

Feature Review

Analogies for modeling belief dynamics

Henrik Olsson^{1,2,*} and Mirta Galesic^{1,2,3,*}

Belief dynamics has an important role in shaping our responses to natural and societal phenomena, ranging from climate change and pandemics to immigration and conflicts. Researchers often base their models of belief dynamics on analogies to other systems and processes, such as epidemics or ferromagnetism. Similar to other analogies, analogies for belief dynamics can help scientists notice and study properties of belief systems that they would not have noticed otherwise (conceptual mileage). However, forgetting the origins of an analogy may lead to some less appropriate inferences about belief dynamics (conceptual baggage). Here, we review various analogies for modeling belief dynamics, discuss their mileage and baggage, and offer recommendations for using analogies in model development.

Analogies for belief dynamics

Belief dynamics (see [Glossary](#)) is at the core of many challenges that our societies currently face, from misinformation to polarization and violent conflict. They emerge from the interaction between cognitions and **social networks** over time. This interaction has been difficult to study in our siloed scientific environment. Different fields have been focusing on different aspects of the complex sociocognitive system underlying belief dynamics. Psychologists and cognitive scientists tend to focus on understanding cognitive processes within individual minds. Sociologists, economists, and computational social scientists tend to focus on social network structures, while applied mathematicians and statistical physicists focus on developing analytic and computational models without necessarily tying them to empirical data. To understand and model belief dynamics and the underlying complex cognitive–social system, researchers have been using analogies with more familiar phenomena and methods ([Box 1](#)). Some analogies for belief dynamics are popular across different disciplines, such as epidemics, **ferromagnetism**, and thresholds analogies. Others are used more often in particular disciplines, such as forces in psychology, evolution in anthropology, weighted additive models (WADDs) in psychology and sociology, and Bayesian updating in economics.

Recognizing analogies that underlie different **models of belief dynamics** is important because every **analogy** can not only provide conceptual mileage, but also bring some conceptual baggage ([Box 2](#)). Understanding the limits of their analogies would help scientists to develop better models of when and why people's **beliefs** change and anticipate previously unexpected societal trajectories, from vaccine resistance to climate change denial and radicalization. Finding appropriate analogies for belief dynamics can also help improve communication with the public about vulnerabilities in our belief systems that facilitate manipulation or disregard of useful information.

In this review, we describe analogies for belief dynamics in more detail, illustrating how they have inspired sociocognitive elements of different models ([Figure 1](#)) and describing their conceptual mileage and baggage (summarized in [Table 1](#)). We do not aim to provide a comprehensive review of all belief dynamics models or how well they can account for empirical phenomena. There are several excellent articles that do this (e.g., [1–4]). Instead, we provide a conceptual exploration of analogical reasoning used to study belief dynamics.

Highlights

Analogies from other systems provide theoretical insights, modeling tools, and empirical methods for studying belief dynamics. Frequently used analogies include epidemics, ferromagnetism, forces, evolution, thresholds, weighted additive models, and Bayesian learning.

Analogies that inspire models of belief dynamics can provide conceptual mileage by highlighting valuable insights and similarities between the base and target phenomena.

However, they can also introduce conceptual baggage when certain aspects of the analogy are inappropriately extended to the target domain, leading to incorrect conclusions.

Cognitive scientists could advance the study of belief dynamics by considering analogies with other systems, providing translations of analogical concepts into relevant cognitive and social processes, and comparing implications of models based on different analogies with empirical data.

¹Santa Fe Institute, Santa Fe, NM 87501, USA

²Complexity Science Hub, 1080 Vienna, Austria

³Vermont Complex Systems Center, University of Vermont, Burlington, VT 05405, USA

*Correspondence: olsson@csh.ac.at (H. Olsson) and galesic@santafe.edu (M. Galesic).



Box 1. Use of analogies in science

The use of analogies in research on belief dynamics is not unusual: when trying to understand a novel system, scientists often use analogies with phenomena [130,131] and methodological tools they know well [88]. This has been observed across sciences, from astronomy [132] and ecology [133] to economics [134] and cognitive science [135]. The scientific use of analogies is not very different from the way people in daily life use analogies and metaphors from familiar domains to develop and communicate solutions to novel and complex problems [136,137].

Analogies shape the way scientists conceptualize the phenomena they study, formulate questions they are asking, and choose methodological tools. Analogies can provide not only useful insights or ‘conceptual mileage’, but also unnecessary ‘conceptual baggage’ (see Box 2 in the main text). Recognizing the baggage of an analogy and correcting these wrong assumptions can lead to a more nuanced understanding of the phenomenon of interest [138]. For example, the revolving of planets around the Sun has been used as an analogy for the revolving of electrons around the center of the mass of the atom. However, understanding the limits of this analogy helped develop insights into the quantum nature of atomic structure [139].

Belief dynamics models vary significantly, addressing different cognitive and social, individual, and collective aspects, and using different terms for the phenomena they are trying to explain. They also differ in what they call these phenomena (beliefs, opinions, or attitudes). We choose to use the term ‘beliefs’ throughout this article (Box 3). Nevertheless, different models that target different levels and different dimensions of belief dynamics can be inspired by the same analogies and can carry similar mileage and baggage. While we focus on the most frequently used analogies for belief dynamics, some researchers have been using other analogies, combining several analogies, or trying to implement plausible cognitive and social mechanisms directly (Box 4). We provide several recommendations to researchers for how to recognize and use analogies when developing models of belief dynamics (Box 5).

Epidemics

One of the most frequently used analogies for belief dynamics is that of epidemics, whereby the spread of beliefs (target of the analogy) is compared with the spread of disease (base of the analogy). Epidemiological models originally developed to understand the spread of disease, such as the **susceptible–infected–recovered (SIR)** and susceptible–infected–susceptible (SIS) models, have been used as simple descriptions of the spread of beliefs on social networks [5]. Models based on the epidemics analogy typically involve the spread of a single belief at a time, and beliefs are represented as binary variables (a belief is present or absent; Figure 1).

Traditionally, the updating process has been assumed to resemble a simple contagion whereby beliefs have a chance to spread after each exposure to an already ‘infected’ person. In line with the rich literature on the important role of **frequency-dependent social learning** and deciding

Box 2. Conceptual mileage and baggage of analogies

What is a good analogy? As with any analogy, some of the analogies for belief dynamics might be more useful than others [140]. Most theoretical accounts of analogies emphasize the importance of similarity between the target (the novel phenomenon) and the base (the known phenomenon) of an analogy. However, more similarity is not necessarily better. The similarity of features of objects involved in the target and the base of an analogy (feature similarity) is less important than the similarity of relationships between these objects [141–143]. For example, in the analogy ‘A lion is among animals [target] as a king is among humans [base],’ the lion and the king, as well as animals and humans, have relatively few features in common [143]. However, the relationship between lions and animals is similar to the relationship between kings and humans, and the analogy provides insight, or as we say here, conceptual mileage, into the position of lions in a typical food chain.

An important prerequisite for successful use of analogies is understanding their limitations. Given that analogies are ‘borrowed’ from other domains, not all properties of the base phenomenon will correspond to the target phenomenon. Forgetting this can lead to incorrect conclusions and hypotheses about the domain of interest [133]. For example, even though it is apt to say that a lion is like a king, it would be overreaching to expect that lions also wear crowns. This assumption would be a conceptual baggage of the analogy. This points to the importance of carefully mapping relevant properties of the analogy to theoretical constructs and ultimately to specific implementations in computational models (see Box 5 in the main text).

Glossary

Analogy: comparison that aims to explain the relationship between objects in one, typically less well understood domain (the target of the analogy) by the relationship between objects in another, typically better understood domain (the base of the analogy).

Belief dynamics: set of interacting cognitive and social processes leading to change of individual beliefs and their spread on social networks; see also Boxes 3 and 4 in the main text.

Belief measurement: different methodologies for inferring people’s beliefs about different issues, from asking them directly about their specific beliefs using a variety of rating scales, to inferring beliefs from their behaviors and communication traces.

Beliefs: encompass conceptualizations of beliefs used in different fields, from assumptions about states of the world, views and opinions on moral and political issues to cognitive aspects of attitudes; see also Box 3 in the main text.

Deficit model of science communication: theoretical framework that assumes public skepticism or hostility toward science is due to a lack of knowledge, and that providing more information will change attitudes and behaviors.

Ferromagnetism: property of certain materials to exhibit magnetization in the presence of a magnetic field.

Frequency-dependent social learning: strategies for updating beliefs and behaviors based on the number or fraction of others who have a specific belief. Prominent examples are majority, unanimity, and minority strategies.

Ising model: mathematical model of ferromagnetism that describes interactions between spins on a lattice, where each spin can be in one of two states. This model has been used to study the dynamics of beliefs in belief and social networks.

Models of belief dynamics: analytic or computational models that aim to describe and predict changes of beliefs in individual minds and on social networks.

Social networks: way to represent relationships between individuals. Networks can be as simple as a dyad and as complex as multilayer networks describing different social structures.

