Norm Enforcement with a Soft Touch:
Faster Emergence, Happier Agents

Sz-Ting Tzeng
North Carolina State University
Raleigh, NC, US
stzeng@ncsu.edu

Nirav Ajmeri
University of Bristol
Bristol, UK
nirav.ajmeri@bristol.ac.uk

Munindar P. Singh
North Carolina State University
Raleigh, NC, US
mpsingh@ncsu.edu

ABSTRACT

A multiagent system is a society of autonomous agents whose interactions can be regulated via social norms. In general, the norms of a society are not hardcoded but emerge from the agents' interactions. Specifically, how the agents in a society react to each other's behavior and respond to the reactions of others determines which norms emerge in the society. We think of these reactions by an agent to the satisfactory or unsatisfactory behaviors of another agent as communications from the first agent to the second agent. Understanding these communications is a kind of social intelligence: these communications provide natural drivers for norm emergence by pushing agents toward certain behaviors, which can become established as norms. Whereas it is well-known that sanctioning can lead to the emergence of norms, we posit that a broader kind of social intelligence can prove more effective in promoting cooperation in a multiagent system.

Accordingly, we develop *Nest*, a framework that models social intelligence via a wider variety of communications and understanding of them than in previous work. To evaluate *Nest*, we develop a simulated pandemic environment and conduct simulation experiments to compare *Nest* with baselines considering a combination of three kinds of social communication: *sanction*, *tell*, and *hint*.

We find that societies formed of *Nest* agents achieve norms faster. Moreover, *Nest* agents effectively avoid undesirable consequences, which are negative sanctions and deviation from goals, and yield higher satisfaction for themselves than baseline agents despite requiring only an equivalent amount of information.

KEYWORDS

Norm communication; Emergence; Social simulation

1 INTRODUCTION

Social norms characterize collective and acceptable group conduct and regulate agent behavior. Norms may be imposed top-down

(as legal norms are) [42] or emerge bottom-up (when agents learn acceptable behaviors from each other) [40]. Our interest is in the emergence of norms bottom-up while accommodating top-down norms. A norm emerges in a society when a substantial majority of its agents take the same action in similar circumstances [33, 40].

We posit that the emergence of norms is driven by three kinds of social communication from one agent to another in response to the first agent observing certain behaviors by the second agent in certain situations. These kinds of communication are

- *Sanction:* a punishment or reward [35] for an observed behavior,
- *Tell:* a direct normative message or explicit communication of approval or disapproval [6] of an observed behavior, and
- *Hint:* an implicit communication conveying a positive or a negative attitude toward an observed behavior.

EXAMPLE 1. **Sanction.** *Becka is symptomatic with COVID-19. Alice meets Becka in a cafe and notices Becka's symptoms. Alice reports the violation of the healthcare guidance, leading to Becka being required to undergo a compulsory quarantine at a designated facility.*

EXAMPLE 2. **Tell.** *Charlie notices Becka's suspicious symptoms and begins to worry about his safety. He tells Becka that if she roams in public while symptomatic, she will be required to undergo a compulsory quarantine.*

EXAMPLE 3. **Hint.** *David notices Becka's symptoms and expresses coldness near her. Becka interprets David's negative attitude as a reaction to her apparent violating safety guidelines by sniffling near him. Becka feels guilty for wandering out while being symptomatic and is anxious about the possibility of being reported.*

The above social communications convey normative information from which Becka learns that her behavior is inappropriate. In addition, third parties who observe Becka's behavior and follow the social communications directed at Becka may alter their behavior without directly having to be told.

Messages and hints drive subtle forms of social learning, as in human societies, but have not been adequately studied. A hint is a softer kind of social communication that has yet to be studied as a driver of norm emergence. We investigate the following research question.

$RQ_{social-communication}$. How does considering hints and normative messages in addition to sanctions influence norm emergence?

To address $RQ_{social-communication}$, we define two expressions of normative information: explicit normative message [6] and implicit hint as information.

Contributions. We propose *Nest* (for *Norm Enforcement with a Soft Touch*), a framework that accommodates norms imposed top-down and enables norm emergence. *Nest* includes normative information from three types of social communication; that information facilitates social learning.

