, the universality we identified in the ABMs of pluralistic ignorance is much less formal. We did not introduce any measures, such as critical exponents, that would demonstrate a correspondence between different models, which, by the way, would be a desirable yet challenging research task for the future. For now, we have only observed that the same phenomenon can emerge in very different models. However, this observation brings us to the question of the role of ABMs in social science. Are they a substitute for large-scale psychological experiments, a forecasting tool, a means to build scenarios that can qualitatively answer what-if questions, or a way to formalize conceptual theories? This question, which we are not attempting to answer here, raises new ones. Is it important to evaluate which ABM is best for a given purpose or to answer a given question? How should we evaluate it?

Much has been written about the validation of agent-based models (ABMs) and how challenging it can be, as discussed recently by Jedrzejewski et al. (2025). The natural answer is that we should validate ABMs based on empirical data. However, our experience is that we can relatively easily calibrate models to empirical data, or even reproduce empirical data within relatively simple ABMs (Jedrzejewski et al., 2025; Sznajd-Weron & Weron, 2002; Sznajd-Weron et al., 2008). Of course, we can reject some models because they fail to describe the data. We can also compare different versions of models and evaluate which ones better reproduce the data (Jedrzejewski et al., 2025; Sznajd-Weron et al., 2008), but that still does not mean we have found the one true model. Therefore, in our opinion, it is not the particular ABM that matters; it is not about proving that there is one and only one correct model. It is about identifying the most important mechanisms that influence a given phenomenon. We can then also ask under which circumstances a given phenomenon will not appear, or will be weaker.

We would like to highlight that what we wrote above is just our private opinion, which we decided to express even though we believe that it may not be popular. After writing this concluding sentence, one more thought came to mind: perhaps our private opinion is in fact shared by others, and we are simply experiencing pluralistic ignorance.

Acknowledgments

This research was funded by National Science Centre Poland Grant 2019/35/B/HS6/02530. The authors would like to thank Paul R. Nail for confirming the validity of the symbolic scheme of the Social Response Context Model shown in Fig. 2, and Marcin Weron creating Fig. 1.

References

Abraham, M., Collischon, M., Grimm, V., Kreuter, F., Moser, K., Niessen, C., Schnabel, C., Stephan, G., Trappmann, M., & Wolbring, T. (2022). COVID-19, normative attitudes and pluralistic ignorance in employer-employee relationships. *Journal for Labour Market Research*, *56*(1), 19. https://doi.org/10.1186/s12651-022-00325-4

Andersen, H. C. (1837). *The Emperor's New Clothes.* (Lit2Go Edition). https://etc.usf.edu/lit2go/68/fairy-tales-and-other-traditional-stories/5637/the-emperors-new-clothes/

Anderson, B. D. O., & Ye, M. (2019). Recent Advances in the Modelling and Analysis of Opinion Dynamics on Influence Networks. *International Journal of Automation and Computing*, *16*(2), 129–149. https://doi.org/10.1007/s11633-019-1169-8

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*(9), 1–70. https://doi.org/10.1037/h0093718

Barabási, A.-L., & Pósfai, M. (2017). *Network science*. Cambridge University Press.

Bicchieri, C. (2005, December). *The Grammar of Society*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616037

Bicchieri, C., & Fukui, Y. (1999). The Great Illusion: Ignorance, Informational Cascades, and the Persistence of Unpopular Norms. *Business Ethics Quarterly*, *9*(1), 127–155. https://doi.org/10.2307/3857639

Bilewicz, M., & Soral, W. (2020). *Hate Speech Epidemic. The Dynamic Effects of Derogatory Language on Intergroup Relations and Political Radicalization* (tech. rep. No. S1).

Brown, G. D. A., Lewandowsky, S., & Huang, Z. (2022). Social sampling and expressed attitudes: Authenticity preference and social extremeness aversion lead to social norm effects and polarization. *Psychological Review*, *129*(1), 18–48. https://doi.org/10.1037/rev0000342

Centola, D., & Macy, M. (2007). Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, *113*(3), 702–734. https://doi.org/10.1086/521848

Centola, D., Willer, R., & Macy, M. (2005, January). The emperor's dilemma: A computational model of self-enforcing norms. https://doi.org/10.1086/427321

De Langhe, R. (2014). To specialize or to innovate? An internalist account of pluralistic ignorance in economics. *Synthese*. https://doi.org/10.1007/s11229-014-0436-z

Dong, J., Hu, J., Zhao, Y., & Peng, Y. (2024). Opinion formation analysis for Expressed and Private Opinions (EPOs) models: Reasoning private opinions from behaviors in group decision-making systems. *Expert Systems with Applications*, *236*, 121292. https://doi.org/10.1016/j.eswa.2023.121292