[6–8], researchers have moved beyond the assumption of simple contagion to allow for a complex contagion process [9,10], whereby individuals need exposure to several ‘infected’ contacts carrying a new belief before adopting it themselves.

The simplicity of epidemiological models enables investigations of belief spread on a variety of network structures, going beyond pairwise interactions toward simplicial complexes [11] and hypergraphs [12], which capture hierarchies and multiway relationships. The epidemics analogy is also useful for modeling joint dynamics of disease and vaccine hesitancy beliefs [13,14], or of disease and fear [15]. Basic epidemiological models can also be extended to include updating of social network links [16] and joint spread of several related beliefs or behaviors [17].

Conceptual mileage

The epidemics analogy brings a lot of conceptual mileage to the study of belief dynamics (Table 1). Consider the recent example of a SIR model of the spread of antivaccine beliefs in social networks with different characteristics [14]. The model starts with a subset of individuals who initially hold these beliefs and share them with their contacts. Everyone has an equal chance of being influenced, making a permanent decision based on the persuasiveness of the beliefs, measured by a factor ranging from 0 to 1. The spread continues through social networks, step by step, only reaching new individuals via those convinced in the previous step. The process ends when no new individuals are convinced. The outcome is a population divided into those who adopt antivaccine beliefs, those exposed but not convinced, and those never exposed. Only individuals convinced by antivaccine beliefs refuse vaccination, while all others are willing to be immunized. Using a SIR-like model on a social network highlights the potential to understand more clearly how network characteristics contribute to belief change in different directions. The results in [14] suggest that even if beliefs against vaccination have minimal persuasiveness, they can

Susceptible–infected–recovered (SIR) models: class of epidemiological models used to describe the spread of infectious diseases within a population, categorizing individuals as susceptible (S), infected (I), or recovered (R). These and similar models (e.g., susceptible–infected–susceptible or SIS models, in which individuals can move between being susceptible and infected repeatedly) have been used to simulate how beliefs spread through social networks.

	Structure									Process									
	Belief format		Belief number and relationships		Social influence		Updating rule			Updating moderators		Goal	Outcome						
	Discrete	Continuous	One belief	Several unrelated beliefs	Several related beliefs	Summary representation	Social networks	Frequency dependence	Averaging	Multiplication	Imitation of a model	Susceptibility/prior beliefs	Threshold	(Felt) dissonance	Consistency	Correspondence	Belief updating	Belief relationships updating	Social network updating
Epidemics																			
Ferromagnetism																			
Thresholds																			
Forces																			
Evolution																			
WADD models																			
Bayesian learning																			

Trends in Cognitive Sciences

Figure 1. Translation of different analogies to sociocognitive elements of belief dynamics models. These elements can be roughly divided into the underlying structure and the process of belief dynamics [36]. Elements of the structure include the representation of beliefs (whether they are discrete or continuous, their number and inter-relationships) and the representation of the social environment (as a summary representation or as social networks). Elements of the process include updating rules specifying how individuals adopt new beliefs (e.g., by choosing the most frequent belief in their social environments, averaging beliefs of different social contacts, multiplying past and new potential beliefs, or by simply imitating a specific model individual); moderating factors that determine whether, and to what extent, the updating rule will be implemented in the first place (e.g., depending on prior beliefs, similarity thresholds, or felt dissonance); the goal of the updating process (typically achieving consistency with other beliefs in one’s mind or in one’s social environment, and sometimes also correspondence with the actual truth); and outcome (only belief updating or also describing the dynamics of the relationships between beliefs and between social contacts). Colored blocks denote the typical translations of structure and process of belief dynamics models inspired by each analogy. Each analogy has its own distinct color. Striped colors mark translations that exist but have been done rarely or only recently. White blocks show translations that are typically not (although in some cases could be) implemented by models guided by a specific analogy. Abbreviation: WADD, weighted additive model.

Table 1. Analogies for belief dynamics

Core analogy	Modeling tools	Conceptual mileage	Conceptual baggage
Epidemics			
People infect each other with beliefs just like they infect each other with diseases	Epidemiological models (SIR, SIS, and similar models)	Some beliefs increase the chance of 'contracting' other beliefs 'Immunization' from harmful beliefs might be possible Repeated exposure to a belief increases likelihood of 'infection'	In its basic form, cannot account for importance of similarity of people 'infecting' each other, occurrence of belief backlash, and phenomenon that 'infection' with some beliefs can strengthen overall belief system
Ferromagnetism			
People reduce dissonance by aligning their beliefs, just like particle systems reduce energy by aligning their spins	Models of interacting particle systems (Ising, voter, and similar models)	Easy mapping of ferromagnetic concepts to belief dynamics: spins as beliefs, energy as dissonance, attention as inverse temperature, couplings as relationships Many aspects can be empirically measured Minimal but powerful models to explain phenomena	Easy mapping might lead to unwarranted generalizations from ferromagnetic models to human sociocognitive systems Relationships between beliefs or between individuals are typically assumed to be fixed Focus can be on analytical solvability rather than explaining real-world belief systems
Forces			
Belief systems are subject to psychological and social forces, just like physical systems are subject to physical forces	Triangle balance	Simple representations of alliances and animosities in belief and social networks Simple algorithms for updating both beliefs and relationships between beliefs and individuals	Triadic closure and balance might not be most crucial aspects of belief dynamics
Thresholds			
Beliefs update only when certain threshold is crossed, just like physical systems change state only when they cross a threshold	Bounded-confidence models, drift diffusion models	Can easily produce patterns of polarization and fragmentation Grounded in actual neural mechanisms	Implausible assumption of thresholds that are stable across individuals, contexts, and time
Evolution			
Beliefs evolve through process of cultural evolution, just like genes evolve through process of genetic evolution	Evolutionary models	Stresses importance of theoretical and empirical comparison of different updating rules Recognizes differences between horizontal and vertical social transmission Highlights relevance of cultural attractors Recognizes relevance of cumulative cultural evolution	Beliefs are not discrete memes that get transmitted unchanged Network structure is rarely explicitly modeled
WADDs			
People update their beliefs by weighing and adding individual and social influences, just like information is integrated in regression models	Regression analysis, matrix algebra	Relatively easy to fit to empirical data Enable estimating relative influence of individual and social factors	Can be cognitively demanding Might overfit because of too many parameters
Bayesian learning			
People update their beliefs just like likelihoods become updated in Bayesian models	Bayes theorem, Bayesian networks	Provides normative standard for belief updating	People might not have capacity to update beliefs according to Bayes theorem

quickly reach a large segment of the population. Additionally, the more centrally positioned individuals in the network, who are more frequently exposed to these views, are more inclined to embrace them.

Another example of conceptual mileage is the recognition that some beliefs can increase the chance of adopting other beliefs, just like some diseases can increase the likelihood of getting another disease. For example, higher trust in government can facilitate positive beliefs about vaccination [18], just like some immune diseases can increase the chances of respiratory infections. Further example of conceptual mileage of the epidemics analogy is the idea that it might be possible to 'immunize' people from harmful beliefs just like it is possible to vaccinate people against

Box 3. Beliefs, opinions, and attitudes

Beliefs have been defined in numerous ways that often overlap with definitions of opinions, attitudes, and related terms, reflecting different ontological pathways to understanding the same or similar phenomena [144]. For example, beliefs can be conceptualized as cognitive aspects of attitudes [145]. The American Psychological Association defines an opinion as ‘an attitude, belief, or judgment’ⁱ, and having opinions is often defined as believing somethingⁱⁱ. The ‘Stanford Encyclopedia of Philosophy’ defines beliefs as propositional attitudes comprising a subject, proposition, and ‘attitude, stance, take, or opinion’ that the subject has about the proposition^{iv}. These examples illustrate that there is considerable conceptual confusion in the literature regarding the definition of beliefs, opinions, and attitudes (e.g., see [146]). Despite different names, the underlying processes are similar, and analogies and models used to study them often overlap (see also ‘belief dynamics’ in the Glossary in the main text). Here, we use the term ‘belief’ broadly to encompass conceptualizations of beliefs used in different fields, from assumptions about states of the world, views and opinions on moral and political issues to cognitive aspects of attitudes.

Belief dynamics is a set of interacting cognitive and social processes leading to the change of individual beliefs and their spread on social networks. This phenomenon is sometimes referred to as ‘opinion dynamics’, although the analogies and models are often similar or identical. For instance, Brandt and Sleegers applied an Ising model to analyze what they term a ‘political belief system’, conceptualizing the nodes as attitudes or opinions [30]. Galesic and Stein used the framework of Ising models to study belief dynamics, defining the nodes as beliefs and framing the process as belief dynamics [34], and van der Maas *et al.* developed a hierarchical Ising model to study opinions as a system of beliefs and behaviors associated with a specific attitudinal object [32].

diseases. This idea has been used to design educational interventions to counter misinformation [19]. Finally, repeated exposure to a belief is often related to a higher likelihood of adopting it [20], similar to how repeated exposure to a virus makes one more likely to get infected by that virus.

