

Social norm dynamics in a behavioral epidemic model

Christos Charalambous^[0000–0002–0752–1007]

Department of Economics, University of Cyprus,
PO Box 20537, 1678 Nicosia, Cyprus
`charalambous.christos.2@ucy.ac.cy`

Abstract. Understanding the social determinants of preventive behavior is vital for epidemic modelling and effective policy making. Traditional models emphasize imitation or rational trade-offs, but recent evidence highlights the role of social norms. We develop a behavioral epidemic model of seasonal disease on multilayer networks, where vaccination decisions combine learning from experience with coevolving social norms. The framework distinguishes descriptive norms (what others do) from injunctive norms (what others think ought to be done), while incorporating cognitive dissonance, social projection and logical consistency. Simulations show that norm dynamics yield markedly different vaccination uptake and infection levels compared to considering solely payoff-driven learning. Injunctive norms exert stronger and more persistent effects than descriptive norms. Interventions targeting injunctive expectations significantly affect outcomes, while those on descriptive norms can be weaker or even counterproductive. Norm-based models, once empirically validated, can better capture human behavior and guide strategies for collective action problems even beyond pandemics.

Keywords: First keyword · Second keyword · Another keyword.

1 Introduction

Vaccine hesitancy is one of the most severe major global health threats [1], made more pressing by the experience of COVID-19 [2] and the persistence of seasonal influenza. As with many collective-action problems, formal regulation and mandates are not always sufficient, feasible, or timely on their own [3]. This has motivated growing interest in behavioural approaches that act through communication, trusted authorities, peer influence, and public messaging to shape perceived social norms and encourage cooperation [4, 21].

Models that couple disease dynamics with human behaviour describe how infection and protective actions influence one another [5–7], while game-theoretic approaches account for strategic vaccination decisions made under bounded rationality [8–11]. However, these models usually capture social influence in overly narrow ways—most often as simple imitation—thereby missing important psychological mechanisms such as cognitive dissonance [12], social projection [13],

and higher-order expectations [14]. These mechanisms shape how individuals interpret social information, form intentions, and respond to perceived expectations from others. To reflect real decision-making, models need to distinguish between descriptive norms—what others do—and injunctive norms—what others think should be done [15, 16]. When these two coincide, compliance and cooperation are typically much stronger [16, 17].

Empirical evidence consistently shows that pro-vaccination social norms lead to higher vaccine uptake across different diseases [18–20], and public-health strategies increasingly focus on interventions that harness these norms [21]. In most modelling approaches, descriptive norms are introduced through social learning processes [22, 23], while injunctive norms—anchored in moral and reputational expectations—have received much less attention [24–26], even though they are crucial for sustaining long-term behavioural change. Only a limited number of modelling approaches have explicitly integrated material incentives with norm-sensitive behaviour, especially in vaccination settings [24, 27, 28]. More recent psychologically grounded formulations offer a firmer foundation for doing so by treating actions, personal norms, and expectations as coevolving state variables [29].

Most behavioural–epidemic models represent social influence through imitation or by averaging payoffs, treating norms as fixed factors that modify utility rather than as evolving psychological states [8–10]. In contrast, the framework introduced here distinguishes between personal norms, empirical expectations, and normative expectations, linking their joint evolution to experience-weighted learning. This approach moves beyond imitation and payoff-based mechanisms, allowing us to explore how different norm-related variables—personal norms, empirical expectations, and normative expectations—coevolve with epidemic risk.

An expanding line of research is now developing cognitively grounded, empirically informed agent-based models (ABMs) to investigate how norms emerge, exert downward causation, and shape long-term intervention outcomes [30–35]. In their “transmission game,” Woike et al. [36] found that communicating injunctive norms significantly reduced risk-taking, while merely providing descriptive information or case counts had little effect. Building on this evidence, we introduce an ABM that integrates a behavioral epidemic model with realistic social-norm dynamics [29], following the approach of Realpe-Gómez et al. [33], to explore how vaccination intentions and norms evolve over successive seasonal outbreaks. By “norm dynamics” we refer specifically to the coupled evolution of three agent-level variables: personal attitudes y_i , empirical expectations \tilde{x}_i , and normative expectations \tilde{y}_i , each of which changes over time in response to behaviour, social observation, and external influence.

We first describe the model—its decision process, Experience-Weighted Attraction (EWA) learning [37], and norm dynamics—and then analyse how social norms influence infection and vaccination patterns, including cases of stubborn adherence and external interventions. The framework is suitable for future laboratory validation [38].

Our contributions are fourfold: (i) we couple epidemic and norm dynamics through adaptive learning weights [39–41]; (ii) we integrate cognitive dissonance, social projection, and logical consistency into norm evolution [24, 29]; (iii) we show that injunctive norms exert stronger long-term effects on vaccination than descriptive ones, consistent with experimental findings [36, 42]; and (iv) we model external interventions acting directly on norms rather than merely on information flows [43–45].

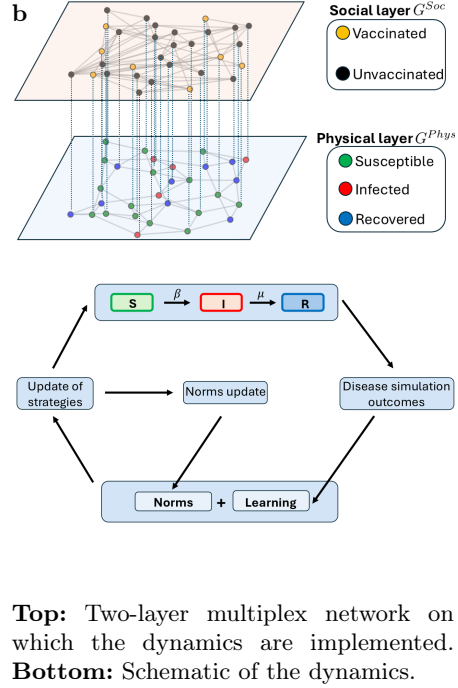
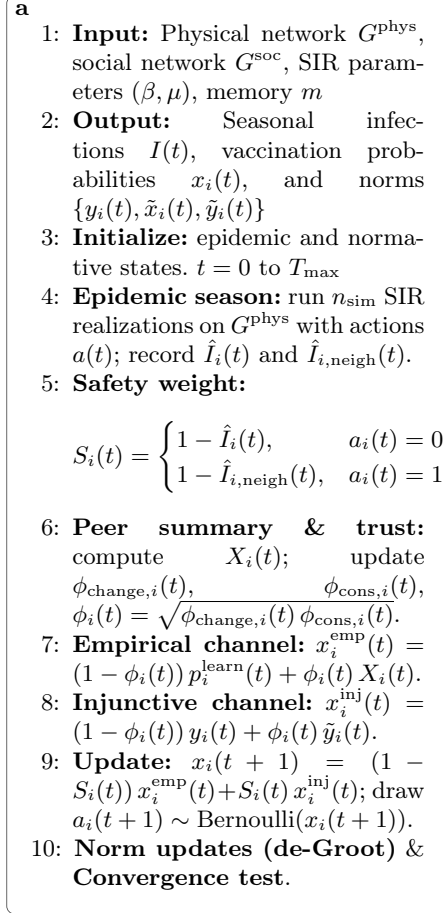
2 Model

The model evolves on two partially overlapping networks. The *physical layer* represents contacts through which infection can spread. The *social layer* represents observation and influence: agents observe neighbours’ vaccination behaviour on this layer and update their intentions and norm-related variables accordingly. Some links can belong to both layers, capturing the fact that people often receive social information from individuals with whom they also interact physically.