Duggins, P. (2017). A Psychologically-Motivated Model of Opinion Change with Applications to American Politics. *Journal of Artificial Societies and Social Simulation*, *20*(1). https://doi.org/10.18564/jasss.3316

Epstein, J. M., & Axtell, R. L. (1996, October). *Growing Artificial Societies*. The MIT Press. https://doi.org/10.7551/mitpress/3374.001.0001

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

Gaisbauer, F., Olbrich, E., & Banisch, S. (2020). Dynamics of opinion expression. *Physical Review E*, *102*(4), 042303. https://doi.org/10.1103/PhysRevE.102.042303

Geiger, S. J., Köhler, J. K., Delabrida, Z. N. C., Garduño-Realivazquez, K. A., Haugestad, C. A. P., Imada, H., Iyer, A., Maharja, C., Mann, D. C., Marczak, M., Melville, O., Nijssen, S. R. R., Powdthavee, N., Praptiwi, R. A., Ranade, G., Rosa, C. D., Vitale, V., Winkowska, M., Zhang, L., & White, M. P. (2025). What We Think Others Think and Do About Climate Change: A Multicountry Test of Pluralistic Ignorance and Public-Consensus Messaging. *Psychological Science*, *36*(6), 421–442. https://doi.org/10.1177/09567976251335585

Halbesleben, J. R., & Buckley, M. R. (2004). Pluralistic ignorance: historical development and organizational applications. *Management Decision*, *42*(1), 126–138. https://doi.org/10.1108/00251740310495081

Halbesleben, J. R., Wheeler, A. R., & Buckley, M. R. (2007). Understanding pluralistic ignorance in organizations: application and theory. *Journal of Managerial Psychology*, *22*(1), 65–83. https://doi.org/10.1108/02683940710721947

Hoffmann, T., Ye, M., Zino, L., Cao, M., Rauws, W., & Bolderdijk, J. W. (2024). Overcoming inaction: An agent-based modelling study of social interventions that promote systematic pro-environmental change. *Journal of Environmental Psychology*, *94*, 102221. https://doi.org/10.1016/j.jenvp.2023.102221

Holland, J. H., & Miller, J. H. (1991). Artificial Adaptive Agents in Economic Theory. *The American Economic Review*, *81*(2), 365–370.

Huang, C. Y., & Wen, T. H. (2014). A novel private attitude and public opinion dynamics model for simulating pluralistic ignorance and minority influence. *JASSS*, *17*(3), 09. https://doi.org/10.18564/jasss.2517

Jackson, J. C., Rand, D., Lewis, K., Norton, M. I., & Gray, K. (2017). Agent-Based Modeling. *Social Psychological and Personality Science*, *8*(4), 387–395. https://doi.org/10.1177/1948550617691100

Jedrzejewski, A., Kowalska-Pyzalska, A., Pawłowski, J., & Sznajd-Weron, K. (2025). Everyone is different, but does it matter? The role of heterogeneity in empirically grounded agent-based models of alternative fuel vehicles diffusion. *Operations Research and Decisions*, *35*(1), 45–80. https://doi.org/10.37190/ord250103

Jedrzejewski, A., Marcjasz, G., Nail, P. R., & Sznajd-Weron, K. (2018). Think then act or act then think? (S. Gonçalves, Ed.). *PLOS ONE*, *13*(11), e0206166. https://doi.org/10.1371/journal.pone.0206166

Kamińska, B., & Sznajd-Weron, K. (2025). Impact of cognitive dissonance on social hysteresis: Insights from the expressed and private opinions model. *Expert Systems with Applications*, *273*, 126851. https://doi.org/10.1016/j.eswa.2025.126851

Katz, D., & Allport, F. H. (1931). *Students' Attitudes: A Report of the Syracuse University Reaction Study*. Craftsman Press.

Kuran, T. (1987). Preference Falsification, Policy Continuity and Collective Conservatism. *The Economic Journal*, *97*(387), 642. https://doi.org/10.2307/2232928

Kuran, T. (1998, June). *Private Truths, Public Lies*. Harvard University Press. https://doi.org/10.2307/j.ctvt1sgqt

Ladyman, J., & Wiesner, K. (2020, August). *What Is a Complex System?* Yale University Press. https://doi.org/10.2307/j.ctv14rmpwc

MacDonald, G., Nail, P. R., & Levy, D. A. (2004). Expanding the Scope of the Social Response Context Model. *Basic and Applied Social Psychology*, *26*(1), 77–92. https://doi.org/10.1207/s15324834basp2601{_}7

Madsen, J. K., Bailey, R., Carrella, E., & Koralus, P. (2019). Analytic Versus Computational Cognitive Models: Agent-Based Modeling as a Tool in Cognitive Sciences. *Current Directions in Psychological Science*, *28*(3), 299–305. https://doi.org/10.1177/0963721419834547