Conceptual baggage

Even though the epidemics analogy brings a lot of conceptual mileage to the study of belief dynamics, it can also bring some conceptual baggage and erroneous expectations. Public communication about consumer products, science, and politics often relies on the expectation that exposure and attention to facts supporting a belief will enhance the acceptance of that belief, following the **deficit model of science communication** (reviewed in [21]). However, exposure to beliefs that are radically different from one’s own can sometimes lead to backfire effects, whereby individual beliefs become even less similar than before to the belief they were exposed to [21]. For instance, studies such as the one in [22] have demonstrated that exposing people to scientific facts about vaccination can paradoxically make them less likely to vaccinate. Furthermore, it has been shown empirically that the spread of beliefs is enhanced when people share other relevant beliefs and characteristics [23,24], but the spread of infectious diseases typically does not depend on whether two people share other similar diseases. Finally, a disease typically weakens the organism so that the likelihood of other infections increases. By contrast, being ‘infected’ by some beliefs can ‘strengthen’ the belief system against a whole class of beliefs. For example, believing that ‘many very important things happen in the world, which the public is never informed about’ (and similar items from the Conspiracy Mentality Questionnaire [25]) is strongly related to the support of many different antivaccination arguments [26] as well as to the resistance to government-recommended preventative measures against coronavirus disease 2019 (COVID-19) [27].

Ferromagnetism

Another popular analogy in the field of belief dynamics is ferromagnetism. Models inspired by this analogy assume that people reduce dissonance by aligning their beliefs to fit with the beliefs of other people around them and their own belief systems, similar to how particle systems reduce energy by aligning with each other and their magnetic fields. For example, belief dynamics models inspired by the **Ising model** [28–30] represent beliefs as spins valued -1 or $+1$, positioned in a network of other spins, which can be other beliefs one has or beliefs of other people. The relationships between beliefs are represented by couplings between spins. The cognitive dissonance

stemming from conflicting beliefs is represented by the energy of the system [31], which can be lowered over time by adjusting beliefs to better align with each other. The extent of this adjustment depends on the attention to dissonance, represented as inverse temperature in the Ising model [31,32]. For example, when considering who to vote for, some individuals might seek consistency between their beliefs about a certain candidate and their other personal beliefs, such as about moral and economic issues.

Particle systems have been studied in statistical physics using other models beyond Ising and similar models. The most prominent among them is the voter model [33], whereby individuals adopt the belief of a randomly drawn neighbor.

Models of belief dynamics inspired by ferromagnetism can include further psychologically plausible assumptions. For example, while the original Ising model assumes that spins update to align with the majority of other neighboring spins, it is possible to implement other updating rules, such as random copying and following specific individuals (e.g., experts [34]). The model can also be extended to represent beliefs with more than two discrete values [35], and to represent additional internal and external influences as internal and external magnetic fields (Figure 1) [36], and both cognitive and social networks [32,37]. The ferromagnetism analogy has been used to develop many other variations of belief dynamics models [1].

Conceptual mileage

Concepts of spins, energy, couplings, temperature, and spin-updating rules map well to the concepts of belief, dissonance, network links, attention to dissonance, and belief updating in actual human belief dynamics [31]. This allows researchers to design empirical measures of different aspects of these models. For example, Galesic and Stein evaluated the ability of Ising-like models with different belief-updating rules to account for the dynamics of belief change in political and health-related beliefs in real-world populations [34].

These simple models can describe a range of phenomena observed in the real world, including radicalization, polarization, and clustering of opinions [1,38], and can be exactly solvable for certain social network structures, such as regular lattices. As such, they can be a rich source of inspiration to social scientists when developing minimal computational models that explain different patterns with a very limited set of parameters.

Conceptual baggage

The minimalistic representation of belief dynamics as updating spin systems omits much of the richness of human experience, such as the effect of emotions, memory processes, and various individual differences. This is also the case for other analogies, but the mapping of ferromagnetic systems of particles to human cognitive–social systems is relatively straightforward and can mask the many ways in which these systems are not similar. This can increase the chances for generalizations that might not hold for human sociocognitive systems. For example, belief updating might not always be related to dissonance reduction [39], and attention to dissonance might not always be needed for belief change. In addition, in ferromagnetic systems, the relationships between spins are often fixed in time or quenched, although they can evolve over time in so-called ‘annealing systems’ [40].

The main conceptual baggage of simple ferromagnetic models is that they often predict consensus and have difficulty reproducing other empirically demonstrated outcomes, such as disagreement and polarization. Therefore, numerous extensions of these models have been proposed, including representations of stubborn and confident individuals, diverse social network structures, other updating rules, such as majority rule, beliefs with three states, and the coevolution

of both beliefs and network links [38,41,42]. However, models in this framework typically remain confined to the dynamics of a single belief with discrete values, with the main outcome being belief rather than network updating.

In addition, the drive for mathematical elegance and theoretical closure in these models sometimes leads to a focus on solving the model rather than solving the problem. This can result in beautifully constructed models that operate well under theoretical conditions but falter in practical scenarios where the unpredictability of human behavior and the irregularity of social interactions come into play.

Thresholds

In many physical and biological systems, significant changes occur only when a certain threshold is crossed. For example, water tends to freeze when its temperature drops to 0°C, and neurons fire only when their membrane potential depolarizes to around -55 mV. These thresholds exemplify a broader principle of emergent collective effects observed at the macro level, where systems experience phase transitions, shifting from one dynamic regime (e.g., consensus) to another (e.g., polarization).

In a similar way, some belief dynamics models assume a certain threshold for updating to occur (Figure 1). For example, in so-called 'bounded confidence models', individuals update their beliefs toward each other only if the difference between them is not too large [43–46]. In drift-diffusion models (DDM), which are often used to model preferences, the concept of threshold also has an important role. In recently proposed social DDM models [47], it is assumed that each individual initially accumulates personal information about the state of the world. The accumulated evidence acts as a starting point for the social phase of the process, in which individuals can gather additional social information. The decision is made once an individual has collected enough evidence, surpassing the threshold.

These models lend themselves to various extensions, for example, allowing for the fact that some individuals have more opportunities for updating their beliefs [48], or that individuals update either their beliefs or their social networks [49,50], or that the communication of beliefs is noisy [51,52]. Beliefs in threshold-based models are typically represented as continuous variables, but the analogy has also been applied to categorical belief variables. For example, in Axelrod's model of cultural dissemination, individuals copy each other's beliefs if they have at least some beliefs (or more generally, features) in common [53].

Conceptual mileage

Models inspired by the thresholds analogy can straightforwardly produce patterns of fragmentation and polarization of societies. To illustrate this, consider the updating process in bounded confidence models. Assume that each individual in a collective has a belief on a scale from 0 to 1. If the threshold is set at, for example, 0.2, this means that that an individual will update their belief toward another individual only if their beliefs differ by no more than 0.2. Setting a low threshold typically results in the population dividing into separate clusters, whereas a higher threshold promotes consensus. As the threshold decreases, belief clusters become smaller and more numerous. The initial distribution of beliefs and the threshold levels influence this dynamic, potentially leading to various forms of polarization [2]. Furthermore, the mechanism suggested by this analogy has been strongly grounded in neuroscience findings, where a perceptual decision occurs when the firing rate of a specific cortical neural population hits a certain threshold [54].

Conceptual baggage

The assumption of threshold can be a mere redescription rather than an explanation of the observed behavior. Thresholds are likely highly dependent on situational factors and on the specific

belief [43], but many models specify them exogenously and assume they are stable across contexts. This can obscure the actual mechanisms that give rise, endogenously, to different thresholds for the same individuals in various situations. For instance, a person might readily update their beliefs in a supportive group setting but resist change when faced with opposition or during a crisis. Recent developments in the field have attempted to address these limitations. For example, models of bounded confidence now include mechanisms for threshold adaptation [55], acknowledging that individuals might adjust their thresholds based on new information, changing circumstances, or shifts in the social context.

Forces

Several classes of model are based on the analogy with physical forces. In these models, often inspired by Gestalt psychology, beliefs and their relationships are embedded in psychological fields and are affected by various psychological and social forces, similar to how physical objects are embedded in phase spaces defined by different forces that affect them, influencing their formation, strength, and interaction with other beliefs [56]. In these models, a psychological field is a construct that represents the space in which beliefs exist and interact. Each belief can be thought of as a point or vector within this field, having qualities, such as direction (representing attitude or stance) and magnitude (representing the strength or conviction of the belief). The field is dynamic, constantly influenced by internal and external changes, similar to electromagnetic or gravitational fields in physics (Figure 1).

