Learned Futility: How Social Learning Can Lead to the Diffusion of Ineffective Strategies

Veronika Kurchyna[1,6], Lilian Kojan[2], Jan Schneider[3], Bernd Wurpts[4], Anastasia Golovin[5], André Calero Valdez[2], Jan Ole Berndt[1], and Ingo J. Timm[1]

[1] German Research Center for Artificial Intelligence, Cognitive Social Simulation
[2] University of Luebeck
[3] Wroclaw University of Science and Technology
[4] Independent Researcher
[5] Max Planck Institute for Dynamics and Self-Organization
[6] Corresponding Author: veronika.kurchyna@dfki.de

Abstract. When facing a crisis, such as a novel type of disease, individuals need to learn about effective health measures and practices to prevent the spread of illness. They do so both through reflection about their own actions as well as the communal experience of their peers. Here, we present an agent-based model to examine the resulting dynamics in the diffusion of health behaviours and practices. In the model, we employ reinforcement learning and bounded confidence opinion dynamics to model varying degrees of external, e.g. social, and internal knowledge gains in the context of protective measures against a novel disease as use case. Our study shows that social influence is critical for the adoption of potentially effective low-cost strategies, while individual learning modes limit the spread of potentially harmful high-cost strategies. On the downside, social learning also facilitates the spread of ineffective or even harmful health measures and practices. Our findings suggest that cultural variation emerges in times of crisis among learning individuals.

Keywords: Agent-based Modelling · Multi-Agent Reinforcement Learning · Opinion Dynamics · Social Simulation

1 Introduction

The recent COVID-19 pandemic is the reflection of a broader pattern in the evolution of human history: increaseing numbers of emerging diseases. In the decades after the Second World War, hundreds of new diseases were identified, including dangerous pathogens that caused serious epidemics [24]. Increasing health threats through diseases call for a better understanding of the human responses that mediate the harmful impact of pathogens. While many of the emergency responses in modern societies are coordinated by governments and international organisations, in times of crisis, people frequently draw on their own knowledge and learn from their social environment. Lacking the knowledge

about effective health measures and practices, individuals fall back on established measures from other contexts or develop health strategies in coordination with their social contacts. These micro-interactions and social learning experiences can produce cultural variation in the societal responses to epidemic diseases. This paper uses agent-based modelling to uncover some of the conditions under which effective and ineffective cultural variation emerges during times of crisis. Often certain groups display health behaviour that strongly differs from the majority. In the past, as example for effective group practices, we observed Jewish hygiene customs that safeguarded during the Black Death [27]. A contemporary example of ineffective or even harmful actions is the questionable use of antiparasitic drugs during the recent COVID-19 pandemic [26]. In social simulation, such phenomena are typically framed as processes of opinion spread and assimilation [12], which could be interpreted as peer pressure, a contagion of fear, or behavioural imitation. At the same time, it is self-evident that individuals are not solely imitating the behaviour of others but are also capable of accumulating and evaluating their own experiences and will adjust their behaviour accordingly, something represented in agent-based modelling as agent learning [30].

In this paper, we combine both mechanisms to describe how cultural practices spread in the context of disease prevention. The distinctive innovation of our method lies in its capacity to facilitate the co-learning of various actions, encompassing a wide range of effectiveness, from potentially harmful to highly effective. In contrast to traditional reinforcement learning, this allows a convergence towards cultural practices that are actually ineffective in preventing diseases. Through this novel approach, we investigate the interplay of learning processes, social influence dynamics, and the propagation of diseases, with a particular emphasis on the emergence of seemingly ineffective strategies, despite the potential for knowledge acquisition.

2 Background, Motivation and Research Question

To understand the current relevance of this research, it is important to consider it in its historical and social context.

2.1 Epidemic Diseases and Societal Responses

While epidemic diseases played only a minor role in social scientific research for many decades, the recent COVID-19 pandemic led to new efforts to improve public health and investigate the change of societies. Social scientists widely agree that many factors drive social change. They are related to social and technological innovations, changing demographics, revolutions, wars, or economic crises. Infectious diseases are also increasingly recognised as contributors towards change. Major killers such as plague, cholera, smallpox, HIV/AIDS and other illnesses were important causes in the generation of different forms of organised social responses and the advancement of methods in public health [24].

Historically, epidemics are distinct from other, less destructive and widespread forms of disease, such as chronic or genetic ones. Infectious diseases are ongoing threats and human societies stay susceptible to the transmission of known and novel pathogens. Records of responses to epidemics are indicative of human creativity and diversity across societies and periods. As the dominant frameworks on the causes of diseases changed, so did the ways individuals and groups responded to changing health threats [24]. Reactions are intrinsically tied to dominant medical doctrines and prevalent structures of power. As an example, the handling of the reduced mortality among patients of homeopathic practitioners during a cholera outbreak illustrates two phenomena: For one, although homeopaths might attribute success to homeopathic treatments, it was likely due to differing hygiene and rehydration practices. It is difficult to identify effective practices in the absence of rigorous scientific methods, since individuals typically engage in multiple practices concurrently. For another, the medical establishment outright disregarded the success of the homeopaths without investigating the reasons of success [9]. Regardless of success, people are more likely to adopt practices of similar individuals than those perceived as different.

Jumping to the 21st century, the COVID-19 pandemic marks an important contemporary case signalling the importance of dominant medical paradigms of disease and power structures. In the immediate pandemic period, individuals in the U.S. and many other countries applied various non-pharmaceutical interventions to confront the disease, partially because effective pharmaceutical interventions (e.g. medications and vaccines) were not yet available. The disease spread largely depended on the personal behavioural choices. Physical distancing measures (such as self-isolation) in combination with hygienic practices (including frequent hand-washing or wearing face masks) were highly effective. However, social distancing through contact reduction was also associated with major social and economic costs. Isolation negatively impacted mental health and disrupted the daily lives of individuals across the globe. Additionally, we observed various processes of social contagions in the dissemination of misinformation, fear and health advice. Observable behaviours, such as mask wearing, influenced group behaviour. Misinformation about cures caused many cases of self-injuries because harmful substances such as bleaching agents were praised as potent solutions. Less harmful home remedies such as eating garlic or drinking ginger tea were also falsely named effective on social media [7].