In each season, agents first decide whether to vaccinate. Conditional on these actions, the epidemic unfolds on the physical layer according to a Susceptible–Infected–Recovered (SIR) process. To estimate the risk associated with the current seasonal vaccination profile, we run n_{sim} independent SIR realizations on the *same* physical network and with the *same* vaccination configuration. These repeated realizations provide estimates of each agent’s own infection probability and of the infection exposure in its physical neighbourhood. After these epidemic realizations are completed, agents update their vaccination intentions and their norm-related variables; only then does the next season begin. Thus, agents do not revise behaviour continuously during an outbreak, but only between seasons. The present formulation is most appropriate for recurrent infectious diseases in which protective decisions are revised between epidemic waves or seasons, such as seasonal influenza or other diseases with repeated campaigns and repeated behavioural adjustment. Replacing SIR with another compartmental process such as SIRD would not alter the behavioural architecture of the model; it would only modify how epidemic outcomes enter the perceived-risk terms and payoff estimates.

Unless otherwise stated, the physical layer follows a small-world topology [46] with mean degree 6 and rewiring probability 0.1, whereas the social layer follows a Klimek–Thurner network [47, 48] with parameters $c = 0.58$, $r = 0.12$, and $m = 1$. Results were qualitatively robust under alternative topologies, including the case in which both layers are Erdős–Rényi networks. The two-layer architecture is illustrated in Fig. 1.

To model behavioral adaptation, we employ a game-theoretic framework in which vaccination payoffs depend on others’ choices. Decisions are influenced not only by epidemiological risk but also by personal beliefs and social norms. Following [29], each individual i is characterized by (i) a personal attitude y_i (belief about the appropriate behavior), (ii) an empirical expectation \tilde{x}_i (belief about others’ intentions), and (iii) a normative expectation \tilde{y}_i (belief about others’ attitudes).



Seasonal update of intentions and norms.

Fig. 1: Dynamics' algorithm, Two-layer network and schematic of dynamics. **a.:** Algorithm illustrating the system dynamics. **b. Top:** Vaccination game on a two-layer network. The physical layer hosts a seasonal SIR epidemic, while the social layer governs vaccination decisions based on information from both layers. The physical layer follows a small-world topology and the social layer a Klimek–Thurner network, with partial overlap reflecting that many physical contacts also serve as information links. **Bottom:** Schematic of the algorithm. Each season, the SIR dynamics run on the physical layer to estimate infection probabilities. Agents then decide whether to vaccinate, considering both payoff-based learning and normative factors, update their norms accordingly, and begin the next season.

Model overview

The model combines a seasonal epidemic process with a behavioural updating process on a multiplex network. In each season, agents first choose whether to vaccinate. Given these choices, a seasonal outbreak unfolds on the physical contact layer according to a compartmental epidemic model. Repeating the epidemic simulation multiple times for the same seasonal configuration provides each agent with estimated infection risks. These risks, together with past material outcomes and social information from the social layer, determine the probability that the agent vaccinates in the next season. Social norms do not enter as a single scalar factor. Instead, each agent is characterized by three distinct variables: a personal attitude y_i , an empirical expectation \tilde{x}_i about others' behaviour, and a normative expectation \tilde{y}_i about what others think should be done. These three variables are updated every season and jointly shape subsequent vaccination decisions.

At a high level, one season consists of four steps:

1. agents choose vaccination actions;
2. the epidemic unfolds on the physical network and generates infection-risk estimates;
3. agents update vaccination intentions using both payoff-based learning and norm-related variables;
4. personal attitudes and expectations are updated on the social network.

This overview is intended to separate the logic of the model from the mathematical details introduced below. Before introducing the equations, we also summarize the behavioural variables. For each agent i , $x_i(t)$ denotes the vaccination intention, i.e. the probability of vaccinating in season t ; $a_i(t) \in \{0, 1\}$ denotes the realized action, with $a_i(t) = 1$ meaning vaccination; $y_i(t)$ denotes the agent's personal attitude about whether vaccination is the appropriate behaviour; $\tilde{x}_i(t)$ denotes the agent's empirical expectation about how likely others are to vaccinate; and $\tilde{y}_i(t)$ denotes the agent's normative expectation about how strongly others approve vaccination. These quantities are conceptually distinct: $x_i(t)$ is a behavioural propensity, whereas $y_i(t)$, $\tilde{x}_i(t)$, and $\tilde{y}_i(t)$ are norm-related state variables that help determine it. For readability, we first show how these variables enter the vaccination decision, and only afterwards specify their own update equations.

2.1 The decision making process

We define the dynamics of an agent's vaccination intention $x_i(t)$, i.e. the probability of choosing vaccination, which we assume to be driven by both material and normative considerations. Empirical evidence suggests that communities under risk rely more on experience and descriptive norms than on moral beliefs or injunctive norms [40, 41]. In the absence of precise data on how perceived risk shapes this balance, we assume a simple linear relation between risk and the weight placed on empirical factors, following standard practice in norm-utility

models [44]. Therefore we split the contribution to the intention $x_i(t+1)$ update as follows:

$$x_i(t+1) = (1 - S_i(t)) x_i^{empirical}(t) + S_i(t) x_i^{injunctive}(t) \quad (1)$$

which depends on the safety parameter $S_i(t)$ that quantifies the safety an agent feels and takes values $0 < S_i(t) < 1$. The safety is assumed to be given by the following formula

$$S_i(t) = \begin{cases} 1 - \widehat{I}_i(t), & \text{if agent is unvaccinated,} \\ 1 - \widehat{I}_{i,\text{neighbors}}(t), & \text{if agent is vaccinated.} \end{cases} \quad (2)$$

where $\widehat{I}_i(t) = \frac{n_i^{\text{inf}}(t)}{n_{\text{sim}}}$ is the estimated probability that agent i becomes infected during season t across the n_{sim} independent SIR realizations performed on the physical network for that season, and $n_i^{\text{inf}}(t)$ is the number of those realizations in which agent i becomes infected. Likewise, $\widehat{I}_{i,\text{neighbors}}(t)$ denotes the average fraction of i 's neighbors in the *physical* network that become infected across the same n_{sim} realizations. Thus, both quantities are physical-layer risk estimates computed after actions are fixed and before intentions are updated for the next season.

We define the contribution from what is or what agents think is actually happening, i.e. the empirical contribution to the intention, as

$$x_i^{empirical}(t) = (1 - \phi_i(t)) p_i^{\text{learn}}(t) + \phi_i(t) \tilde{x}_i(t) \quad (3)$$

where we see contributions from the agent's learning of material payoffs $p_i^{\text{learn}}(t)$ (see next subsection, $0 \leq p_i^{\text{learn}}(t) \leq 1$) and what the agent thinks that the rest of the community actually chose on average $\tilde{x}_i(t)$ (empirical expectations, $0 \leq \tilde{x}_i(t) \leq 1$). Their relative weight depends on $\phi_i(t)$, which determines whether decisions rely more on personal experience or on social information

$$\phi_i(t) = \sqrt{\phi_i^{\text{change}}(t) \phi_i^{\text{consensus}}(t)}. \quad (4)$$

$\phi_i^{\text{change}}(t)$ is the change-detector function and is defined as in [39]. $\phi_i^{\text{consensus}}(t) \in [0, 1]$ is the consensus function given by

$$\phi_i^{\text{consensus}}(t) = |2 * X_i - 1| \quad (5)$$

where X_i is the average fraction of neighbors vaccinated as observed by the focal individual, and hence $\phi_i^{\text{consensus}}(t)$ is equal to 1 when neighbors agree and 0 if only half are vaccinated.