A well-known analogy stemming from this field-theoretical approach is the analogy of balance, proposed by Heider [57]: When beliefs are imbalanced, the system will experience tension, and forces will arise to make beliefs and/or the relationships between them balanced. Heider elaborated on systems of three relations (triads), and the idea has been subsequently generalized to larger cycles [58]. Consider a triad of individuals, let's say Alice, Bob, and Carol, whose relationships can be positive (+) or negative (-). The relationships in the triad are deemed balanced if the product of their signs is positive. For example, if Alice likes Bob (+), Bob likes Carol (+), and Alice likes Carol (+), the triad is balanced. Conversely, a triad where Alice likes Bob (+), Bob likes Carol (+), but Alice dislikes Carol (-) is imbalanced. To reduce the tension in the imbalanced case, Alice could change her feelings about Carol to positive, or one of the other relationships could flip from positive to negative. The analogy of balance provides a simple quantitative framework for intuitive notions, such as 'a friend of my friend is my friend' and 'an enemy of my enemy is my friend'. As such, it has been used over decades and across different fields to understand the dynamics of social interactions and processes leading to the establishment of cooperative or conflict relationships, as well as fragmentation and polarization of societies (for recent examples, see [59–63]).

Another well-known model based on the analogy with physical forces is the social impact theory of Latané [64], who proposed that beliefs change under the combined influence of several distinct social forces: the strength of social contacts in terms of their status or similar characteristics, their immediacy or closeness in space or time, and the number of social contacts present. Implemented in a computational model within dynamic social impact theory [65], these forces can produce a variety of patterns observed in the real world. The theory has been used to illuminate the belief dynamics involved in the US capitol riots on January 6, 2021 [66].

Conceptual mileage

The analogy brings conceptual mileage by providing a simple representation of alliances and animosities as well as simple algorithms for modeling their changes. It can be applied to many different types of network, from pure belief or social networks to their various combinations. Furthermore, this framework can be extended to belief networks, offering a model for analyzing

how beliefs and social networks evolve over time. For example, Pham *et al.* present a model of a process that can account for the evolution of balance in a social network, as well as the evolution of a range of beliefs [61]. In this model, everyone begins with a binary belief and one individual is randomly selected to potentially change their belief based on the ‘stress’ from conflicting beliefs within their social circle. If flipping the belief significantly reduces this stress, the belief changes. Simultaneously, the relationship between each pair of neighbors is evaluated, determining a value that is positive if they agree and negative if they disagree. This dual assessment allows for the evolution of personal beliefs and interpersonal balance at the same time.

Conceptual baggage

Belief and social systems often contain various patterns beyond triangles, and while triadic relationships can often be inferred from the existing links, this does not mean that they exist or will be formed. In addition, people might not be aware of relationships beyond their own dyadic connections, and recent research suggests that similar societal-level patterns can occur by only dyadic updating [61]. Alternatively, people might be seeking balance in larger structures beyond triadic [58,67]. Finally, many other factors beyond balance might influence the belief dynamics [68]. For example, information integration in the form of summation and averaging can produce belief change [39]. In this view, people are not trying to achieve balance, but might simply be influenced by people around them without ever considering dissonance or balance.

Evolution

A rich body of literature uses the analogy of the evolutionary process to understand the spread of beliefs and practices over longer periods of time and across populations [7,69]. Similar to genetic evolution, evolution of cultural traits, such as beliefs, can be seen as driven by three principles [70]: diversity of cultural traits across individuals and groups (equivalent to the principle of ‘variation’ in the language of genetic evolution); differences in their persistence and spread (‘differential fitness’); and transmission with probability related to their cultural fitness (or ‘principle of inheritance’).

Conceptual mileage

The evolution analogy helps to recognize that different social learning strategies, such as conformism [7] and status-based learning [71], produce different patterns of belief dynamics [8,72]. The evolution analogy also helps us to understand the differences between horizontal and vertical transmission of beliefs and other cultural traits that is, spread within and between generations; [73], as well as the importance of cultural attractors and related cognitive biases [74,75]. It has also inspired research on the evolution of covert and overt signaling strategies, and the coevolution of beliefs and covert signals [76]. Its conceptual mileage further includes the concepts of cumulative cultural evolution, whereby culturally acquired traits (which might also include beliefs) develop over generations [77]. Finally, models built within this analogy have stimulated experimental and other empirical research on social learning and belief updating (for recent examples, see [78–81]), as well as tournaments to directly compare the success of different strategies for social learning [82].

Conceptual baggage

The conceptual baggage of the analogy includes the unwarranted conceptualization of beliefs as discrete memes that get transmitted unchanged, or that cultural evolution should bring societal progress [70]. This view can be traced back to older studies that somewhat exaggerated the similarities between genes and memes. For instance, Dawkins [83] implied that memes replicate and spread through populations in a manner akin to biological genes. However, this comparison can be misleading because memes, unlike genes, often change and adapt in their transmission due to the influence of complex social and individual factors (e.g., [84]). They are not bound by the rigid structures of biological replication and can morph in response to feedback from their cultural

environment. Therefore, while the meme analogy has been useful in drawing parallels between biological and cultural evolution, it oversimplifies the multifaceted nature of how beliefs change over time. Furthermore, unlike other analogies, the evolution analogy does not have a well-developed representation of different plausible social network structures. In models inspired by this analogy, it has often been tacitly assumed that a population is a well-connected graph. However, this has been changing, with empirical data showing the value of intermediate levels of connectedness in social networks [85] and the unique human multigroup social structure [86], and simulations showing how different social network structures can evolve due to cultural selection [87].

Weighted additive models

This and the next analogy are somewhat different than the preceding ones. Instead of known systems or system properties to explain belief dynamics, the base of these analogies are methodological tools developed in other areas of cognitive and social sciences, and now used as approximations of sociocognitive processes underlying belief dynamics. The existence of these analogies is in line with the observation that various methodological tools, such as hypothesis-testing algorithms and analysis of variance, have been used before to explain cognitive phenomena [88].

Weighted additive models, often implemented as linear regression formulations, have been used widely in social sciences to estimate underlying relationships in the data. They have also been proposed as actual cognitive mechanisms through which people perceive and react to the world. Examples include additive utility models, such as prospect theory [89,90], additive versions of information integration theory that have been applied to a variety of perceptual and cognitive tasks [91], the lens model framework for studying how individuals use information from the task environment to make judgments [92,93], and several other judgment and decision models (e.g., [94]).

In belief dynamics, such models have been used to describe the way in which people integrate their own beliefs and the beliefs of others. In the Theory of Reasoned Action [95], and its later extension, the Theory of Planned Behavior [96,97], own attitudes and perceived social norms are combined as a weighted average (with the addition of perceived behavioral control in the latter theory) to predict behavioral intention. In the French–DeGroot model [98,99], individuals adopt beliefs that are an average of their own beliefs and the beliefs of each of their social contacts, weighted by their influence. In the Friedkin–Johnsen model [100], the equation also includes individual stubbornness or a desire to keep one’s initial belief, allowing for richer belief dynamics that do not always converge to a consensus for connected networks. These models are suited for representing the dynamics of one or more continuous beliefs, and they describe how these beliefs get updated. The relationships between beliefs, as well as relationships between individuals, are typically fixed (Figure 1).

Conceptual mileage

Weighted additive models can be used to model the dynamics not only in social networks, but also in belief networks [101]. They can fit the empirical data well [102] and enable estimation of the relative influence of individual and social factors. They have been used to analyze real-world phenomena from filter bubbles [103] to signaling private opinions [104]. They have been related to the more general psychological process of social sampling to create agent-based models of belief updating where utility of different beliefs is a weighted average of utilities from aligning to social environments and to own attitudes [105].

Conceptual baggage

Weighting and averaging of beliefs of all social contacts, each with their separate weight, might be cognitively unrealistic for representing actual human processing. People might instead use more

simplified, heuristic summaries to navigate and represent the complexities of their social environments efficiently [106]. This critique of the weighted additive models aligns with broader discussions in cognitive science, where it has been shown that simpler heuristic approaches can sometimes outperform the more computation-intensive strategies of weighted averaging in tasks related to judgment and decision-making (e.g., [107,108]). Nevertheless, the weighted average approach can also be executed as a more cognitively feasible controlled serial process, wherein a repetitive cycle unfolds in real time, with each subsequent cycle refining a prior estimate into a new one based on newly considered evidence [102,109].

Bayesian learning

Bayes [110] developed his theorem to compute conditional probabilities of an event based on prior beliefs and new evidence. The theorem, refined by Laplace decades after Bayes' death, has been used to estimate the likelihood of different outcomes in numerous applications from engineering and computer science to genetics and medicine. In cognitive science, it has been used to model human cognition [111]. Although the appropriateness of viewing the mind as Bayesian is frequently discussed [112], Bayesian calculus remains a frequently used analogy for cognitive processes.

In belief dynamics, Bayesian learning has been used to model how people update their beliefs upon receiving new information (e.g., [113]). These models represent how people should optimally update their beliefs about an issue after receiving new information about it. This information can include the beliefs of their social contacts or payoffs of different options (Figure 1).