These examples suggest that societal responses to epidemic diseases are highly contingent on social context. Political authorities play an active role by coordinating measures and drawing from experiences with previous epidemics. Howeverthe success of public health policies relies on compliance and the active decisions of individuals. During times of crisis, individuals try to protect themselves and their closest contacts using various private and public sources as well as spiritual and non-spiritual goods and services to cope with the various consequences of epidemic diseases.

2.2 Diffusion and Learning of Cultural Practices

The theory of contagion was initially developed to explain diseases, but core ideas were also adopted by social sciences to study the diffusion of behaviours, beliefs, or cultural practices through social reinforcement. An important approach to the explanation of social change is therefore dedicated to the structures of social networks, defined as sets of relations among individuals [8, 6].

A fundamental assumption is that the structural characteristics of social networks are significant for transmission processes. Research shows that human networks are defined by connectivity (average degree of edges above 1) and short paths, referred to as the small world property. These traits provide important structural conditions for the diffusion of diseases, information, or practices [29].

As mobility and the range of contacts increased, so did the potential for epidemic diseases to spread rapidly and widely [18]. Research on the most deadly disease in human history, the Black Death, indicates that the fear response in the form of testaments and wills generated large connected networks that provided opportunities for information transmission and knowledge transfer [31].

Going beyond empirical studies of diseases, social scientists developed a set of learning models that incorporate network dynamics and foster a better understanding of general patterns in decision making processes. Through repeated interaction, information gathering and updating of beliefs, a social network converges towards shared beliefs. Commonly, the emerging consensus depends on the initial opinions of individuals and the degrees of nodes in the network [18].

While consensus may form in many contexts, opinions about health measures such as vaccinations are not homogeneous and large variations can be found in countries such as the United States where groups of "anti-vaxxers" emerged in several regions. Research based on agent-based modelling shows that these types of variation in cultural beliefs are not dependent on social network structures, but rather related to cognitive processes such as the interpretations of the relations among cognitive objects, i.e., associative diffusion rather than social contagion. This research shifts the focus from durable relationships in many network studies of social influence, to short-term observations of behaviours. It is argued that assumptions about balkanised worlds and segregated groups do not sufficiently explain cultural differentiation [13].

Research questions and approach In summary, our research is motivated by the observation that both historically and concurrently, cultural practices in the context of infectious diseases are strongly tied to the patterns of interaction in social groups. Thus, the central research question revolves around investigating the mechanisms through which reinforcement learning and opinion dynamics collectively contribute to the propagation of ineffective practices through incorrect beliefs (or opinions) about their effectiveness.

This is achieved by combining three fundamental mechanisms: Multi-action reinforcement learning, wherein multiple actions can be chosen in parallel; opinion dynamics, were agents can adapt their opinions about the effectiveness of practices based on the opinions of other agents; and an epidemiological model in

which effective actions curb disease transmission. This approach enables us to explore how a lack of understanding regarding the actual significance of individual actions in shaping the overall outcome fosters the development of suboptimal strategies. This view does not imply malice or ignorance on the part of the agents - however, during a pandemic, a successful agent will not be inclined to conduct systematic experiments to determine which of the different preventive measures contributed to their continued health. As such, the multi-agent perspective introduced through opinion dynamic and cooperative multi-agent reinforcement learning is necessary to mimic the way real people learn or abandon practices.

3 Foundations and Related Works

As an introduction to the methods used in this work, we give a brief presentation of basic epidemiological modelling and the modelling of learning agents using reinforcement learning, opinion dynamics and the way they can be combined.

3.1 Epidemiological Modelling

Both equation- and agent-based models can be used to examine questions related to epidemics, typically represented by a type of a Susceptible-Infected-Recovered (SIR) model [21]. The acronym stands for compartments characterising the progression of diseases in different stages. Reinfectability can be introduced by returning to Susceptible after a period of immunity, leading to an SIRS model [16].

3.2 Modelling Learning Agents

Characteristics of agency [30] such as autonomy, social interaction and adaptation to environments are often necessary to portray human behaviour. Some models achieve the simulation of behaviour using routines such as workplaces or education [25] without entering the complexity of human decision-making, while others focus on needs to implement goal-directed agent behaviour [14]. Finally, social components such as pressure or support [5] may also influence behaviour.

Reinforcement Learning Reinforcement Learning (RL) is a method of machine learning rooted in mechanisms observed in nature. *Operant Conditioning* examines behaviour beyond simple stimulus-response pairs and instead considers behaviour a result of rewards and punishments for actions emitted by the actor [28]. Through these responses from the environment, individuals learn which behaviours lead to desirable feedback, increasing the likelihood of repetition.

RL functions on similar principles: learning agents explore the action-state-space, gathering information about the attractiveness of certain actions (or practices) depending on the current state and possible subsequent states [19]. Contrary to goal-directed agents, reinforcement agents seek to maximise their rewards. As such, the definition of states, actions and the balancing of rewards

is crucial to the outcome of the learning process. Typically, implementations of RL will use some form of so-called Q-learning. In Q-learning, agents learn the utility of an action in a certain state - either learning ahead of the actual simulation (Monte Carlo Learning) or during runtime (temporal difference learning). While so far, only a single learning agent was spoken of, the same principle can be expanded to multi-agent reinforcement learning [32] in which agents either cooperatively explore different strategies or each agent learns on their own in competition against others.

Opinion Dynamics This field examines the way social influence leads to emergent patterns of opinion formation in populations. The basis of these models is social influence [11]. In short, agents will adapt their opinion to the opinion of other agents so overall, opinions will assimilate. That is, when two agents interact, either one or both of them will move their own opinion to be closer to the opinion of the other agents. An important extension is that similarity makes assimilation more likely, e.g., by including a bounded confidence mechanism [15]. That means that an agent will only adapt their opinion to that of another agent if the difference between their opinions is below a certain threshold, the confidence level. Conversely, opinion dynamic models can also be extended to include repulsive influence where under certain circumstances, agents actually move their opinion to be further away from the opinion of another agent [11].

Bounded Confidence and Social Learning Bounded confidence models can easily be applied to the concept of social learning. In learning by observing others, individuals will not just imitate any behaviour, but only that of sufficiently similar other individuals [3]. This minimum degree of similarity can be understood as the confidence level. However, while in the opinion dynamics literature, opinions are often examined independent from related behaviour, the social learning literature is concerned with behaviour and its consequences. Behaviour is imitated when these consequences are evaluated positively. So while bounded confidence models of opinion dynamics serve to describe *if* we learn from others, including some processing of behavioural outcomes serves to narrow down *what* we learn from others [22].