Intuitively, $\phi_{\text{change},i}$ measures how volatile neighbours' behaviour has been compared to the recent past, whereas $\phi_i^{\text{consensus}}(t)$, captures the degree of agreement among them. Using the geometric mean ensures that social information becomes influential only when both conditions are met—behaviour is stable and neighbours largely agree—whereas in noisy or polarized environments agents fall back on their own experience and attitudes.

The contribution from the belief dependent part of the agents, which is the mechanism activated when the agent is not at risk takes the following form

$$x_i^{injunctive}(t) = (1 - \phi_i(t)) y_i(t) + \phi_i(t) \tilde{y}_i(t) \quad (6)$$

where again the idea is to split the contributions between the effect of own beliefs and the beliefs of others using $\phi_i(t)$. We assume $0 < \tilde{y}_i(t), y_i(t) < 1$. The dynamics of the norm variables, $y_i(t)$ and $\tilde{y}_i(t)$ are also defined in the next subsections. A diagram with the contributions to the total intention is shown in Fig. 2.

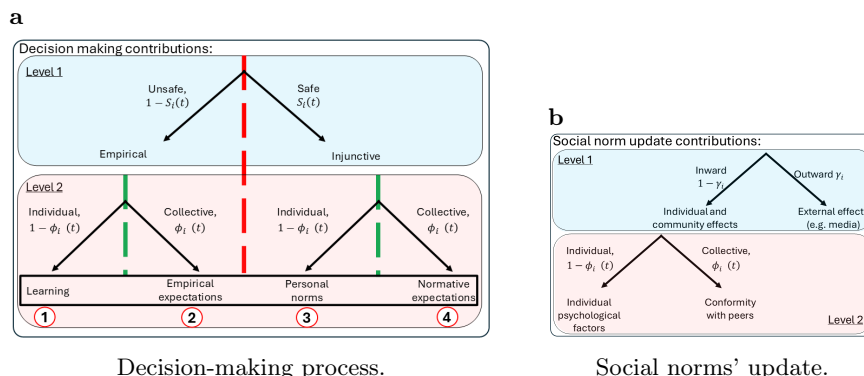


Fig. 2: Decision-making process and social norm dynamics. a.: Schematic of the decision-making process. Vaccination intention combines learning (1), personal norms (3), and social norms (2) and (4), each weighted by its relative influence. Agents rely more on empirical factors when they feel unsafe (low $S_i(t)$) or when their environment is unstable and heterogeneous, and more on normative factors under stable conditions. **b.: Schematic of social-norm dynamics.** Each agent updates norms based on external cues (with probability γ_i), personal beliefs, and peers' actions. In unstable and highly divided environments, agents rely more on internal factors than on social conformity.

We employ one of the simplest mechanisms to model whether individuals rely on personal or collective information. We assume that when individuals perceive instability in their neighbors' behavior and lack of consensus, they rely more on their own judgment, drawing on personal experience or internalized norms. This simplification omits other possible influences, such as the degree of trust in neighbors' ability to make sound decisions under changing epidemiological conditions, which could depend on factors like disease severity and would substantially complicate the model.

Numerous social and behavioral theories explain how people adapt and change behavior [51]. Building on the Continuous Opinions and Discrete Actions (CODA) framework [52], our approach links individual actions to attitudes shaped by observed behavior, consistent with social-psychological models that emphasize cog-

nitive rules over strict rationality in decision-making [53]. A similar perspective was used in [44]. Earlier mathematical models often treated attitudes as static, overlooking their feedback with behavior and thereby misrepresenting adoption dynamics and policy impact.

2.2 The experience dependent mechanism

Human behaviour combines model-free and model-based reinforcement learning [37], integrating past experience with forward-looking reasoning. The Experience-Weighted Attraction (EWA) model captures both [54], unifying reinforcement [55] and belief learning [56] by weighting realized and forgone payoffs equally. Forgone payoffs are inferred from neighbours' average infection rates, since individuals lack direct access to others' outcomes [57]. Agents evaluate payoffs over the last m cycles with a safety-dependent memory decay. EWA captures how individuals integrate past experience with expectations about others. This empirically validated approach [33] provides a cognitively grounded basis for modelling social adaptation and norm compliance.

The payoffs are therefore estimated as following:

$$\Pi_i^{Unvac}(t) = \begin{cases} 1 - c_I \widehat{I}_{i,neighbors}(t) & \text{if vaccinated} \\ 1 - c_I \widehat{I}_i(t) & \text{if not vaccinated} \end{cases} \quad (7)$$

$$\Pi_i^{Vac}(t) = 1 - c_V$$

where $c_I > 0$ is the perceived cost of infection and $c_V > 0$ is the perceived cost of vaccination. In the payoff normalization used here, avoiding infection yields payoff 1, infection lowers payoff by c_I , and vaccination lowers payoff by c_V . The materially-motivated intention of the agent i at time t to vaccinate or not, is assumed to be given by the commonly employed Quantal Response Equilibrium (QRE) for binary choices, which is simply given by the logistic (softmax) function

$$p_i^{learn}(t) = \frac{1}{1 + e^{-\frac{\pi_i^{Unvac}(t) - \pi_i^{Vac}(t)}{\kappa}}} \quad (8)$$

Here $\pi_i^{Vac}(t)$ is the average payoff for the vaccinating option over the last m rounds, while $\pi_i^{Unvac}(t)$ is the average payoff received for not vaccinating over the last m rounds $\pi_i^{Unvac}(t)$, given by

$$\pi_i^{Unvac}(t) = \sum_{j=0}^m \frac{(S_i(t))^j}{\sum_{n=0}^m (S_i(t))^n} \Pi_i^{Unvac}(t-j) \quad (9)$$

where the safety parameter $S_i(t)$ defined above plays the role of a memory decaying function.

2.3 Social norm dynamics

To incorporate social norm dynamics into the EWA framework, we separate material payoffs from normative factors. Experiments show that individuals represent others' preferences and beliefs independently of their own [42, 58]. Accordingly, we model personal attitudes, empirical expectations, and normative expectations as distinct state variables rather than a single imitation process. We are interested in the dynamics of the personal norms y_i , normative expectations \tilde{y}_i and the empirical expectations \tilde{x}_i which we assume as defined in Gavrillets' model [29].

The dynamics in this work follow a De Groot-type update driven by three psychological forces: cognitive dissonance, social projection, and conformity. Cognitive dissonance aligns actions and self-beliefs [59]; social projection reflects the assumption that others are similar to oneself [13]; and logical constraints reduce inconsistencies between beliefs about others' actions and attitudes [60]. Attitudes and expectations thus adapt toward peers' average behaviour X_i [61], or externally promoted standards G_i^l .