Another way in which Bayesian calculus has been introduced to belief dynamics is through Bayesian networks, which specify causal relationships between different beliefs. While they originated in artificial intelligence [114,115] as a tool for modeling causality, Bayesian networks have been widely adopted by cognitive scientists to analyze and explain phenomena from reasoning

Box 4. Beyond simple analogies

Some models of belief dynamics use more than one analogy. For example, some bounded confidence models combine the threshold analogy with the analogy of forces by introducing negative influence mechanisms [147] that enable descriptions of the dynamics of attraction and repulsion [147,148]. Threshold models can also be reformulated in terms of weighted averaging, unveiling an underlying similarity between models originally conceived in different theoretical frameworks [149,150]. Furthermore, threshold models can be seen as a special variant of nonlinear voter models [151], which are in turn directly related to contagion, discussed under the epidemics analogy. Threshold models, in which thresholds are defined as the number or proportion of others who must adopt a certain belief before an individual does so (e.g., [149]), can be viewed as related to the neural analogy of frequency-dependent belief change. Another example of an approach that includes a diverse set of analogies is a recent model of cultural change and diversity that implements mechanisms of cultural drift, a process of belief consistency, as well as leniency contract, where the majority attends to a dissenting in-group minority opinion [152].

Beyond those described in the main text, other analogies have been used to model belief dynamics. One example is percolation, where a belief seeps through a society like a liquid through a substance [153,154]. Another is quantum probability theory, which has been used to explain different cognitive phenomena [120,155], including context effects in **belief measurement** [156]. The analogy of networks is present in many models in addition to other analogies, whereby the relationships between beliefs and between individuals can be represented as physical networks of objects. Analogies with other dynamical systems have been used to provide insights into nonlinear dynamics and bifurcations of beliefs [157,158], and real-world polarization [159].

Some approaches to belief dynamics aim to model it by directly implementing plausible assumptions about cognitive and social processes [160–165]. One strand of such models has been developed within the connectionist or neural network approach [166–168]. For example, the Attitudes as Constraint Satisfaction model [166] has separate banks of nodes for cognitive representations and for persuasive communications, allowing for modeling of the impact of external messages represented on the persuasion units. Another example is Van Overwalle and Siebler's model that is based on an auto-associative recurrent network focused on assimilating new information [168].

about causal learning in children [116] to belief polarization [117]. In belief dynamics, Bayesian networks have been used to understand intuitive theories people have about vaccination [118], and climate change [119] and to shed light on the dysfunctional disagreement that occurs when individuals are unable to converge on similar beliefs, even with repeated iterations and evidence [120].

Conceptual mileage

The Bayesian learning analogy enables explorations of empirical patterns that could be expected if people are optimally updating their beliefs. By pointing out diverse belief patterns that can occur even with optimal reasoning, such analyses show what can be explained without assuming any specific cognitive biases or deficient reasoning. For example, models assuming Bayesian learning

Box 5. Using analogies to develop models of belief dynamics

Analogies can be a useful tool for building models of belief dynamics, provided that researchers remain aware of the differences between the base system and the belief dynamics system. It is expected that analogies will eventually 'break'. The key is to identify when this occurs so that researchers can adapt, or combine different analogies as needed. Here, we provide several guidelines for a fruitful use of analogies in model building.

Map analogical constructs to theoretical constructs

A first step is to identify and map analogical constructs to theoretically grounded components of belief dynamics. These include structural components, such as networks of individual beliefs and social contacts, as well as process components, including different cognitive and social processes that act on that structure [36]. For some analogies, this mapping is easier than for others. For example, for the ferromagnetic analogy, there are clear mappings between the properties of the analogy and constructs in a belief dynamic system, including the mappings spins→beliefs or individuals, couplings→influence between beliefs or individuals, and energy→dissonance between different beliefs or individuals. For other analogies, this can be more challenging, especially when the goal is to model the dynamics of both the network of internal beliefs and the network of social relationships. For example, the epidemics analogy fares well as a base for modeling the spread of a single belief. However, mapping a classic SIR-like process to the dynamics of internal belief networks is more challenging, as internal beliefs do not 'infect' but rather 'affect' each other such that, if one belief changes in one direction, a connected belief might change in either the same or opposite direction. Thus, care must be taken to specify the relevant mappings of the analogy to theoretical constructs.

Implement theoretical constructs in quantitative models

To derive precise predictions of theories based on different analogies, and to be able to compare these theories, it is useful to implement theoretical constructs into quantitative models of belief dynamics. There are many ways in which this can be done. For example, if a belief dynamics model assumes that the influence between individuals is affected by their similarity, that similarity can be calculated in different ways, including the absolute, square, or other distance between the beliefs and/or other traits of those individuals. Models can be developed to describe belief dynamics on individual and/or collective levels, focusing more on cognitive or social aspects of the underlying processes. The choice will depend on the theoretical and empirical considerations and will shape the mileage and baggage of an analogy. To illustrate, the classical SIR model implementation of the epidemics analogy disregards individual differences and specific real-world network structures, and models belief dynamics using deterministic differential equations at the group level. The same analogy, and the same SIR process, can also be implemented on a social network in an agent-based framework that allows for investigation of the effects of both diverse individual-level processes and network characteristics on the spread of beliefs. When choosing the modeling framework, researchers might need to trade off rigorous mathematical results about aggregate behaviors against a richer understanding of processes at the individual level.

Conduct empirical tests and comparisons of quantitative models

Models of belief dynamics are rarely compared with each other or to empirical data, leading to a limited understanding of which models best capture real-world belief dynamics. Often, these models are developed at an abstract level that is not conducive to empirical measurement, or they go untested because they originate from fields that traditionally do not engage in empirical studies. Ideally, every variable and parameter in a model should be directly observable or inferable from empirical data and should have a plausible psychological or sociological interpretation that reflects factors known to influence belief dynamics. Empirical data could come from surveys, laboratory experiments, social media data, and other sources. Models should be able to predict at least rough patterns in these data, and should be compared with plausible alternative models, with model complexity taken into account to avoid overfitting.

of uncertain evidence can produce polarization patterns [117,121]. In the model presented in [121], individuals apply Bayes' rule iteratively: they first use it to interpret individual pieces of information and then again to update their beliefs based on the sequence of these interpreted signals. This process of double updating results in confirmation bias, potentially causing individuals who receive identical information to develop polarized views. In [122], the authors show how reaching a consensus about climate change can be slowed down or impossible when a minority of scientists is contrarian to the evidence, and when the general public observes the scientific debate. The model in [123] shows how echo chambers can occur even when individuals are Bayesian optimal.

Conceptual baggage

The Bayesian analogy may be understood as implicitly assuming that people actually compute Bayesian calculations when updating their beliefs. In reaction to this assumption, largely perceived as unrealistic (e.g., [112]), cognitively simpler models have been proposed [99,124,125]. Populations might include a mixture of Bayesian agents and agents that use various other updating rules [113]. Note, however, that according to one interpretation of Bayesian learning, the updating of beliefs does not imply explicit Bayesian computation, instead it can be construed as an implicit process resulting from the brain computing probabilities given prior knowledge and exposure. Even so, in practice, the models in this approach have explicit Bayesian calculations, even if they are described as 'as-if' models.

Concluding remarks

Analogies are a useful tool for understanding the complex interaction of cognitions and social environments that give rise to belief dynamics. The analogies we reviewed are very different, although they are all used to study the same system of sociocognitive structures and processes (Figure 1). Some frequently used analogies, including epidemics, ferromagnetism, and thresholds, have been developed outside cognitive science, while others have stronger origins in empirical findings of human cognition and sociality (e.g., forces, evolution, weighted additive models, and Bayesian learning).

Independently of their origin, some analogies map well onto different aspects of the underlying sociocognitive systems (e.g., ferromagnetism), while others might be less realistic but still provide useful normative standards for evaluating real-world patterns of belief dynamics (e.g., weighted additive models and Bayesian learning). Analogies also differ in the number of assumptions, or parameters, they introduce, with some analogies (e.g., epidemics or thresholds) being simpler in that respect than others (e.g., weighted additive models or Bayesian networks). Finally, models based on some analogies are easier to translate into empirically measurable constructs that enable model testing and comparison. For example, key parameters of analogies based on ferromagnetism and forces can be measured directly using survey questions (e.g., dissonance and temperature; [36]), while parameters such as thresholds and susceptibility (in thresholds and epidemics analogies, respectively) can be inferred from the data *post hoc* but their direct measures have not yet been well developed.

Each of the analogies brings significant conceptual mileage. For example, they help researchers to appreciate that complex patterns, such as polarization, clustering, fragmentation, and minority influence, can be produced by simple models with only a few parameters, such as those inspired by ferromagnetism and epidemics. Moreover, many analogies introduce useful methodologies for representing and simulating belief dynamics, such as energy minimization, cumulative evolution, structural balance, or Bayesian calculus. However, all analogies also carry conceptual baggage, and the key to the successful use of analogies is to recognize this baggage and change or combine analogies to avoid it. A common issue with many analogies for belief dynamics, particularly those involving thresholds and evolution, is the assumption that beliefs spread independently of other

Outstanding questions

How do we map different plausible analogies for belief dynamics on the underlying cognitive and social processes they aim to represent?

How can we recognize when an analogy is breaking and introducing more conceptual baggage than mileage?

How can we combine analogies to model different sociocognitive aspects and timescales of belief dynamics?

How can we compare predictions of models driven by different analogies to empirical data?