3.3 Related Work

Most generally, it is acknowledged that both social learning as well as learning based on individual experience can be described as a trade-off. Less costly strategies of information acquisition come with the risk of forming superstitions or false beliefs [20]. Works like these do not attribute this trade-off to specific theories or algorithmic approaches such as operant conditioning and social cognitive learning or reinforcement learning and bounded confidence opinion dynamics. Even so, they show that a link between learning strategies and 'superstitious' behaviour is not an entirely novel idea. In particular, social influences facilitating the establishment of irrational beliefs and superstitions is already well-established from

psychological and sociological perspectives. The notion that such beliefs are a byproduct of learning processes [4], as opposed to lack of insight or knowledge, supports this work's assumption that collective learning *supports* the formation of superstitions or ineffective strategies. Further, research also has shown that in situations of risk, stress and danger, these effects are particularly likely to emerge [23]. As such, a novel disease with associated health risks is a suitable use case to demonstrate the rise of ineffective actions despite learning agents.

Some research already applied reinforcement learning to examine the relationship of learning and superstition, noting that low-cost actions are particularly susceptible to the formation of superstition [1]. However, we are unaware of any such works combining reinforcement learning with social components or a multi-action learning algorithm. Generally, related work suggests that while the theoretical link between the different forms of learning and the emergence of superstitious behaviour is established in theoretical research, agent-based simulation is rarely used to incorporate these learning strategies to observe how specific habits of strategies are propagated in a community.

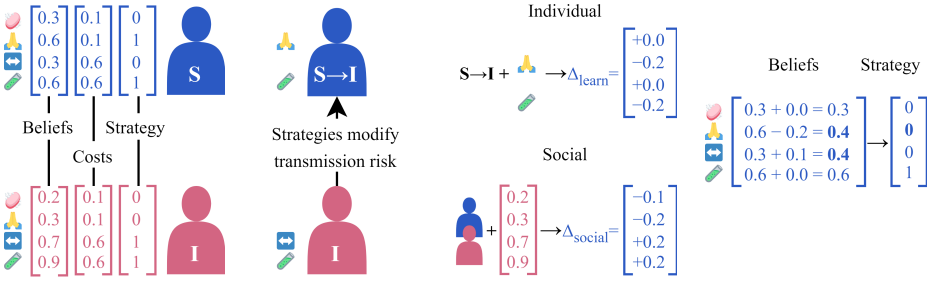
4 Conceptual Model and Mathematical Description

Agents interact and spread diseases in a simple infection model. Agents can reduce the transmission risk by engaging in different practices, each associated with costs and objective effectiveness ratings. The set of practices an agent is currently implementing is that agent's *strategy*. Agents hold *subjective beliefs* about the effectiveness of each practice and communicate about these beliefs. While they correspond to opinions in classical models of opinion dynamics, we will use the term *belief* going forward to emphasise that these cognitions are subjective views 'about' specific practices and their effects. A belief is formed through the combination of individual learning based on success so far, as well as communicated experiences of others.

Before proceeding with a detailed mathematical description, we present a schematic of how one agent interacts with another and updates their beliefs. In figure 1, one can observe how a blue agent (b) changes their disease status, beliefs and strategy based on their interaction with a red (r) agent. Numbers are for illustration purposes only. Also, for easier presentation, exogenous random influences on beliefs are not depicted.

At the outset of each step (1a), all agents hold a set of subjective *beliefs* about the effectiveness of the practices: Hand-washing, praying, physical distancing, and drinking bleach. If that belief is 0.5 or larger, the respective practice is active and part of the *strategy*. The practice *costs* are the same for all agents. Agents also have a disease status: Susceptible, infectious, or recovered. Here, b is susceptible and engages in praying and and drinking bleach, while r is infected and engages in physical distancing and drinking bleach.

In the infection propagation phase (1b), if a susceptible and an infected agent interact, there is the risk of disease transmission. This risk is decreased based on the effectiveness of the practices in the agents' strategies. Here, r infects b.



(a) Outset. (b) Infection propagation. (c) Learning. (d) Strategy update.

Fig. 1: Schematic sequence of one model step and one agent interaction.

In the learning phase (1c), both individual learning and social learning take place. Since b got infected, individual learning leads to b decreasing the belief in the practices that were active. Because b and r have similar enough beliefs, social learning leads to b's beliefs moving closer to r's beliefs.

In the strategy update phase (1d), individual and social learning combined change b's belief vector sufficiently that b's strategy also changes. The belief in praying is below the threshold of 0.5, so the new strategy no longer includes it.

Our goal is to observe the interplay of individual-level learning and collective learning moderated by bounded confidence. Thus, a number of abstractions were made to allow focusing on the core mechanisms of the model:

- Agents do not have a spatial network or activity-based routine. In each time step, agents choose a single other agent they are connected to for interaction.
- While agents know the costs of practices, there is no explicit budget, and the costs are not reflected in the initial belief value. It does, however, impact how agents evaluate their strategies.
- When calculating the payoff of a practice, the only reward weighed against the cost of a practice is the subjective belief about the effectiveness of that practice in protecting from disease. Other rewards, like social rewards for assimilating, are not explicitly modelled.
- The agents encounter a novel disease for which no established knowledge or best practice recommendations exist at the beginning of the simulation. Communication is limited to personal contacts - information and disease travel at the same speed in this simplified model.
- Beliefs are the main basis for social comparisons - our agents are traitless apart from their health status and their individual belief vector.

To facilitate the reading of mathematical expressions, we define the following notations and model parameters:

1. n, m - for agents n and m by unique numeric identifier,

2. i, j - for practices i and j,
3. t, s - for time steps t and s .