Rewriting the equations in [29], we obtain

$$\begin{aligned} y_i' &= y_i + \xi_i^1 \left[\widehat{C}_i^{11} x_i + \widehat{C}_i^{12} X_i + \widehat{C}_i^{13} G_i^1 - y_i \right] \\ \tilde{y}_i' &= \tilde{y}_i + \xi_i^2 \left[\widehat{C}_i^{21} y_i + \widehat{C}_i^{22} X_i + \widehat{C}_i^{23} G_i^2 - \tilde{y}_i \right] \\ \tilde{x}_i' &= \tilde{x}_i + \xi_i^3 \left[\widehat{C}_i^{31} \tilde{y}_i + \widehat{C}_i^{32} X_i + \widehat{C}_i^{33} G_i^3 - \tilde{x}_i \right] \end{aligned} \quad (10)$$

where $\xi_i^l = \sum_j C_i^{lj}$ is interpreted as the rate at which the l^{th} variable ($l = 1$ corresponds to personal norms variable y , $l = 2$ corresponds to normative expectations variable \tilde{y} and $l = 3$ corresponds to empirical expectations variable \tilde{x}) is updated, and $\widehat{C}_i^{lj} = \frac{C_i^{lj}}{\sum_k C_i^{lk}}$ such that $\sum_j \widehat{C}_i^{lj} = 1$. In the rest of the article we assume for simplicity that $\xi_i^1 = \xi_i^2 = \xi_i^3 = 1$. The results were robust across a wide range of parameter values, in particular we examined the cases where $\xi_i^1 < \xi_i^2 < \xi_i^3$ which reflects the empirical observations that personal norms are the slowest to adapt while empirical expectations the fastest. This hierarchy is experimentally justified as evidence shows that personal norms remain stable and predictive even when empirical and normative expectations vary, suggesting that personal moral evaluations are less context-dependent than descriptive or injunctive cues [62]. Future extensions could introduce state-dependent ξ_i^q , potentially yielding qualitative transitions similar to those reported in [63].

Expressing the equations in this form clarifies that the coefficients \widehat{C}_i^{lj} represent the weights determining how each of the three variables influences the evolution of the l^{th} variable. More specifically, in this formulation, the coefficients $\widehat{C}_i^{\ell 1}$ and $\widehat{C}_i^{\ell 2}$ can be interpreted as self- and peer-weights, respectively: $\widehat{C}_i^{\ell 1}$ measures how strongly variable ℓ (personal norm, normative expectation, or empirical expectation) is pulled towards the agent's own state, whereas $\widehat{C}_i^{\ell 2} = 1 - \widehat{C}_i^{\ell 1}$ measures the influence of neighbours' actions. When external authorities are present, $\widehat{C}_i^{\ell 3}$ captures the weight assigned to externally promoted targets G_i^ℓ . In

our baseline analysis we set $\hat{C}_i^{\ell 3} = 0$, so that all changes arise endogenously from peer influence and internal consistency, and introduce non-zero $\hat{C}_i^{\ell 3}$ only when explicitly studying external interventions. Each variable thus depends on two inputs: the individual’s own norms (x_i, y_i, \tilde{y}_i) and the neighbours’ actions X_i . $\widehat{C}_i^{\ell 1}$ denotes the weight on self-referential updating, while $1 - \widehat{C}_i^{\ell 1}$ captures peer influence, corresponding to inward versus outward psychological orientation. We assume that

$$C_i^{\ell 1} = 1 - \phi_i(t) \quad (11)$$

and hence $C_i^{\ell 2} = \phi_i(t)$ where $\phi_i(t)$ as defined in the previous section. This links the coefficients to measurable quantities from the epidemic dynamics rather than assuming arbitrary values, and hence tailors the dynamics to our specific context—the evolution of social norms within a population facing collective epidemic risk. We link these weights to the change-detector $\phi_i^{change}(t)$ and consensus functions $\phi_i^{consensus}(t)$, capturing that agents rely more on personal anchors in unstable environments and on peers in stable ones. This reflects the idea that norm updating accelerates under clear social signals and slows when noise increases (Fig. 2B). From a NorMAS perspective, the model includes three norm-relevant constructs at once: a personal evaluative component (y_i), a descriptive component (\tilde{x}_i), and an injunctive component (\tilde{y}_i). In the present application all three concern the same focal behaviour, namely vaccination, but the framework could be extended so that different norms concern partially distinct behaviours or different sources of approval and disapproval.

3 Results

3.1 Experimental settings

Unless otherwise stated, simulations use a population of $N = 500$ agents and a physical-layer SIR process with infection rate β and recovery rate $\mu = 1$. For each season and each network realization, we run $n_{sim} = 1000$ independent SIR realizations to estimate infection risks. Vaccination decisions are then updated once per season. The bounded rationality parameter is set to $\kappa = 0.1$, the vaccination and infection costs are $c_V = 1$ and $c_I = 0.1$, and the default memory length is $m = 4$. The system is run until an evolutionary steady state is reached, operationalized as a change in the vaccinated fraction of less than 0.025 over the last 50 seasons, with a maximum horizon of 200 seasons. Unless otherwise stated, each reported point is computed over 1000 independent realizations with different network instances, and we report lower and upper quartiles. We now present the main behavioural and epidemic outcomes of the model under the baseline parameterization described above.

The initial values of the three norms associated with each agent are drawn independently from a uniform distribution. Finally, in all the cases studied, we made sure that the system was simulated for sufficiently many seasons such that evolutionary equilibrium was reached, in particular that the percentage of

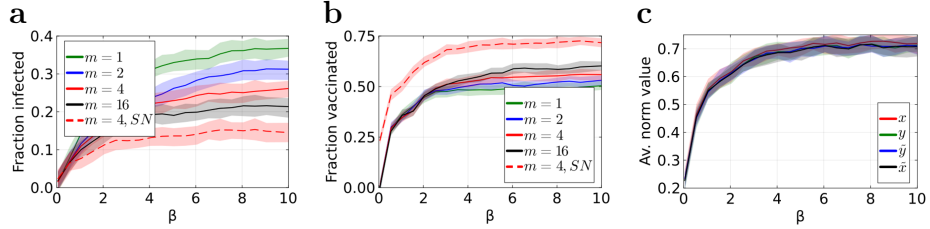


Fig. 3: Role of memory and social norms in epidemic outcomes. **a.** Infected fraction versus infectivity rate β . Infection increases with β and saturates at high values, while longer memory reduces outbreak size. For fixed $m = 4$, adding social norm dynamics further lowers infections. **b.** Vaccination coverage versus β . Coverage stabilizes near 0.5 for large β due to panic effects, but with norms, even small β sustain $\sim 20\%$ coverage as heterogeneous initial norms persist. **c.** Mean values of y , \tilde{y} , \tilde{x} , and x . All converge in the long run, indicating alignment between norms and vaccination intention.

vaccinated individuals did not change by more than 0.025 in the last 50 seasons and with a maximum of 200 seasons. In all of the figures presented in this section we include the lower and upper quartile range as well.