How can we effectively use analogies to communicate scientific findings about belief dynamics to the general public and policy makers?

beliefs (Figure 1). This view has contributed to the widespread deficit model of science communication, which assumes that communicating scientific facts is sufficient to change people's beliefs. Nevertheless, it is now well understood that other beliefs one has, as well as one's social environment, affect the likelihood of belief updating [21]. Another example of baggage is the implicit assumption of the epidemics analogy, that any contact increases the probability of 'contracting' a new belief, while, in reality, belief updating also depends on other beliefs people have, properties of the source, and so on. This assumption can lead to an assumption that, with repetition, almost any message will take hold in the population. Sadly, this is not true for many scientific messages. Yet another frequent baggage of many analogies reviewed here is the assumption that beliefs change but the relationships between them, as well as the underlying social network structure, stay the same (Figure 1). However, the relationship between beliefs and people who share them does change over time, complicating the prediction of the resulting belief dynamics.

One aspect of belief dynamics that we have not addressed here is the dynamic relationship between beliefs and behavior. Typically, behavior is not part of belief dynamics models, despite the clear importance of the relationship between beliefs and behavior. An exception is the theory of planned behavior, which integrates personal beliefs, perceived social norms, and perceived behavioral control to predict behavioral intentions and, ultimately, actual behavior. Meta-analyses have shown that behavior change interventions based on the theory of planned behavior are indeed effective [126].

How can we construct a good analogy for belief dynamics? Aside from satisfying a few basic principles, such as one-to-one correspondence between relationships in the base and the target of the analogy, reducing extraneous associations, and avoiding mixed analogies [127], our review suggests that analogies with almost any other well-known system or methodological tool can illuminate some aspects of belief dynamics. Accordingly, studies of belief dynamics have been using various other analogies beyond those we focused on here, and combinations of several analogies are also common (Box 4). More perspectives on the same phenomenon are useful as long as they have realistic assumptions and are rigorously compared with each other and empirical data [128,129]. Cognitive scientists could significantly advance the study of belief dynamics by considering many different analogies, providing translations of their core concepts into existing cognitive and social processes, developing quantitative models based on these analogies, and comparing their theoretical implications and empirical predictions (see Outstanding questions and Box 5).

Acknowledgments

This work has been supported by NSF grants BCS-1918490, DRMS-1949432, MMS- 2019982, and OIA- 2242829.

Declaration of interests

The authors declare no competing interests.

Resources

ⁱ<https://dictionary.apa.org/opinion>

ⁱⁱwww.merriam-webster.com/dictionary/opinion

ⁱⁱⁱwww.oed.com/dictionary/opinion

^{iv}<https://plato.stanford.edu/archives/spr2024/entries/belief/>

References

1. Castellano, C. *et al.* (2009) Statistical physics of social dynamics. *Rev. Mod. Phys.* 81, 591–646
2. Flache, A. *et al.* (2017) Models of social influence: towards the next frontiers. *J. Artif. Soc. Soc. Simul.* 20, 2
3. Sirbu, A. *et al.* (2017) Opinion dynamics: models, extensions and external effects. In *Participatory Sensing, Opinions and Collective Awareness* (Loreto, V. *et al.*, eds), pp. 363–401, Springer

4. Will, M. *et al.* (2020) Combining social network analysis and agent-based modelling to explore dynamics of human interaction: a review. *Socio-Environ. Syst. Model.* 2, 16325
5. Bettencourt, L.M. *et al.* (2006) The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models. *Physica A* 364, 513–536
6. Condorcet, M. (1785) *Essay on the Application of Analysis to the Probability of Majority Decisions*, Imprimerie Royale
7. Boyd, R. and Richerson, P.J. (1985) *Culture and the Evolutionary Process*, University of Chicago Press
8. Kendal, R.L. *et al.* (2018) Social learning strategies: bridge-building between fields. *Trends Cogn. Sci.* 22, 651–665
9. Guilbeault, D. *et al.* (2018) Complex contagions: a decade in review. In *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks* (Lehmann, S. and Ahn, Y.Y., eds), pp. 3–25, Springer
10. Guilbeault, D. and Centola, D. (2021) Topological measures for identifying and predicting the spread of complex contagions. *Nat. Commun.* 12, 4430
11. Iacopini, I. *et al.* (2019) Simplicial models of social contagion. *Nat. Commun.* 10, 2485
12. Landry, N.W. and Restrepo, J.G. (2020) The effect of heterogeneity on hypergraph contagion models. *Chaos* 30, 103117
13. Cooney, D.B. *et al.* (2022) Social dilemmas of sociality due to beneficial and costly contagion. *PLoS Comput. Biol.* 18, e1010670
14. Prieto Curiel, R. and González Ramírez, H. (2021) Vaccination strategies against COVID-19 and the diffusion of anti-vaccination views. *Sci. Rep.* 11, 6626
15. Epstein, J.M. *et al.* (2021) Triple contagion: a two-fears epidemic model. *J. R. Soc. Interface* 18, 20210186
16. Gross, T. *et al.* (2006) Epidemic dynamics on an adaptive network. *Phys. Rev. Lett.* 96, 208701
17. Wang, W. *et al.* (2019) Coevolution spreading in complex networks. *Phys. Rep.* 820, 1–51
18. Schmelz, K. and Bowles, S. (2021) Overcoming COVID-19 vaccination resistance when alternative policies affect the dynamics of conformism, social norms, and crowding out. *Proc. Natl. Acad. Sci. USA* 118, e2101221118
19. Lewandowsky, S. and Van Der Linden, S. (2021) Countering misinformation and fake news through inoculation and prebunking. *Eur. Rev. Soc. Psychol.* 32, 348–384
20. Weaver, K. *et al.* (2007) Inferring the popularity of an opinion from its familiarity: a repetitive voice can sound like a chorus. *J. Pers. Soc. Psychol.* 92, 821–833
21. Ecker, U.K. *et al.* (2022) The psychological drivers of misinformation belief and its resistance to correction. *Nat. Rev. Psychol.* 1, 13–29
22. Nyhan, B. *et al.* (2014) Effective messages in vaccine promotion: a randomized trial. *Pediatrics* 133, e835–e842
23. Goldstone, R.L. *et al.* (2013) Learning along with others. *Psychol. Learn. Motiv.* 58, 1–45
24. Cialdini, R.B. and Goldstein, N.J. (2004) Social influence: compliance and conformity. *Annu. Rev. Psychol.* 55, 591–621
25. Bruder, M. *et al.* (2013) Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy Mentality Questionnaire. *Front. Psychol.* 4, 225
26. Holford, D.L. *et al.* (2023) Psychological profiles of anti-vaccination argument endorsement. *Sci. Rep.* 13, 11219
27. Soveri, A. *et al.* (2021) Unwillingness to engage in behaviors that protect against COVID-19: the role of conspiracy beliefs, trust, and endorsement of complementary and alternative medicine. *BMC Public Health* 21, 684
28. Galam, S. (2008) Sociophysics: a review of Galam models. *Int. J. Mod. Phys. C* 19, 409–440
29. Dalege, J. *et al.* (2016) Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychol. Rev.* 123, 2–22
30. Brandt, M.J. and Sleegers, W.W. (2021) Evaluating belief system networks as a theory of political belief system dynamics. *Personal. Soc. Psychol. Rev.* 25, 159–185
31. Dalege, J. *et al.* (2018) The attitudinal entropy (AE) framework as a general theory of individual attitudes. *Psychol. Inq.* 29, 175–193
32. van der Maas, H.L. *et al.* (2020) The polarization within and across individuals: the hierarchical Ising opinion model. *J. Complex Netw.* 8, cnaa010
33. Holley, R.A. and Liggett, T.M. (1975) Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* 3, 643–663
34. Galesic, M. and Stein, D.L. (2019) Statistical physics models of belief dynamics: theory and empirical tests. *Physica A* 519, 275–294
35. Dalege, J. and van der Does, T. (2022) Using a cognitive network model of moral and social beliefs to explain belief change. *Sci. Adv.* 8, eabm0137
36. Galesic, M. *et al.* (2021) Integrating social and cognitive aspects of belief dynamics: towards a unifying framework. *J. R. Soc. Interface* 18, 20200967
37. Dalege, J.; *et al.*, Networks of beliefs: an integrative theory of individual- and social-level belief dynamics. *Psychol. Rev.* In Press.
38. Jędrzejewski, A. and Sznajd-Weron, K. (2019) Statistical physics of opinion formation: is it a SPOOF? *C. R. Phys.* 20, 244–261
39. Anderson, N.H. (1971) Integration theory and attitude change. *Psychol. Rev.* 78, 171–206
40. Jędrzejewski, A. and Sznajd-Weron, K. (2017) Person-situation debate revisited: phase transitions with quenched and annealed disorders. *Entropy* 19, 415
41. Jędrzejewski, A. *et al.* (2020) Spontaneous symmetry breaking of active phase in coevolving nonlinear voter model. *Phys. Rev. E* 102, 042313
42. Redner, S. (2019) Reality-inspired voter models: a mini-review. *C. R. Phys.* 20, 275–292
43. Deffuant, G. *et al.* (2000) Mixing beliefs among interacting agents. *Adv. Complex Syst.* 3, 87–98
44. Hegselmann, R. and Krause, U. (2002) Opinion dynamics and bounded confidence models, analysis, and simulation. *J. Artif. Soc. Soc. Simul.* 5, 3
45. Krause, U. (2000) A discrete nonlinear and non-autonomous model of consensus formation. In *Communications in Difference Equations* (Elaydi, S.N. *et al.*, eds), pp. 227–236, Gordon and Breach
46. Weisbuch, G. *et al.* (2002) Meet, discuss, and segregate! *Complexity* 7, 55–63
47. Tump, A.N. *et al.* (2022) Avoiding costly mistakes in groups: the evolution of error management in collective decision making. *PLoS Comput. Biol.* 18, e1010442
48. Li, G.J. and Porter, M.A. (2023) Bounded-confidence model of opinion dynamics with heterogeneous node-activity levels. *Phys. Rev. Res.* 5, 023179
49. Kozma, B. and Barrat, A. (2008) Consensus formation on adaptive networks. *Phys. Rev. E* 77, 016102
50. Kan, U. *et al.* (2023) An adaptive bounded-confidence model of opinion dynamics on networks. *J. Complex Netw.* 11, 415–444
51. Kurahashi-Nakamura, T. *et al.* (2016) Robust clustering in generalized bounded confidence models. *J. Artif. Soc. Soc. Simul.* 19, 7
52. Klemm, K. *et al.* (2003) Global culture: a noise-induced transition in finite systems. *Phys. Rev. E* 67, 045101
53. Axelrod, R. (1997) The dissemination of culture: a model with local convergence and global polarization. *J. Confl. Resolut.* 41, 203–226
54. Gold, J.I. and Shadlen, M.N. (2007) The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574
55. Bernardo, C. *et al.* (2021) Heterogeneous opinion dynamics with confidence thresholds adaptation. *IEEE Trans. Control Netw. Syst.* 9, 1068–1079
56. Lewin, K. (1943) Defining the field at a given time. *Psychol. Rev.* 50, 292
57. Heider, F. (1946) Attitudes and cognitive organization. *J. Psychol.* 21, 107–112
58. Cartwright, D. and Harary, F. (1956) Structural balance: a generalization of Heider's theory. *Psychol. Rev.* 63, 277–293
59. Górski, P.J. *et al.* (2020) Homophily based on few attributes can impede structural balance. *Phys. Rev. Lett.* 125, 078302
60. Pham, T.M. *et al.* (2020) The effect of social balance on social fragmentation. *J. R. Soc. Interface* 17, 20200752
61. Pham, T.M. *et al.* (2022) Empirical social triad statistics can be explained with dyadic homophilic interactions. *Proc. Natl. Acad. Sci. USA* 119, e2121103119