- N : Number of agents/nodes in the graph
- i_c : Base probability that an infection spreads from n to m
- i_{cs} : Probability that m is randomly infected without contact
- t_i : Time an agent remains in the I state
- t_r : Time an agent remains in the R state
- r : Weight for individual learning vs. social learning
- β : Magnitude of assimilation of m to n, even if n may have a lower payoff.
- k_b : Similarity threshold for bounded confidence
- k_a : Activity threshold
- r_b : Base payoff agents receive for remaining healthy while doing nothing
- A : Action Space, consisting of $(i, e_i, c_i), i \in \{1 \dots I\}$
- d : Parameter which controls the influence of random, stochastic perturbation
- c_i : cost associated with the practice of measure i
- e_i : degree of true effectiveness of practice i

4.1 Agent Environment and Network

As described in section 2.2, most individuals (referred to as *agents*) have a small number of contacts with local clustering. To model an appropriate network topology, we employ a Watts–Strogatz model with small-world properties[29].

The probability for an agent n to randomly meet another given agent m at time t is uniform among its network neighbours and 0 if they are not neighbors.

In this model, contacts are not reciprocal. While not reflective of reality, this design choice ensures analytical exactness, in which interactions between agents cannot be doubled by agents choosing each other and interacting twice.

4.2 Disease Modelling

To model the spread of an infectious disease across the network, a SIRS model is used. Each agent's health status is at any time one of the three: either susceptible S, infected I or recovered R. Once a susceptible agent gets infected, it remains in the state I for a fixed amount of time (days) t_i before moving on to the status recovered R. This status provides a temporary immunity period t_r after which it transitions back into the S state and may be infected again.

An individual's probability of getting infected is a function of the agent's strategy, i.e., the set of active practices, with effective practices lowering the risk of infection.The degree to which a given practice is effective may be quantified as on a continuous scale between 0 (*no effect at all*) and 1 (100% *effective*). We number the practices $i = 1 \dots I$ with a degree of effectiveness e_i attributed to each practice i and define a vector of *true* effectiveness:$E := (e_1, e_2, e_3, \dots, e_I)$.

This vector describing the degree of protection (which we call *true effectiveness*) of each practice is global and the same for all agents, and doesn't

change over the course of the simulation time. The same goes for the cost vector $C := (c_1, c_2, \ldots, c_i)$.

We assume first a universal probability of infection per contact with an infected person: i_c, which holds when no effective practices at all are being taken by a given agent. This unprotected probability is then reduced individually for each agent n and at each time step t dependent on which practices are implemented by the agent at the time. With E the vector of true effectiveness of practices given in (4.2) and A the time dependent activity vector (3) the individual protected infectibility of each *susceptible* agent is modelled by equation 1 with i_{cs} as the chance of n being randomly infected without contact.

$$i_c(n,t) = i_c \cdot \prod_{i=1}^{I} \left(1 - a_{(n,t,i)} e_i\right) + i_{cs}, \tag{1}$$

4.3 Agent Beliefs and Strategy

Each agent holds *subjective beliefs* about the effectiveness of different practices. As in the case of true effectiveness e_i, believed effectiveness $q_{(n,t,i)}$ of each measure i is quantified on a continuous scale between 0 and 1 as expressed in equation (2). This *believed effectiveness* can differ from $e_i((4.2))$, and agents can either over- or underestimate the effectiveness of different practices.

$$Q(n,t) := (q_{(n,t,1)}, q_{(n,t,2)}, q_{(n,t,3)}, \ldots, q_{(n,t,I)}), \quad q_{(n,t,i)} \in [0,1]. \tag{2}$$

Beliefs can change over time due to interaction with other agents, due to an agent's own experience, or due to exogenous random influences. These three mechanisms are introduced in the next section. At the beginning of the simulation, the belief values in the agents' belief vectors are randomly generated. Accordingly, agents explore a broad range of strategies which are reduced over the runtime through experience and interaction with others.

To implement a practice, agents need to have a sufficient level of faith in the effectiveness of that practice. This level is expressed by the static threshold k_a. Once at belief value falls beneath the given threshold degree the agent ceases to exercise the practice. Hereinafter, we refer to the vector of those practices which an agent actively pursues as it's *strategy*, denoted by $A(n,t)$

$$A(n,t) := (a_{(n,t,1)}, a_{(n,t,2)}, a_{(n,t,3)}, \ldots, a_{(n,t,I)}), \quad a_{(n,t,i)} \in \{0,1\}. \tag{3}$$

where $a_{(n,t,i)} := \begin{cases} 1 & q_{(n,t,i)} \geq k_a, \\ 0 & q_{(n,t,i)} < k_a. \end{cases}$

Just as real people typically can take multiple protective measures at the same time, such as wearing masks in public spaces while also observing a strict personal hygiene and taking homeopathic supplements, agents are also able to choose multiple practices to implement simultaneously. The higher the effort

(cost) necessary to exercise a given practice, the less willing (ready) any agent will be to adopt it. If taking up a new practice takes a lot of effort, then the agent will be reluctant to take that step. On the other hand: If the effort (necessary cost) to try out a new practice is sufficiently low, then an agent, put simply, *might as well decide to try it out* - cost and readiness to adopt are inversely proportional to each other.

4.4 Learning Mechanisms

As laid out above, agents beliefs about practice effectiveness can change due to three mechanisms, described mathematically in 4:

1. **Social Learning:** Agents assimilate beliefs of others, provided their overall beliefs do not conflict too much.
2. **Individual Learning:** Agents draw on their own experience and evaluate whether their strategy was successful in protecting them from infection.
3. **Exogenous Random Influences:** Agents randomly adjust beliefs. This represents unaccounted-for external influences apart from social contacts.

$$Q(n,t{+}1) = Q(n,t){+}r{\cdot}\Delta_{social}Q(n,t){+}(1{-}r){\cdot}\Delta_{learn}Q(n,t){+}\Delta_{stoch}Q(n,t) \quad (4)$$

with the model parameter r setting the ratio of influence between social and individual learning. In our definitions of the two learning effects, Δ_{learn} and Δ_{social}, we make use of the two concepts of *bounded confidence* between two agents and of the *payoff function* related to each agent at each time.

Bounded Confidence in Interactions As described previously, agents have a network on which contacts occur randomly. During these interactions, each agent's beliefs about the effectiveness of protective measures is transparent to others. While selective information-sharing is an interesting phenomenon to examine, it is not part of the current study.