The role of social norms Initially, we study the role of memory on epidemic outcomes without considering social norms. Figure 3a shows that the infected fraction increases with the infectivity rate β , remaining close to zero for small β and saturating at high values. Longer memory reduces the outbreak size, as agents incorporate a richer history of payoffs when deciding whether to vaccinate. Figure 3b displays the corresponding vaccination coverage: for small β , coverage is close to zero in the absence of norms, while for large β it rises toward 0.5, driven by what we call a panic effect: as infection risk becomes sufficiently high, the payoff disadvantage of remaining unvaccinated becomes large for many agents at the same time, so vaccination increases sharply even without strong normative reinforcement. Increasing memory boosts coverage further, though not linearly. This is both due to network effects, but also due to the fact that high β values reduce the safety factor $S_i(t)$ and thus amplify the role of recent outcomes in shaping decisions. When social norm dynamics are introduced, both infection levels and coverage change significantly. Noteworthy, norms sustain a non-trivial fraction of vaccination ($\approx 20\%$) even at low β , reflecting the persistence of initial heterogeneous attitudes. Figure 3c illustrates how the three norm variables evolve. As expected, at equilibrium, the three norms coincide, as well as with the vaccination intention.

The role of external factors We study the role of an external factor affecting the evolution of social norms, assuming that the external effect can be present independently of the epidemiological state of the community. This is introduced by setting $\widehat{C}_i^{l3} = \gamma_i^l$ where $0 \leq \gamma_i^l \leq 1$ for $l \in \{1, 2, 3\}$ is the strength of the

external factor signal, and G_i^l is the target value for the norms. Intuitively, G_i^l is not a directly imposed behaviour, but the value toward which an external influence pulls the corresponding norm-related variable. For example, a campaign emphasizing that “vaccination is the right thing to do” would act on personal attitudes y_i or normative expectations \tilde{y}_i , whereas a message stating that “most people vaccinate” would act on empirical expectations \tilde{x}_i . Thus, G should be interpreted as the behavioural content of the external message once translated into the $[0, 1]$ scale of the corresponding variable, while γ measures how strongly agents are exposed to or persuaded by that message. $\widehat{C}_i^{lj} = (1 - \gamma_i^l) \widehat{C}_i^{lj}$, where \widehat{C}_i^{l1} and $\widehat{C}_i^{l2} = 1 - \widehat{C}_i^{l1}$ as before. Note that the condition that the coefficients of the variables sum up to 1 is still satisfied such that these can be still interpreted as relative contributions. In our studies, we assume the same strength γ_i^l and target G_i^l for all agents, i.e. $\gamma_i^l = \gamma^l$ and $G_i^l = G^l \forall i \in \{1, \dots, N\}$.

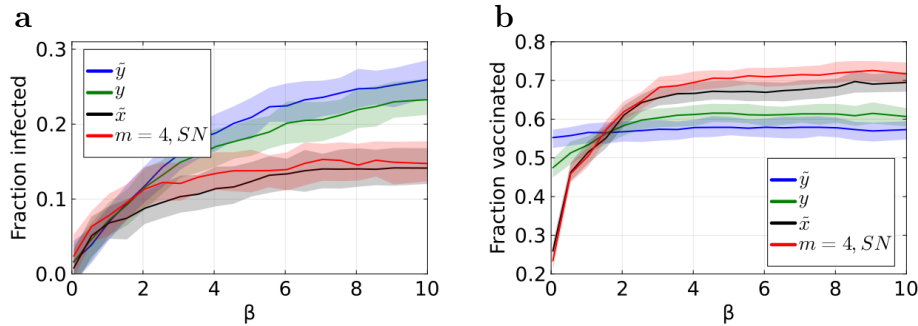


Fig. 4: **Role of external factors on social norms.** **a.** Infected fraction versus infectivity rate β when an external factor influences y_i , \tilde{y}_i , or \tilde{x}_i . The factor reduces perceived disease severity, driving norms toward $G_i = 0.6$, below the equilibrium value 0.7 in Fig. 3. It increases infections when applied to y_i or \tilde{y}_i , but has negligible effect on \tilde{x}_i . **b.** Vaccination coverage versus β for the same cases. Parameters: $\gamma_i = 1.0$, $G_i = 0.6$.

Figures 4 and 5 analyse how external influences on social norms alter epidemic outcomes. Fig. 4 shows that when an external factor plays down disease severity by shifting norms toward a lower target ($G = 0.6$), outbreaks become larger if the intervention acts through personal norms or injunctive expectations, while effects through empirical expectations remain negligible. Moving to Fig. 5, we observe richer dynamics. Panel A demonstrates that the outbreak size depends not only on which norm is targeted but also on the strength of the intervention γ and its target value G . Driving personal norms or injunctive expectations toward lower values reliably increases outbreak size, whereas targeting empirical expectations can sometimes backfire against the external factor, demonstrating a resilience of the system against such attempts: for intermediate values of G ($G \approx 0.6 - 0.7$), the community resists, and infection levels fall below

the baseline without external influence. Panel B further highlights these nonlinearities, showing that for target values close to $G \approx 0.7$ outbreaks can shrink under both normative expectation and empirical expectation interventions, with empirical expectations providing the strongest buffer, while for low G values ($G < 0.6$) the outbreak size grows as expected. For $G \approx 0.4 - 0.6$ targeting personal or injunctive norms remains more effective, however for sufficiently low G targeting empirical expectations might be more effective. Finally, Fig. 6 examines how the average values of each norm evolve under external driving. Only the targeted norm consistently converges to the imposed value, but interventions on injunctive expectations also drag the other norms along, underscoring their coordinating role.

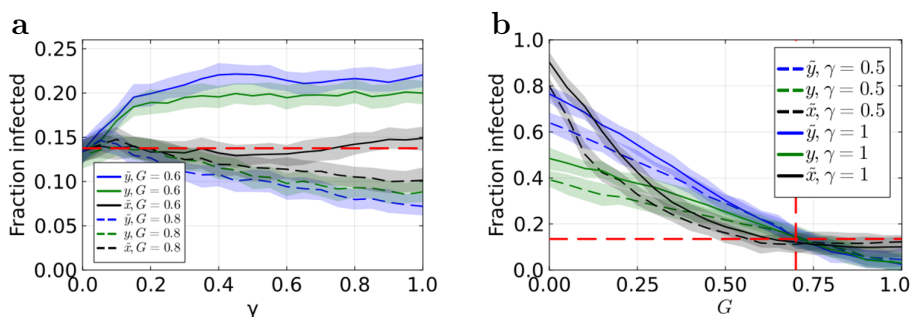


Fig. 5: Role of the external factor's strength γ and target value G . **a.** Outbreak size versus coupling strength γ when the external factor acts on y_i , \tilde{y}_i , or \tilde{x}_i , with targets $G = 0.6$ (lower severity perception) and $G = 0.8$ (higher). The red horizontal dashed line marks the baseline for the proportion of infected individuals without intervention. Driving y_i or \tilde{y}_i toward $G = 0.6$ increases infections, while the same external force applied on empirical expectations \tilde{x}_i has little effect. For a driving guiding vaccination level to $G = 0.8$ a reduction in the outbreak size is observed for all 3 case, however the intervention is more effective when applied on the normative expectation \tilde{y}_i and less effective when applied to empirical expectations \tilde{x}_i . **b.** Outbreak size versus G for $\gamma = 0.5$ and $\gamma = 1$. The red horizontal dashed line marks the baseline for the proportion of infected individuals without intervention, while the vertical represents the corresponding proportion for vaccinated individuals. For large G (nudging toward vaccination), the intervention is succesful and the fraction of infected individuals approaches towards zero when the intervention is applied on the moral-related norms y_i or \tilde{y}_i . The intervention is however much less effective when applied on \tilde{x}_i ; for small G , outbreaks grow. For values of $G \approx 0.6 - 0.7$ interventions on \tilde{x}_i could even buffer against external influence (for $\gamma = 0.5$, black dashed line), i.e. the outbreak size is even smaller than what would have been had there been no intervention.