62. Rodriguez, N. *et al.* (2016) Collective dynamics of belief evolution under cognitive coherence and social conformity. *PLoS ONE* 11, e0165910
63. Schweighofer, S. *et al.* (2020) A weighted balance model of opinion hyperpolarization. *J. Artif. Soc. Soc. Simul.* 23, 5
64. Latané, B. (1981) The psychology of social impact. *Am. Psychol.* 36, 343–356
65. Nowak, A. *et al.* (1990) From private attitude to public opinion: a dynamic theory of social impact. *Psychol. Rev.* 97, 362–376
66. Harton, H.C. *et al.* (2022) 'I'll be there with you': social influence and cultural emergence at the capitol on January 6. *Group Dyn. Theory Res. Pract.* 26, 220–238
67. Kargaran, A. *et al.* (2020) Quartic balance theory: global minimum with imbalanced triangles. *Phys. Rev. E* 102, 012310
68. Yap, J. and Harrigan, N. (2015) Why does everybody hate me? Balance, status, and homophily: the triumvirate of signed tie formation. *Soc. Networks* 40, 103–122
69. Richerson, P.J. and Boyd, R. (2008) *Not by Genes Alone: How Culture Transformed Human Evolution*, University of Chicago Press
70. Mesoudi, A. (2016) Cultural evolution: a review of theory, findings and controversies. *Evol. Biol.* 43, 481–497
71. Henrich, J.P. and Gil-White, F.J. (2001) The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.* 22, 165–196
72. Kandler, A. and Powell, A. (2018) Generative inference for cultural evolution. *Philos. Trans. R. Soc. B* 373, 20170056
73. Anderson, K.A. and Creanza, N. (2023) A cultural evolutionary model of the interaction between parental beliefs and behaviors, with applications to vaccine hesitancy. *Theor. Popul. Biol.* 152, 23–38
74. Sperber, D. (1996) *Explaining Culture: A Naturalistic Approach*, Oxford University Press
75. Mesoudi, A. (2021) Cultural selection and biased transformation: two dynamics of cultural evolution. *Philos. Trans. R. Soc. B* 376, 20200053
76. Smaldino, P.E. *et al.* (2018) The evolution of covert signaling. *Sci. Rep.* 8, 4905
77. Mesoudi, A. and Thornton, A. (2018) What is cumulative cultural evolution? *Proc. R. Soc. B* 285, 20180712
78. Derex, M. and Boyd, R. (2016) Partial connectivity increases cultural accumulation within groups. *Proc. Natl. Acad. Sci.* 113, 2982–2987
79. Mercier, H. *et al.* (2018) Willingness to transmit and the spread of pseudoscientific beliefs. *Appl. Cogn. Psychol.* 32, 499–505
80. Morgan, T.J. *et al.* (2022) The experimental evolution of human culture: flexibility, fidelity and environmental instability. *Proc. R. Soc. B* 289, 20221614
81. Toyokawa, W. *et al.* (2019) Social learning strategies regulate the wisdom and madness of interactive crowds. *Nat. Hum. Behav.* 3, 183–193
82. Rendell, L. *et al.* (2010) Why copy others? Insights from the social learning strategies tournament. *Science* 328, 208–213
83. Dawkins, R. (1976) *The Selfish Gene*, Oxford University Press
84. Atran, S. (2001) The trouble with memes: inference versus imitation in cultural creation. *Hum. Nat.* 12, 351–381
85. Derex, M. *et al.* (2018) Divide and conquer: intermediate levels of population fragmentation maximize cultural accumulation. *Philos. Trans. R. Soc. B* 373, 20170062
86. Hill, K.R. *et al.* (2014) Hunter-gatherer inter-band interaction rates: implications for cumulative culture. *PLoS ONE* 9, e102806
87. Smolla, M. and Akçay, E. (2019) Cultural selection shapes network structure. *Sci. Adv.* 5, eaaw0609
88. Gigerenzer, G. (1991) From tools to theories: a heuristic of discovery in cognitive psychology. *Psychol. Rev.* 98, 254–267
89. Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291
90. Tversky, A. and Kahneman, D. (1992) Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain.* 5, 297–323
91. Anderson, N.H. (1981) *Foundations of Information Integration Theory*, Academic Press
92. Hammond, K.R. *et al.* (1964) Analyzing the components of clinical inference. *Psychol. Rev.* 71, 438–456
93. Hirsch, C.J. *et al.* (1964) Some methodological considerations in multiple-cue probability studies. *Psychol. Rev.* 71, 42–60
94. Juslin, P. *et al.* (2003) Exemplar effects in categorization and multiple-cue judgment. *J. Exp. Psychol. Gen.* 132, 133–156
95. Fishbein, M. and Ajzen, I. (1975) *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Addison-Wesley
96. Ajzen, I. (1991) The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211
97. Fishbein, M. and Ajzen, I. (2010) *Predicting and Changing Behavior: The Reasoned Action Approach*, Psychology Press
98. French Jr., J.R. (1956) A formal theory of social power. *Psychol. Rev.* 63, 181–194
99. DeGroot, M. (1974) Reaching a consensus. *J. Am. Stat. Assoc.* 69, 118–121
100. Friedkin, N. and Johnsen, E. (1990) Social influence and opinions. *J. Math. Sociol.* 15, 193–205
101. Friedkin, N.E. *et al.* (2016) Network science on belief system dynamics under logic constraints. *Science* 354, 321–326
102. Friedkin, N.E. and Bullo, F. (2017) How truth wins in opinion dynamics along issue sequences. *Proc. Natl. Acad. Sci. USA* 114, 11380–11385
103. Chitra, U. and Musco, C. (2020) Analyzing the impact of filter bubbles on social network polarization. In *WSDM '20: Proceedings of the 13th International Conference on Web Search and Data Mining* (Caverlee, J. *et al.*, eds), pp. 115–123, ACM
104. Ye, M. *et al.* (2019) An influence network model to study discrepancies in expressed and private opinions. *Automatica* 107, 371–381
105. Brown, G.D. *et al.* (2022) Social sampling and expressed attitudes: authenticity preference and social extremeness aversion lead to social norm effects and polarization. *Psychol. Rev.* 129, 18–48
106. Hertwig, R., Hoffrage, U., eds (2013) *Simple Heuristics in A Social World*, Press, Oxford University
107. Gigerenzer, G. and Brighton, H. (2009) Homo heuristics: why biased minds make better inferences. *Top. Cogn. Sci.* 1, 107–143
108. Gigerenzer, G. *et al.* (1999) *Simple Heuristics that Make Us Smart*, Oxford University Press
109. Juslin, P. *et al.* (2008) Information integration in multiple cue judgment: a division of labor hypothesis. *Cognition* 106, 259–298
110. Bayes, T. (1763) An essay towards solving a problem in the doctrine of chances. *Philos. Trans. R. Soc. Lond.* 53, 370–418
111. Chater, N. *et al.* (2006) Probabilistic models of cognition: conceptual foundations. *Trends Cogn. Sci.* 10, 287–291
112. Jones, M. and Love, B.C. (2011) Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behav. Brain Sci.* 34, 169–188
113. Acemoglu, D. and Ozdaglar, A. (2011) Opinion dynamics and learning in social networks. *Dyn. Games Appl.* 1, 3–49
114. Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of plausible inference*, Morgan Kaufmann
115. Pearl, J. (2000) *Causality: Models*, Cambridge University Press, Reasoning and Inference
116. Gopnik, A. *et al.* (2004) A theory of causal learning in children: causal maps and Bayes nets. *Psychol. Rev.* 111, 3–32
117. Jern, A. *et al.* (2014) Belief polarization is not always irrational. *Psychol. Rev.* 121, 206–224
118. Powell, D. *et al.* (2023) Modeling and leveraging intuitive theories to improve vaccine attitudes. *J. Exp. Psychol. Gen.* 152, 1379–1395
119. Cook, J. and Lewandowsky, S. (2016) Rational irrationality: modeling climate change belief polarization using Bayesian networks. *Top. Cogn. Sci.* 8, 160–179
120. Pothos, E.M. *et al.* (2021) Information overload for (bounded) rational agents. *Proc. R. Soc. B* 288, 20202957
121. Fryer Jr., R.G. *et al.* (2019) Updating beliefs when evidence is open to interpretation: implications for bias and polarization. *J. Eur. Econ. Assoc.* 17, 1470–1501
122. Lewandowsky, S. *et al.* (2019) Influence and seepage: an evidence-resistant minority can affect public opinion and scientific belief formation. *Cognition* 188, 124–139
123. Madsen, J.K. *et al.* (2018) Large networks of rational agents form persistent echo chambers. *Sci. Rep.* 8, 12391