Agents will not be inclined to accept the beliefs of others when the degree of disagreement exceeds a threshold k_b. In our model this degree of disagreement is measured by the Euclidean distance between their respective belief vectors.

$$dist(m,n,t) = \sqrt{\sum_{i=1}^{I} \left(q_{(m,i,t)} - q_{(n,i,t)}\right)^2} \quad (5)$$

Agents must acknowledge at least a minimum of similarity to each other before being willing to consider the experiences of one another. If the distance of beliefs between them exceeds the model threshold k_b, then no social learning interaction between the two agents will take place. This may be modelled by the following threshold function:

$$b(n,m,t) = \begin{cases} 1 & dist(n,m,t) > k_b, \\ 0 & dist(n,m,t) \leq k_b. \end{cases} \quad (6)$$

Payoff function We further introduce the *payoff* $u(n,t)$, an index which incorporates the costs incurred by a given agent when exercising a given strategy at time t, and at the same time imposes a penalty of -1 for the time the agent is infected. At time $t = 0$, before agents start exercising any practices this index $u(n,0) = r_b$, that is a base value. During the simulation, the payoff index goes through a trajectory dependent on changes in strategy and health status:

$$u(n,t) = r_b - \sum_{i}^{I} c_i \cdot a_{(i,n,t)} - h(n,t), \tag{7}$$

where $h(n,t) = \begin{cases} -r_b & \text{if the agent is infected at time } t, \\ 0 & otherwise. \end{cases}$

This ensures that the payoff of agents who pursue only few or only low-cost practices will be higher than that of agents who engage in *all* possible practices.

When two agents m and n interact, the difference in payoffs determines how strongly m will assimilate their own beliefs to n - meaning that agents who either chose very effective low-cost strategies or who were plain lucky so far will exert a stronger influence on their peers than agents who remained healthy due to intense self-protection efforts.

This is modelled mathematically by the following weight function $w(n,m,t)$:

$$w(m,n,t) = 1 \,/\, (1 + e^{-\beta \cdot [u(m,t) - u(n,t)]}) \tag{8}$$

Using the individual agents' payoffs $u(m,t)$ and $u(n,t)$ defined as in (7), which in turn define the weight function $w(n,m,t)$ (8) as well as the bounded confidence threshold function $b(n,m,t)$ (6), the $\Delta Q(n,t)_{social}$ for agent n interacting with m is computed for each practice as follows:

$$\Delta_{social}\, q_{(n,t,i)} = b(n,m,t) \cdot w(m,n,t) \cdot \big(q(m,i,t) - q(n,i,t)\big), \tag{9}$$

$$\Delta_{social}\, Q(n,t) = \big(\Delta_{social}\, q_{(n,t,1)}, \Delta_{social}\, q_{(n,t,2)}, \ldots, \Delta_{social}\, q_{(n,t,I)}\big)\,. \tag{10}$$

Multi-Action Reinforcement Learning In this use case, agents distinguish between two states: being healthy (S or R) and being sick I.

Each day an agent succeeds to remain healthy, their belief in active practices is reinforced by a fraction until the degree of certainty = 1 is reached. Since agents cannot discern which of the practices in their strategy truly led to continued health, all active practices are being reinforced.

Once the agent gets infected, belief in the effectiveness of their strategy is penalised. This means that the belief values of the active practices decrease, while the values for inactive measures remain unchanged. These penalties on failing practices occur in function of their implementation cost - expensive practices will be met with more criticism in case of failure.

Depending again on the agent's payoff function $u(n,t)$ the change in belief may occur in either direction: up or down. This direction is established by (11)

which leads to (12)

$$dir(n,t) = \left(e^{\beta \cdot u(n,t)} - 1\right) / \left(e^{\beta \cdot u(n,t)} + 1\right) \tag{11}$$

$$\Delta_{learn}\, q_{(n,t,i)} = dir(n,t) \cdot a(n,i,t) \cdot (1 - q(n,i,t)) \cdot 0.5, \tag{12}$$

but to ensure that the resulting belief degree $q(n, t+1, i)$ stays within the limits $0 \leq q(n, t+1, i) \leq 1$ we define (13) and (14).

$$\Delta_{learn}\, q_{(n,t,i)} = \begin{cases} \min\left(dir(n,t) \cdot a(n,i,t) \cdot \frac{1-q(n,i,t)}{2}, 1 - q_{(n,t,i)}\right) & \text{if } dir(n,t) \geq 0, \\ \max\left(dir(n,t) \cdot a(n,i,t) \cdot \frac{1-q(n,i,t)}{2}, -q_{(n,t,i)}\right) & \text{if } dir(n,t) \leq 0, \end{cases} \tag{13}$$

$$\Delta_{learn}\, Q(n,t) = \left(\Delta_{learn}\, q_{(n,t,1)}, \Delta_{learn}\, q_{(n,t,2)}, \ldots, \Delta_{learn}\, q_{(n,t,I)}\right). \tag{14}$$

Exogenous random influences We assume that agents may change their beliefs not only due to social interactions and their own experience, but also due to other outside influences not explicitly included in our model. Example of such outside influences could happen when listening to a radio broadcast, watching a TV show, or just walking the streets and seeing street signs with related messaging. The model sometimes gives a random nudge in either direction for beliefs towards individual practices, with the stochastic update Δ_{stoch} being a random value on the interval $[-d, d]$ added to beliefs. The value of this nudge is, as defined in 15, limited to values that would not result in $q(n,t,i)$ exceeding 1.

$$\Delta_{stoch} = \min\left(U[-d,d], 1 - q(n,t,i)\right) \tag{15}$$

5 Methodology and Design of Experiments

We perform experiments to examine the model behaviour in three tiers:

1. *Experiment 1*: Development of Transmission and Agent Behaviour with **one** possible effective practice - in this use case, agents only can choose to engage in regular *Hand-washing* or not.
2. *Experiment 2*: Development of Transmission and Agent Behaviour with **two** practices - one effective (*Hand-washing*) and one ineffective (*Praying*). Agents can choose to engage in both practices concurrently, in only one of the practices, or in none of the two.
3. *Experiment 3*: Development of Transmission and Agent Behaviour with **four** practices: *Hand-washing* is cheap and effective, *Praying* is cheap and ineffective. *Physical Distancing* is effective but also costly. *Drinking Bleach* is both ineffective and costly. In their strategy, the agents can choose to activate any combination of these practices.