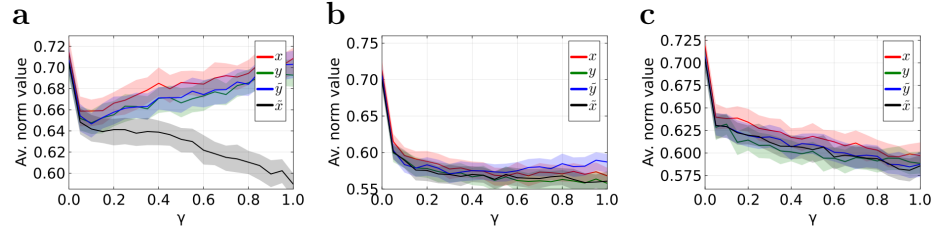


Fig. 6: **Dependence of social norms values on the external factor’s coupling strength γ .** Each scenario applies external coupling to one norm variable (y , \tilde{y} , or \tilde{x}) with target value $G = 0.6$. Average values across agents are shown. **a.** Coupling to \tilde{x} drives it toward the target as γ increases, while other norms remain unaffected. **b.** Coupling to \tilde{y} causes all norms to follow its value. **c.** Coupling to y similarly aligns the other norms with it.

Policy relevance In this interpretation, interventions do not mechanically set vaccination behaviour. Rather, they alter the target toward which one component of the decision process is pulled. A government campaign, media message, or trusted-authority recommendation can therefore be represented as a change in G and γ , depending on both the content of the message and its persuasive strength. Distinguishing injunctive from descriptive norm interventions is essential for effective public-health messaging. Injunctive messages communicate what is socially or morally expected—such as appeals from trusted authorities or statements like “vaccination is a responsible act.” Descriptive messages instead report what most people do (e.g., vaccination rates). Our simulations show that descriptive messages can backfire when they highlight uptake lower than the one without intervention (e.g., $G = 0.6$ and $\gamma = 1$ for a vaccination uptake without intervention of 0.7), inadvertently signalling that non-vaccination is common.

Empirical studies by Heiman et al. [40] and Szekely et al. [42] confirm that interventions targeting perceived approval and moral norms can strongly shift vaccination intentions, especially under uncertainty about others’ behavior. Our model formalizes how such cues, combined with probabilistic decision-making and social learning, affect collective behavior. While multilayer co-evolutionary frameworks coupling information, vaccination, and epidemics [65] emphasize structural feedbacks between awareness and contagion, our approach focuses on behavioral feedbacks from evolving normative expectations and adaptive learning. Together, these results show that external influences affect norms asymmetrically: interventions on injunctive norms propagate system-wide, whereas those on empirical expectations can create resilience or even neutralize attempts to downplay risk. This aligns with the experimental evidence of Szekely et al. [42] and the theoretical analysis of Roozmand et al. [66], which identify normative variables as central to norm emergence.

Finally, the simulations highlight three general principles of norm–epidemic coevolution. (i) Even simple norm dynamics can suppress outbreaks by sustaining vaccination at low infectivity and smoothing behavioral fluctuations, consistent with earlier behavioral–epidemic models where imitation or reinforcement

alone cannot maintain cooperation [8, 9]. (ii) Norm components are asymmetric: external driving applied on normative expectations exerts stronger collective influence than if applied on empirical expectations, consistent with evidence that perceived approval is a key coordination device [36, 42] (for not too low values of target values G). (iii) External interventions on norms can have nonlinear and counterintuitive effects [40, 2]. These patterns illustrate how the structure of norm dynamics shapes epidemic and behavioral outcomes, complementing earlier analyses of awareness–disease coupling in multilayer networks [5, 6].

4 Discussion

Epidemic–behavior models often rely on imitation or payoff-based reinforcement, which capture some dynamics but overlook how attitudes, expectations, and social pressures shape preventive choices. To address this, we developed a behavioral epidemic model coupling disease spread with a decision process based on the Experience-Weighted Attraction (EWA) framework [37]. EWA integrates reinforcement and belief learning by considering both realized and forgone payoffs, offering a more realistic account of human decision making. We extend this by embedding social norm dynamics, distinguishing between descriptive and injunctive norms [16, 17, 29]. The model captures cognitive dissonance, social projection, and consistency mechanisms that explain how individuals align with peers or authorities.

Our analysis shows that coevolving norm dynamics markedly alter outbreak size. External interventions targeting personal or injunctive expectations strongly affect vaccination coverage, whereas those acting on descriptive norms can be weaker or even counterproductive. Most importantly, the model’s qualitative predictions align with recent experimental findings. For example, the strong coordinating role of normative expectations echoes the “transmission game” study by Woike *et al.*, where injunctive-norm messages reduced risk-taking more effectively than descriptive information or case counts [36]. These parallels suggest a pathway for empirical calibration: longitudinal experiments or panel surveys could track the joint evolution of personal norms, empirical expectations, and perceived approval under repeated exposure to norm-based interventions, allowing the coefficients in our norm-update equations to be estimated and tested [2]. Our model could also be employed to examine the asymmetry in the recovery of social norms observed in Vriens *et al.* [63].

Our framework is stylized and not predictive, but explanatory, designed to provide theoretical insights on how social norms could potentially coevolve with epidemic risk and shape vaccination behavior. While parameters are not empirically calibrated and agents are homogeneous in risk perception, trust, and norm sensitivity, these simplifications allow a controlled exploration of core behavioral mechanisms. The current setup also assumes fixed social networks and stable information environments, abstracting from media or institutional feedbacks. A recent complementary model incorporating regret and uncertainty [67], examines the role of more emotional and informational mechanisms in the co-

evolving dynamics of disease outbreaks and social norms evolution. Future extensions could incorporate heterogeneous agents, adaptive or multilayer networks [5, 48], and empirical calibration using experimental or survey data, thereby enhancing realism and policy relevance. Such developments would retain the model’s interpretability while connecting it more closely to real-world interventions. More broadly, empirically grounded norm-based models can guide communication strategies and interventions across collective-action domains, from vaccination and misinformation to environmental cooperation [20, 21, 65, 68–70]. Future work could also extend the framework to include homophily in intentions or vaccination status [69], and peer sanctions and rewards in norm evolution [70].

Acknowledgments. I am thankful to Prof. Sandro Meloni for insightful discussions. This project received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No 101034403.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Code and data availability

The simulation code and scripts used to generate the figures are available at: <https://github.com/Christos3788/Social-norm-dynamics-in-a-Behavioral-epidemic-model>.