124. Golub, B. and Jackson, M.O. (2010) Naive learning in social networks and the wisdom of crowds. *Am. Econ. J. Microecon.* 2, 112–149
125. Molavi, P. *et al.* (2018) A theory of non-Bayesian social learning. *Econometrica* 86, 445–490
126. Steinmetz, H. *et al.* (2016) How effective are behavior change interventions based on the Theory of Planned Behavior? *Z. Psychol.* 224, 216–233
127. Gentner, D. and Jeziorski, M. (1993) The shift from metaphor to analogy in Western science. In *Metaphor and Thought* (2nd edn) (Ortony, A., ed.), pp. 447–480, Cambridge University Press
128. Gyllingberg, L. *et al.* (2023) The lost art of mathematical modeling. *Math. Biosci.* 362, 109033
129. Smaldino, P.E. (2017) Models are stupid, and we need more of them. In *Computational Social Psychology* (Vallacher, R.R. *et al.*, eds), pp. 311–331
130. Hesse, M.B. (1966) *Models and Analogies in Science*, Notre Dame University Press
131. Kuhn, T. (1979) Metaphor in science. In *Metaphor and Thought* (Ortony, A., ed.), pp. 409–419, Cambridge University Press
132. Gentner, D. *et al.* (1997) Analogical reasoning and conceptual change: a case study of Johannes Kepler. *J. Learn. Sci.* 6, 3–40
133. Olson, M.E. *et al.* (2019) A user's guide to metaphors in ecology and evolution. *Trends Ecol. Evol.* 34, 605–615
134. Morgan, M. (2012) *The World in the Model: How Economists Work and Think*, Cambridge University Press
135. Gigerenzer, G. and Goldstein, D.G. (1996) Mind as computer: birth of a metaphor. *Creat. Res. J.* 9, 131–144
136. Brand, C.O. *et al.* (2021) Analogy as a catalyst for cumulative cultural evolution. *Trends Cogn. Sci.* 25, 450–461
137. Gentner, D. *et al.*, eds (2001) *The Analogical Mind: Perspectives from Cognitive Science*, MIT Press
138. Oppenheimer, R. (1956) Analogy in science. *Am. Psychol.* 11, 127–135
139. Arabatzis, T. (2019) *Representing Electrons: A Biographical Approach to Theoretical Entities*, University of Chicago Press
140. Simon, H.A. (1962) The architecture of complexity. *Proc. Am. Philos. Soc.* 106, 467–482
141. Gentner, D. (1983) Structure-mapping: a theoretical framework for analogy. *Cogn. Sci.* 7, 155–170
142. Tourangeau, R. and Sternberg, R. (1981) Aptness in metaphor. *Cogn. Psychol.* 13, 27–55
143. Tourangeau, R. and Sternberg, R. (1982) Understanding and appreciating metaphors. *Cognition* 11, 203–244
144. Dubova, M. and Goldstone, R.L. (2023) Carving joints into nature: reengineering scientific concepts in light of concept-laden evidence. *Trends Cogn. Sci.* 27, 656–670
145. Ajzen, I. (2001) Nature and operation of attitudes. *Annu. Rev. Psychol.* 52, 27–58
146. Bergman, M.M. (1998) A theoretical note on the differences between attitudes, opinions, and values. *Swiss Polit. Sci. Rev.* 4, 81–93
147. Jager, W. and Amblard, F. (2005) Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. *Comput. Math. Organ. Theory* 10, 295–303
148. Axelrod, R. *et al.* (2021) Preventing extreme polarization of political attitudes. *Proc. Natl. Acad. Sci. USA* 118, e2102139118
149. Granovetter, M. (1978) Threshold models of collective behavior. *Am. J. Sociol.* 83, 1420–1443
150. Lopez-Pintado, D. and Watts, D.J. (2008) Social influence, binary decisions and collective dynamics. *Ration. Soc.* 20, 399–443
151. Schweitzer, F. and Behera, L. (2009) Nonlinear voter models: the transition from invasion to coexistence. *Eur. Phys. J. B* 67, 301–318
152. Jung, J. *et al.* (2021) Cultural drift, indirect minority influence, network structure, and their impacts on cultural change and diversity. *Am. Psychol.* 76, 1039–1053
153. Solomon, S. *et al.* (2000) Social percolation models. *Physica A* 277, 239–247
154. Schweitzer, F. (2021) Social percolation revisited: from 2D lattices to adaptive networks. *Physica A* 570, 125687
155. Pothos, E.M. and Busemeyer, J.R. (2022) Quantum cognition. *Annu. Rev. Psychol.* 73, 749–778
156. Wang, Z. *et al.* (2014) Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl. Acad. Sci. USA* 111, 9431–9436
157. Bizyaeva, A. *et al.* (2023) Nonlinear opinion dynamics with tunable sensitivity. *IEEE Trans. Autom. Control* 68, 1415–1430
158. Yang, V.C. *et al.* (2021) Dynamical system model predicts when social learners impair collective performance. *Proc. Natl. Acad. Sci. USA* 118, e2106292118
159. Leonard, N.E. *et al.* (2021) The nonlinear feedback dynamics of asymmetric political polarization. *Proc. Natl. Acad. Sci. USA* 118, e2102149118
160. Banisch, S. and Shamon, H. (2023) Biased processing and opinion polarization: experimental refinement of argument communication theory in the context of the energy debate. *Sociol. Methods Res.*, Published online July 24, 2023. <https://doi.org/10.1177/00491241231186658>
161. Duggins, P. (2017) A psychologically-motivated model of opinion change with applications to American politics. *J. Artif. Soc. Soc. Simul.* 20, 1–13
162. Epstein, J.M. (2014) *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science*, Princeton University Press
163. Lorenz, J. *et al.* (2021) Individual attitude change and societal dynamics: computational experiments with psychological theories. *Psychol. Rev.* 128, 623–642
164. Miller, J.H. and Page, S.E. (2007) *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, Princeton University Press
165. Smith, E.R. and Conrey, F.R. (2007) Agent-based modeling: a new approach for theory building in social psychology. *Personal. Soc. Psychol. Rev.* 11, 87–104
166. Monroe, B.M. and Read, S.J. (2008) A general connectionist model of attitude structure and change: the ACS (Attitudes as Constraint Satisfaction) model. *Psychol. Rev.* 115, 733–759
167. Van Overwalle, F. and Heylighen, F. (2006) Talking nets: a multiagent connectionist approach to communication and trust between individuals. *Psychol. Rev.* 113, 606–627
168. Van Overwalle, F. and Siebler, F. (2005) A connectionist model of attitude formation and change. *Personal. Soc. Psychol. Rev.* 9, 231–274