Each action space option is combined with the following parameter variation:

1. The ratio of individual and social learning is *balanced* with $r = 0.5$.
2. Agents favour individual learning over social learning with $r = 0.8$
3. Agents favour social learning over individual learning with $r = 0.2$

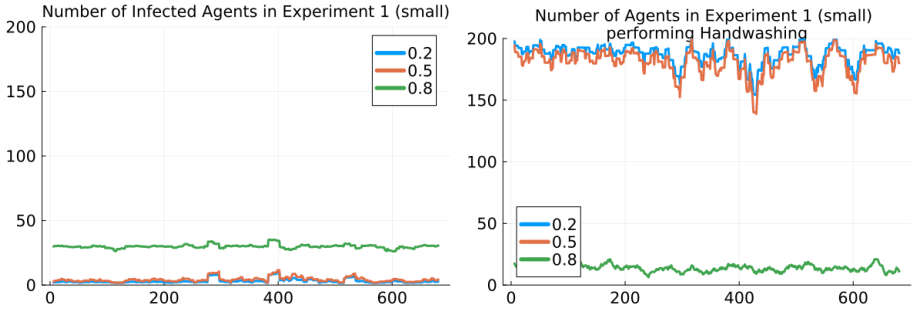
Thus, we are looking at 9 different experimental setups to examine the way reinforcement-based learning and social learning interact and impact the results. Each of these experimental setups is tested on an agent-based model with 200 agents and a model run duration of 700 model steps representing days passed. We record the number of infected agents as well as the number of agents performing available practices at each time step. Further, we run 300 replicates for each setup to rule out coincidence and randomness skewing the results.

6 Results

In *Experiment 1*, we observe that a high value for r, i.e., individual learning being favoured over social learning, leads to significantly higher numbers of infections compared to even a moderately higher degree of social learning, as shown in figures 2a and 2b. Through the inclusion of social learning, the effective strategy is adopted across the population rapidly, greatly reducing the number of infections early into the simulation.

Neither of the two practices reaches widespread adoption under individual learning, though the effective practice is slightly more common than the other one, as seen in figures 3a and 3b. Again, the implementation rates for both practices are very similar for high rates of social learning regardless of effectiveness.

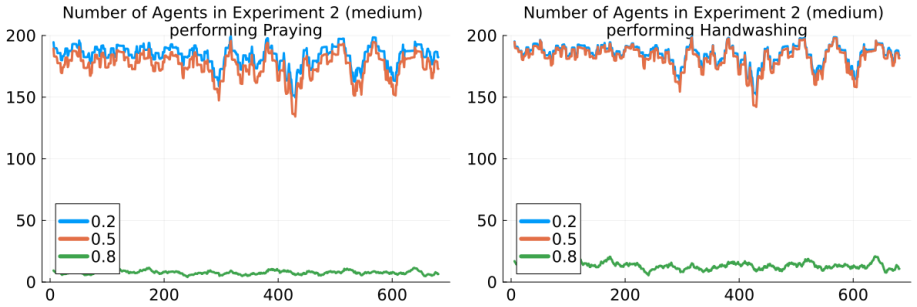
Interestingly, this pattern is not universal. While this is not visible when averaging over 300 model runs, there appears to exist a 'break point' after which agents adopt a practice even for a high r, such as shown in figures 4a and 4b.



(a) Number of infected agents when only *Hand-washing* is available.

(b) Number of agents *Hand-washing*.

Fig. 2: Experiment 1 (single low-cost effective practice): Infection and behaviour results averaged over 300 model runs.



(a) Number of agents *Praying*.

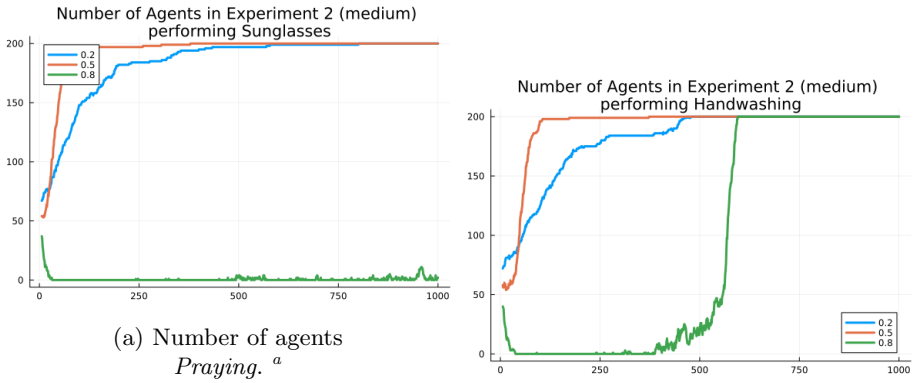
(b) Number of agents *Hand-washing*.

Fig. 3: Experiment 2 (two low-cost practices of different effectiveness): Behaviour results averaged over 300 model runs.

We see a clear point where most agents will cross k_a in favour of hand-washing consistently, leading to an exponential rise in agents performing that action.

Simultaneously, hardly any agent performs the ineffective action, which aligns with the results shown in figure 3a. This result shows that such outlier results may occur. Still, self-learning agents remaining largely inactive appears to be the dominant outcome of most repetitions.

Finally, in *experiment 3*, we can observe that the degree of social learning has a strong impact on the system dynamics in the presence of high-cost options. The two low-cost practices show broadly similar trajectories with low and