References

1. WHO. Ten threats to global health in 2019. *WHO* (2019).
2. Vriens, E., Tummolini, L. & Andrighetto, G. Vaccine-hesitant people misperceive the social norm of vaccination. *PNAS Nexus* **2**, pgad132; 10.1093/pnasnexus/pgad132 (2023).
3. Walker, B. et al. Looming global-scale failures and missing institutions. *Science* **325**, 1345–1346; 10.1126/science.1175325 (2009).
4. Nyborg, K. et al. Social norms as solutions. *Science* **354**, 42–43 (2016).
5. Wang, Z., Andrews, M. A., Wu, Z. X., Wang, L. & Bauch, C. T. Coupled disease–behavior dynamics on complex networks: a review. *Phys. Life Rev* **15**, 1–29; 10.1016/j.plev.2015.07.006 (2015).
6. Granell, C., Gómez, S. & Arenas, A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Phys. Rev. Lett.* **111**, 128701; 10.1103/PhysRevLett.111.128701 (2013).
7. Han, D. & Wang, X. Impact of positive and negative information on epidemic spread in a three-layer network. *Chaos, Solitons Fractals* **186**, 115264; 10.1016/j.chaos.2024.115264 (2024).
8. Bauch, C. T. & Earn, D. J. D. Vaccination and the theory of games. *Proc. Natl Acad. Sci. USA* **101**, 13391–13394; 10.1073/pnas.0403823101 (2004).
9. Reluga, T. C. Game theory of social distancing in response to an epidemic. *PLOS Comput. Biol.* **6**, e1000793; 10.1371/journal.pcbi.1000793 (2010).

10. Shi, B. E. et al. Quantifying the interaction between infectious diseases and preventive behavior. *Phys. Rev. E* **100**, 022310; 10.1103/PhysRevE.100.022310 (2019).
11. Bhattacharyya, S. & Bauch, C. T. A game dynamic model for delayer strategies in vaccinating behaviour for pediatric infectious diseases. *J. Theor. Biol.* **267**, 276–282; 10.1016/j.jtbi.2010.08.036 (2011).
12. Festinger, L. *A Theory of Cognitive Dissonance*. (Stanford University Press, 1957).
13. Krueger, J. I. & Clement, R. The truly false consensus effect: an ineradicable and egocentric bias in social perception. *J. Pers. Soc. Psychol.* **92**, 582–595; 10.1037/0022-3514.92.4.582 (2007).
14. Apperly, I. *Mindreaders: the cognitive basis of "theory of mind"*. (Psychology Press, 2010).
15. Cialdini, R. B., Reno, R. R. & Kallgren, C. A. A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* **58**, 1015–1026; 10.1037/0022-3514.58.6.1015 (1990).
16. Bicchieri, C. *The Grammar of Society: The Nature and Dynamics of Social Norms*. (Cambridge University Press, Cambridge, 2005).
17. Ostrom, E. Collective action and the evolution of social norms. *J. Econ. Perspect.* **14**, 137–158; 10.1257/jep.14.3.137 (2000).
18. Agranov, M., Elliott, M. & Ortoleva, P. The importance of social norms against strategic effects: the case of COVID-19 vaccine uptake. *Econ. Lett.* **206**, 109979; 10.1016/j.econlet.2021.109979 (2021).
19. Graupensperger, S., Abdallah, D. A. & Lee, C. M. Social norms and vaccine uptake: college students' COVID vaccination intentions, attitudes, and estimated peer norms and comparisons with influenza vaccine. *Vaccine* **39**, 2060–2067; 10.1016/j.vaccine.2021.03.018 (2021).
20. Moehring, A. et al. Providing normative information increases intentions to accept a COVID-19 vaccine. *Nat. Commun.* **14**, 126; 10.1038/s41467-022-35052-4 (2023).
21. Brewer, N. T., Chapman, G. B., Rothman, A. J., Leask, J. & Kempe, A. Increasing vaccination: putting psychological science into action. *Psychol. Sci. Public Interest* **18**, 149–207; 10.1177/1529100618760521 (2017).
22. Bauch, C. T. & Bhattacharyya, S. Evolutionary game theory and social learning can determine how vaccine scares unfold. *PLOS Comput. Biol.* **8**, e1002452; 10.1371/journal.pcbi.1002452 (2012).
23. d'Onofrio, A., Manfredi, P. & Poletti, P. The interplay of public intervention and private choices in determining the outcome of vaccination programmes. *PLOS ONE* **7**, e45653; 10.1371/journal.pone.0045653 (2012).
24. Oraby, T., Thampi, V. & Bauch, C. T. The influence of social norms on the dynamics of vaccinating behaviour for paediatric infectious diseases. *Proc. R. Soc. B* **281**, 20133172; 10.1098/rspb.2013.3172 (2014).
25. Battigalli, P. & Dufwenberg, M. Belief-dependent motivations and psychological game theory. *J. Econ. Lit.* **60**, 833–882; 10.1257/jel.20201378 (2022).
26. López-Pérez, R. Aversion to norm-breaking: a model. *Games Econ. Behav.* **64**, 237–267; 10.1016/j.geb.2007.10.009 (2008).
27. Realpe-Gómez, J., Vilone, D., Andrighetto, G., Nardin, L. G. & Montoya, J. A. Learning dynamics and norm psychology supports human cooperation in a large-scale prisoner's dilemma on networks. *Games* **9**, 90; 10.3390/g9040090 (2018).
28. Andrighetto, G., Brandts, J., Conte, R., Sabater-Mir, J., Solaz, H. & Villatoro, D. Punish and voice: punishment enhances cooperation when combined with norm-signalling. *PLOS ONE* **8**, e64941; 10.1371/journal.pone.0064941 (2013).
29. Gavrillets, S. Coevolution of actions, personal norms, and beliefs about others. *Evol. Hum. Sci.* **3**, e44; 10.1017/ehs.2021.40 (2021).