Merdes, C. (2017). Growing Unpopular Norms. *Journal of Artificial Societies and Social Simulation*, *20*(3). https://doi.org/10.18564/jasss.3374

Miller, D. T. (2023). A century of pluralistic ignorance: what we have learned about its origins, forms, and consequences. https://doi.org/10.3389/frsps.2023.1260896

Nail, P. R., Di Domenico, S. I., & MacDonald, G. (2013). Proposal of a double diamond model of social response. *Review of General Psychology*, *17*(1), 1–19. https://doi.org/10.1037/a0030997

Nail, P. R., MacDonald, G., & Levy, D. A. (2000). Proposal of a four-dimensional model of social response. *Psychological Bulletin*, *126*(3), 454–470. https://doi.org/10.1037/0033-2909.126.3.454

Olsson, H., & Galesic, M. (2024, October). Analogies for modeling belief dynamics. https://doi.org/10.1016/j.tics.2024.07.001

Schelling, T. C. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology*, *1*(2), 143–186. https://doi.org/10.1080/0022250X.1971.9989794

Seeme, F., Green, D., & Kopp, C. (2019). Pluralistic Ignorance: A Trade-Off Between Group-Conformity and Cognitive Dissonance. In T. Gedeon, K. W. Wong, & M. Lee (Eds.), *Neural information processing* (pp. 695–706, Vol. 11954). Springer International Publishing. https://doi.org/10.1007/978-3-030-36711-4{_}58

Seeme, F. B., & Green, D. G. (2016). Pluralistic ignorance: Emergence and hypotheses testing in a multi-agent system. *2016 International Joint Conference on Neural Networks (IJCNN)*, 5269–5274. https://doi.org/10.1109/IJCNN.2016.7727896

Shamir, J., & Shamir, M. (1997). *Pluralistic Ignorance Across Issues and Over Time: Information Cues and Biases* (tech. rep. No. 2).

Sparkman, G., Geiger, N., & Weber, E. U. (2022). Americans experience a false social reality by underestimating popular climate policy support by nearly half. *Nature Communications*, *13*(1), 4779. https://doi.org/10.1038/s41467-022-32412-y

Sparkman, G., Ginn, J., Attari, S. Z., & Weber, E. U. (2025). Americans and policymakers underestimate endorsement for the most popular climate solution narrative, combining personal and political action. *Communications Earth & Environment*, *6*(1), 525. https://doi.org/10.1038/s43247-025-02496-7

Sznajd-Weron, K., & Weron, R. (2002). A simple model of price formation. *International Journal of Modern Physics C*, *13*(1). https://doi.org/10.1142/S0129183102003000

Sznajd-Weron, K., Weron, R., & Włoszczowska, M. (2008). Outflow dynamics in modeling oligopoly markets: The case of the mobile telecommunications market in Poland. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(11). https://doi.org/10.1088/1742-5468/2008/11/P11018

Sznajd-Weron, K., Jedrzejewski, A., & Kamińska, B. (2024). Toward Understanding of the Social Hysteresis: Insights From Agent-Based Modeling. *Perspectives on Psychological Science*, *19*(2), 511–521. https://doi.org/10.1177/17456916231195361

Tang, T., Ghorbani, A., & Chorus, C. G. (2022). Hiding opinions by minimizing disclosed information: an obfuscation-based opinion dynamics model. *The Journal of Mathematical Sociology*, *46*(4), 315–341. https://doi.org/10.1080/0022250X.2021.1929968

Taylor, D. G. (1982). Pluralistic Ignorance and the Spiral of Silence: A Formal Analysis. *Public Opinion Quarterly*, *46*(3), 311. https://doi.org/10.1086/268729

Thurner, S., Klimek, P., & Hanel, R. (2018, November). *Introduction to the Theory of Complex Systems* (Vol. 1). Oxford University Press. https://doi.org/10.1093/oso/9780198821939.001.0001

Topuz, E., & Yücel, G. (2024). Analyzing the Emergence and Dynamics of Pluralistic Ignorance with Agent-Based Models. https://doi.org/10.1007/978-3-031-57785-7{_}32

Wenzel, M. (2005). Misperceptions of social norms about tax compliance: From theory to intervention. *Journal of Economic Psychology*, *26*(6), 862–883. https://doi.org/10.1016/j.joep.2005.02.002

Ye, M., Qin, Y., Govaert, A., Anderson, B. D., & Cao, M. (2019). An influence network model to study discrepancies in expressed and private opinions. *Automatica*, *107*, 371–381. https://doi.org/10.1016/j.automatica.2019.05.059