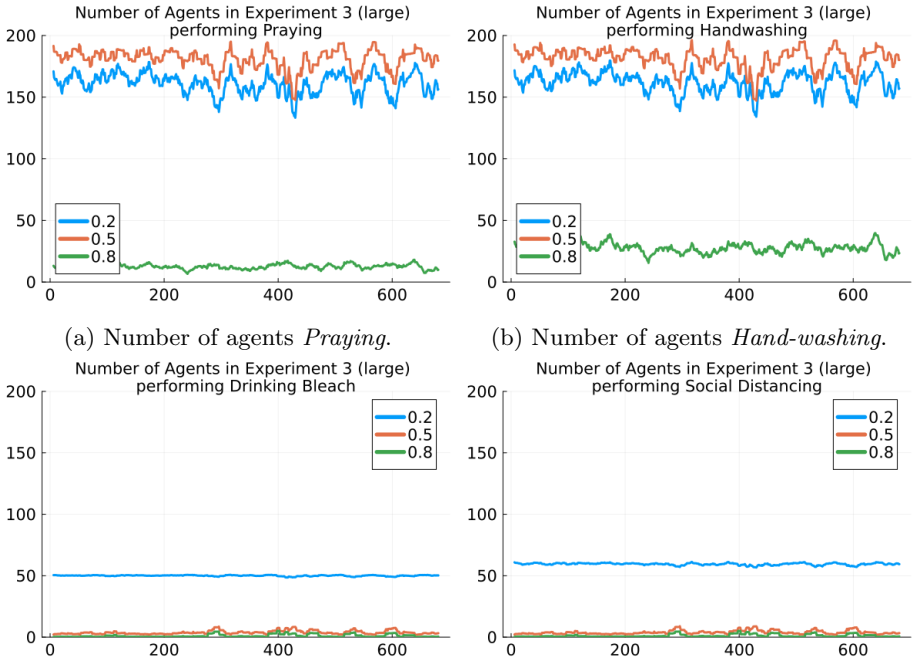


(a) Number of agents *Praying*. [a]

(b) Number of agents *Hand-washing*.

[a] This graphic of early runs still bears the placeholder action name.

Fig. 4: Experiment 2 (two low-cost practices of different effectiveness): Behaviour results of an individual sample model run with 1000 ticks.



(a) Number of agents *Praying*.

(b) Number of agents *Hand-washing*.

(c) Number of agents *Drinking Bleach*.

(d) Number of agents *Physically Distancing*.

Fig. 5: Experiment 3 (four practices of different cost and effectiveness): Behaviour results averaged over 300 model runs.

medium r-values leading to high adoption independent of effectiveness. High r-values hinder widespread adoption, though the effective practiceis chosen slightly more often. For the high-cost practices, the trend is almost reversed: Balanced learning methods will reduce the choice of high-cost practices, regardless of their effectiveness, while social learning being favoured will lead to an ongoing propagation of high-cost practices, albeit to a lesser degree than low-cost ones.

A major factor in these observations is the overall dynamic of the infections which remains similar to the one shown in figure 2a. Social learning, even to a moderate degree, will help flatten the curve rapidly, while purely individualistic learning is slower to reach similar levels of protection in the population.

The results are clearly influenced by model design and protective measures - in the presence of effective low-cost practices, there is hardly a significant benefit to choosing high-cost measures of similar effectiveness instead. Likewise, overly effective methods that stop spread altogether will also eradicate the disease rapidly, preventing the establishment of 'common knowledge'.

7 Discussion and conclusion

Throughout history, social context and past experience have shaped individual and societal response to crises such as epidemics. By combining RL and BCOD, we have presented a model which can explain how mixed learning types can lead to the propagation of ineffective strategies for disease prevention. Our experiments show, cycling back to this work's title, how agents 'learn futility' - especially when it comes to largely benign superstitions.

From a methodological point of view, this combination of two different learning techniques is novel in the space of agent-based social simulation. While Opinion Dynamics are commonly used, reinforcement learning is typically found in other domains of multi-agent systems research. However, this system still needs further refinement. When examining the results of the three different experiments, we can observe that the true efficacy of a practice is overall less important for its adoption than the practice costs and how strongly an agent relies on others' experience. This observation matches not only the historical examples cited at the outset [27, 9], it also fits phenomena observed today. Individual learning is slow to catch on to cheap practices and rapidly eliminates expensive ones, while a high social degree of learning leads to a fast adoption of cheap practices and maintains a certain degree of popularity for expensive strategies. As such, a more refined model is needed to better understand the interplay of the learning techniques by providing an environment in which the individual characteristics of the chosen concepts get to showcase their strengths and weaknesses and how they complement each other as a result.

However, these results still hold some interesting findings - the original hypothesis of this work, that social learning can increase the propagation of ineffective strategies, holds true. Moreover, a balanced ratio of social and individual learning seems to perform best in encouraging the adoption of low-cost strategies while also reducing the spread of potentially harmful high-cost options such as bloodletting, without fully eliminating potentially useful practices.

While we now have advanced scientific methodology that allows us to better infer the health benefits of different interventions, this does not translate into the universal acceptance of beneficial interventions. A prominent example is vaccine hesitancy which is strongly tied to one's social context and narratives about the risks of vaccination, i.e., its potential costs [10]. Conspiracy beliefs about, inter alia, vaccinations are strongly tied to intergroup behaviour [17]. And in a fragmented society, conflicting health practices that are central to group identity will only exacerbate existing fault lines.

In future work, we could adapt our model to even better reflect such a fragmented society with segregated streams of information. An increased tendency to polarise could be modelled by reducing the confidence threshold. This would result in an agent having less other agents to learn from and therefore having to rely more on individual learning. An open question is how the effects of varying the confidence threshold differ from the effects of varying the ratio of individual to social learning.

Beyond the theoretical implication of comparing the effects of changes in these two parameter values, it is obvious that the model is highly sensitive to changes in the parameter values in general. A systematic sensitivity analysis can aid in understanding the central mechanisms of the model. This will not only provide improved insights into the underlying patterns and dependencies between parameters, but also help explaining the model observations from a better validated position.

Overall, computer simulations provide a promising tool set in the exploration of dynamic interaction and the formation of social phenomena. An important line of research on the dissemination of culture demonstrated that individual preferences and exposure in micro-interactions cannot only explain patterns of convergence, but also divergence among individuals and groups [2]. This paper extended this classical research by using more elaborate social learning models, but widely confirms that processes of local convergence can generate cultural variation for the context of biological and social crisis. During epidemics, social learning cannot only explain the spread of effective health interventions, but also destructive cultural variation and the diffusion of ineffective or even harmful health behaviours. Future research should expand and refine our efforts to study the sources of destructive learning behaviours using models that incorporate more complex patterns of seasonality, social structure, and associations among cultural practices.

Acknowledgements. We thank Michael Maes for his substantial support in developing the original version of this model. Bernd Wurpts is grateful for financial support from the University of Lucerne research committee (FoKo). Supported by SEMSAI (German Federal Ministry of Education and Research, Code: 031L0295A) and OptimAgent (German Federal Ministry of Education and Research, Code: 031L0299D).

References

[1] Kevin R. Abbott and Thomas N. Sherratt. "The Evolution of Superstition through Optimal Use of Incomplete Information". In: *Animal Behaviour* 82.1 (July 1, 2011), pp. 85–92. ISSN: 0003-3472.