30. Li, X., Molleman, L. & van Dolder, D. Do descriptive social norms drive peer punishment? conditional punishment strategies and their impact on cooperation. *Evol. Hum. Behav.* **42**, 469–479; 10.2139/ssrn.3571220 (2021).
31. Bravo, G., Squazzoni, F. & Boero, R. Trust and partner selection in social networks: an experimentally grounded model. *Soc. Netw.* **34**, 481–492; 10.1016/j.socnet.2012.03.001 (2012).
32. Taghikhah, F., Filatova, T. & Voinov, A. Where does theory have it right? A comparison of theory-driven and empirical agent-based models. *J. Artif. Soc. Soc. Simul.* **24**, 4; 10.18564/jasss.4573 (2021).
33. Realpe-Gómez, J., Andrighetto, G., Nardin, L. G. & Montoya, J. A. Balancing selfishness and norm conformity can explain human behavior in large-scale prisoner’s dilemma games and can poise human groups near criticality. *Phys. Rev. E* **97**, 042321; 10.1103/PhysRevE.97.042321 (2018).
34. Campbell, D. T. ‘Downward causation’ in hierarchically organised biological systems. In *Studies in the Philosophy of Biology: Reduction and Related Problems* 179; 10.1007/978-1-349-01892-5_11 (1974).
35. Castelfranchi, C. Simulating with cognitive agents: the importance of cognitive emergence. In *Multi-Agent-Based Simulation* 26–44; 10.1007/10692956_3 (1998).
36. Woike, J. K., Hafenbrädl, S., Kanngiesser, P. & Hertwig, R. The transmission game: testing behavioral interventions in a pandemic-like simulation. *Sci. Adv.* **8**, eabk0428; 10.1126/sciadv.abk0428 (2022).
37. Camerer, C. & Ho, T.-H. Experience-weighted attraction learning in normal form games. *Econometrica* **67**, 827–874; 10.1111/1468-0262.00054 (1999).
38. Chapman, G. B., Li, M., Vietri, J., Ibuka, Y., Thomas, D., Yoon, H. & Galvani, A. Using game theory to examine incentives in influenza vaccination behavior. *Psychol. Sci.* **23**, 1008–1015; 10.1177/0956797612437606 (2012).
39. Ho, T.-H., Camerer, C. F. & Chong, J.-K. Self-tuning experience weighted attraction learning in games. *J. Econ. Theory* **133**, 177–198; 10.1016/j.jet.2005.12.008 (2007).
40. Heiman, S. L., Garrett, N., Vosgerau, J., Tinghög, G. & Västfjäll, D. Descriptive norms caused increases in mask wearing during the COVID-19 pandemic. *Sci. Rep.* **13**, 11856; 10.1038/s41598-023-38593-w (2023).
41. Gelfand, M. J. & Harrington, J. R. The motivational force of descriptive norms: for whom and when are descriptive norms most predictive of behavior? *J. Cross-Cult. Psychol.* **46**, 1273–1278; 10.1177/0022022115600796 (2015).
42. Szekely, A., Lipari, F., Antonioni, A., Paolucci, M., Sánchez, A., Perc, M. & Andrighetto, G. Evidence from a long-term experiment that collective risks change social norms and promote cooperation. *Nat. Commun.* **12**, 5452; 10.1038/s41467-021-25734-w (2021).
43. Meiske, B., Álvarez Benjumea, A., Andrighetto, G. & Polizzi, E. Nudging punishment against sharing of fake news. *Eur. Econ. Rev.* **168**, 104795; 10.1016/j.euroecorev.2024.104795 (2024).
44. Tverskoi, D., Babu, S. & Gavrilets, S. The spread of technological innovations: effects of psychology, culture and policy interventions. *R. Soc. Open Sci.* **9**, 211833; 10.1098/rsos.211833 (2022).
45. Gavrilets, S. & Richerson, P. J. Authority matters: propaganda and the coevolution of behaviour and attitudes. *Evol. Hum. Sci.* **4**, e51; 10.1017/ehs.2022.48 (2022).
46. Salathé, M., Kazandjieva, M., Lee, J. W., Levis, P., Feldman, M. W. & Jones, J. H. A high-resolution human contact network for infectious disease transmission. *Proc. Natl. Acad. Sci. USA* **107**, 22020–22025; 10.1073/pnas.1009094108 (2010).

47. Klimek, P. & Thurner, S. Triadic closure dynamics drives scaling laws in social multiplex networks. *New J. Phys.* **15**, 063008; 10.1088/1367-2630/15/6/063008 (2013).
48. Charalambous, C., Sanchez, D. & Toral, R. Language competition on coevolving networks: an agent-based approach of nodes and links coevolution. *Front. Complex Syst.* **1**, 1304448; 10.3389/fcpxs.2023.1304448 (2023).
49. Cardillo, A., Reyes-Suárez, C., Naranjo, F. & Gómez-Gardeñes, J. Evolutionary vaccination dilemma in complex networks. *Phys. Rev. E* **88**, 032803; 10.1103/PhysRevE.88.032803 (2013).
50. Kiss, I. Z., Miller, J. C. & Simon, P. L. *Mathematics of epidemics on networks*. Springer, Cham; 10.1007/978-3-319-50806-1 (2017).
51. Davis, R., Campbell, R., Hildon, Z., Hobbs, L. & Michie, S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychol. Rev.* **9**, 323–344; 10.1080/17437199.2014.941722 (2015).
52. Martins, A. C. R. Continuous opinions and discrete actions in opinion dynamics problems. *Int. J. Mod. Phys. C* **19**, 617–624; 10.1142/S0129183108012339 (2008).
53. Jager, W., Janssen, M. A., De Vries, H. J. M., De Greef, J. & Vlek, C. A. J. Behaviour in commons dilemmas: Homo economicus and Homo psychologicus in an ecological-economic model. *Ecol. Econ.* **35**, 357–379; 10.1016/S0921-8009(00)00220-2 (2000).
54. Galla, T. & Farmer, J. D. Complex dynamics in learning complicated games. *Proc. Natl. Acad. Sci. USA* **110**, 1232–1236; 10.1073/pnas.1109672110 (2013).
55. Sutton, R. S. & Barto, A. G. *Reinforcement learning: an introduction*. MIT Press (2018).
56. Feltovich, N. Belief-based learning models. *Encyclopedia of the Sciences of Learning*, 444–447; Springer; 10.1007/978-1-4419-1428-6_111 (2012).
57. Sánchez, A. Physics of human cooperation: experimental evidence and theoretical models. *J. Stat. Mech.* **2018**, 024001; 10.1088/1742-5468/aaa388 (2018).
58. Hedden, T. & Zhang, J. What do you think I think you think?: strategic reasoning in matrix games. *Cognition* **85**, 1–36; 10.1016/S0010-0277(02)00054-9 (2002).
59. Rabin, M. Cognitive dissonance and social change. *J. Econ. Behav. Organ.* **23**, 177–194; 10.1016/0167-2681(94)90066-3 (1994).
60. Friedkin, N. E., Proskurnikov, A. V., Tempo, R. & Parsegov, S. E. Network science on belief system dynamics under logic constraints. *Science* **354**, 321–326; 10.1126/science.aag2624 (2016).
61. Kashima, Y. et al. Social transmission of cultural practices and implicit attitudes. *Organ. Behav. Hum. Decis. Process.* **129**, 113–125; 10.1016/j.obhdp.2014.05.0 (2015).
62. Catola, M., D’Alessandro, S., Guarnieri, P. & Pizziol, V. *Personal norms in the online public good game*. *Economics Letters* **207**, 110024 (2021).
63. Vriens, E., Andrighetto, G. & Tummolini, L. Risk, uncertainty, & social norms in vaccination decisions. *Philos. Trans. R. Soc. B* **379**, 20230035; 10.1098/rstb.2023.0035 (2024).
64. Goeree, J. K., Holt, C. A. & Palfrey, T. R. Quantal response equilibria. In *Encyclopedia of Complexity and Systems Science* (eds Meyers, R. A.) 234–242 (Springer, 2010).
65. Yin, Q., Wang, Z., Xia, C. & Bauch, C. T. Impact of co-evolution of negative vaccine-related information, vaccination behavior and epidemic spreading in multilayer networks. *Commun. Nonlinear Sci. Numer. Simul.* **109**, 106312; 10.1016/j.cnsns.2022.106312 (2022).

66. Roozmand, O., Deffuant, G., Andrighetto, G. & Paolucci, M. Simulating collective risk management from experimental data. In *Social Simulation Conference 2022* (2022).
67. Charalambous, C. Regret, Uncertainty, and Bounded Rationality in Norm-Driven Decisions. arXiv:2511.10342 (2025).
68. Epstein, J. M. *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science*. (Princeton University Press, 2014).
69. Holme, P. & Newman, M. E. J. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* **74**, 056108; 10.1103/PhysRevE.74.056108 (2006).
70. Gavrillets, S. & Richerson, P. J. Collective action and the evolution of social norm internalization. *Proc. Natl Acad. Sci. USA* **114**, 6068–6073; 10.1073/pnas.1703857114 (2017).