[2] Robert Axelrod. "The dissemination of culture: A model with local convergence and global polarization". In: *Journal of conflict resolution* 41.2 (1997), pp. 203–226.

[3] ALbert Bandura. *Social foundations of thought and action: A social cognitive theory*. NY: Prentice-hall, 1986.

[4] Jan Beck and Wolfgang Forstmeier. "Superstition and Belief as Inevitable By-Products of an Adaptive Learning Strategy". In: *Human Nature* 18.1 (Mar. 1, 2007), pp. 35–46. ISSN: 1936-4776.

[5] Jan Ole Berndt, Stephanie C. Rodermund, and Ingo J. Timm. "Social Congtagion of Fertility: An Agent-Based Simulation Study". In: *2018 Winter Simulation Conference (WSC)*. 2018 Winter Simulation Conference (WSC). Dec. 2018, pp. 953–964.

[6] Damon Centola. *How behavior spreads: The science of complex contagions*. Vol. 3. Princeton University Press Princeton, NJ, 2018.

[7] Nicholas A Christakis. *Apollo's arrow: The profound and enduring impact of coronavirus on the way we live*. Hachette UK, 2020.

[8] Nicholas A Christakis and James H Fowler. *Connected: The surprising power of our social networks and how they shape our lives*. Little, Brown Spark, 2009.

[9] Michael Emmans Dean. "Selective suppression by the medical establishment of unwelcome research findings: the cholera treatment evaluation by the General Board of Health, London 1854". In: *Journal of the Royal Society of Medicine* 109.5 (2016), pp. 200–205.

[10] Eve Dubé et al. "Vaccine Hesitancy". In: *Human Vaccines & Immunotherapeutics* 9.8 (Aug. 8, 2013), pp. 1763–1773. ISSN: 2164-5515. pmid: 23584253.

[11] Andreas Flache et al. "Models of Social Influence: Towards the Next Frontiers". In: *Journal of Artificial Societies and Social Simulation* 20.4 (2017), p. 2. ISSN: 1460-7425.

[12] Sebastian Funk et al. "The spread of awareness and its impact on epidemic outbreaks". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 106.16 (Apr. 2009), pp. 6872–6877. ISSN: 1091-6490.

[13] Amir Goldberg and Sarah K Stein. "Beyond social contagion: Associative diffusion and the emergence of cultural variation". In: *American Sociological Review* 83.5 (2018), pp. 897–932.

[14] Lysa Gramoli et al. "Needs Model for an Autonomous Agent during Long-term Simulations". In: *2021 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE. 2021, pp. 134–138.

[15] Rainer Hegselmann and Ulrich Krause. "Opinion dynamics and bounded confidence: Models, analysis and simulation". In: *JASSS* 5.3 (2002).

[16] Herbert W. Hethcote. "Qualitative analyses of communicable disease models". In: *Mathematical Biosciences* 28.3 (Jan. 1976), pp. 335–356. ISSN: 0025-5564.

[17] Matthew J. Hornsey et al. "Individual, Intergroup and Nation-Level Influences on Belief in Conspiracy Theories". In: *Nature Reviews Psychology* (Nov. 22, 2022), pp. 1–13. ISSN: 2731-0574.

[18] Matthew O Jackson. *The human network: How your social position determines your power, beliefs, and behaviors*. Vintage, 2019.

[19] L. P. Kaelbling, M. L. Littman, and A. W. Moore. "Reinforcement Learning: A Survey". In: *Journal of Artificial Intelligence Research* 4 (May 1996), pp. 237–285.

[20] R. L. Kendal et al. “Trade-Offs in the Adaptive Use of Social and Asocial Learning”. In: *Advances in the Study of Behavior* 35 (Dec. 1, 2005). ISSN: 0065-3454.

[21] William Ogilvy Kermack and Anderson G McKendrick. “A contribution to the mathematical theory of epidemics”. In: *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115.772 (1927), pp. 700–721.

[22] Veronika Kurchyna et al. “Health and Habit: An Agent-based Approach”. In: *KI 2022: Advances in Artificial Intelligence*. Ed. by Ralph Bergmann et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022, pp. 131–145. ISBN: 978-3-031-15791-2.

[23] Fatik Baran Mandal. “Superstitions: A Culturally Transmitted Human Behavior”. In: *International Journal of Psychology and Behavioral Sciences* 8.4 (2018), pp. 65–69. ISSN: 2163-1956.

[24] Frank M Snowden. *Epidemics and society: From the black death to the present*. Yale University Press, 2019.

[25] Lukas Tapp et al. “School’s out? Simulating schooling strategies during COVID-19”. In: *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer. 2022, pp. 95–106.

[26] Courtney Temple, Ruby Hoang, and Robert G. Hendrickson. “Toxic Effects from Ivermectin Use Associated with Prevention and Treatment of Covid-19”. In: *New England Journal of Medicine* 385.23 (Dec. 2, 2021), pp. 2197–2198. ISSN: 0028-4793. pmid: 34670041.

[27] Joshua Teplitsky. “Imagined Immunities: Medieval Myths and Modern Histories of Jews and the Black Death”. In: *AJS Review: The Journal of the Association for Jewish Studies* 46.2 (2022), pp. 320–346.

[28] David S Touretzky and Lisa M Saksida. “Operant conditioning in Skinnerbots”. In: *Adaptive Behavior* 5.3-4 (1997), pp. 219–247.

[29] Duncan J Watts. *Small worlds: the dynamics of networks between order and randomness*. Vol. 36. Princeton university press, 2004.

[30] Michael Wooldridge and Nicholas R. Jennings. “Intelligent agents: theory and practice”. en. In: *The Knowledge Engineering Review* 10.2 (June 1995). Publisher: Cambridge University Press, pp. 115–152. ISSN: 1469-8005, 0269-8889.

[31] Bernd Wurpts. “Networks into Institutions or Institutions into Networks?” In: *Evidence from the Medieval Hansa* (2018).

[32] Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. “Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms”. In: *Handbook of Reinforcement Learning and Control*. Ed. by Kyriakos G. Vamvoudakis et al. Studies in Systems, Decision and Control. Cham: Springer International Publishing, 2021, pp. 321–384.