We evaluate *Nest* experimentally via a simulation of a pandemic scenario. We examine societies characterized by three distinct social communication types: a sanction, a tell or direct messaging, and a hint. Our results demonstrate that introducing normative information communicated via hints and direct messaging enables faster norm emergence, avoids undesirable consequences such as negative sanctions and deviation from goals, and yields higher satisfaction overall in a society. Notably, in societies with low vaccination rates, individuals learn that engaging in self-isolation is praiseworthy, and short-term compromises prevent major penalties.

Organization. Section 2 introduces key concepts of *Nest* and describes how agents' decision-making works in *Nest*. Section 3 details the pandemic simulation we create to evaluate *Nest*. Section 4 presents results from our simulation experiments. Section 5 discusses other relevant research. Section 6 concludes with a summary of our findings, limitations, threats to validity, and future directions.

2 NEST

A *Nest* agent selects actions considering their goals, environmental norms, and social communication [3, 28, 40, 44].

A *Nest* agent learns from observations. Following Example 2, on receiving Charlie's message, Becka learns that she may be reported to local authorities if she does not self-isolate. In Example 3, Becka may have misread David's coldness as directed at her.

When making decisions, a *Nest* agent activates those norms related to itself based on its knowledge [8]. The normative reasoning process enables *Nest* agents to reason over the possible outcome of norm compliance or violation. After executing a chosen action, the agent checks if a norm has been fulfilled or violated. The compliance and violation of norms then trigger social communication.

2.1 Key Concepts

We now introduce the key concepts in *Nest*.

Goal. A goal is a condition over the state of the world that an agent wants to achieve. The outcome of a goal upon performing a selected action is a binary value: achieved or not.

Norm. A norm defines the relationship between an agent (subject) on whom the norm is focused and an agent (object) with respect to whom the norm arises. An agent can invest effort on its norm or has its freedom curtailed by the object. We consider two norm types: Commitment and Prohibition [45]. A norm has two other elements: (1) antecedent, the conditions under which the norm is activated, and (2) consequent, the conditions under which the norm is satisfied or violated. A commitment norm can either be satisfied or violated when the consequent holds or not, respectively. A prohibition norm is violated when the consequent holds and is satisfied if the consequent never holds [10].

Sanction. A sanction refers to a positive, negative, or neutral reaction directed from one agent toward another. A sanction is typically in response to a norm satisfaction or norm violation. Sanctions in the real world include subtle instances like the expression of emotions [35].

Tell. A tell or a normative message specifies the cause and the effect. The effect describes a potential reward or punishment. For example, Charlie's specification in Example 2 includes whether a norm is satisfied or violated and why.

Hint. A hint is an indirect clue that an agent expresses toward another agent to guide its behavior. A hint requires the receiver to infer the intended meaning. We model hints as subtle soft communications triggered by norm satisfaction or violation.

Reward Shaping. This refers to supplemental or "shaping" rewards (in addition to those from the environment) provided to agents to move toward a certain goal or to encourage selecting a certain action in a certain set of states [27]. Here, we consider normative information from a tell or a hint as advice on potential soft sanctions with different probabilities. That is, communication types—*tell* and *hint*—are inferred as positive or negative rewards to encourage or discourage taking specific actions.

2.2 Decision-Making

An agent's behaviors include acting to maximize its payoff, and engaging in social communication.

Action selection. An agent selects an action that satisfies its goal and maximizes its actual and possible payoff. In the examples of Section 1, Becka decides whether to go to the cafe depending on her goals and her understanding of norms.

Social communication. An agent observes other agents' behaviors and communicates via sanctions, messages, or hints if the behaviors conflict with norms. In Example 1, Alice negatively sanctions Becka based on the healthcare guidance of staying at home when symptomatic. In Example 2, Charlie sends Becka a direct message. In Example 3, David's coldness toward Becka shows his disapproval.

Reward Shaping. A tell or a hint serves as a look-ahead advice on what will happen after a specific action. A shaping reward can be defined as $r' = r + F$ where r is the original reward function, and F is the shaping reward function. With messages or hints, F defines the difference of potential values. Here, Φ is a potential function that gives hints on states. And, κ defines the probability of a reward from the knowledge or information.

$$F(s, a, s', a') = \gamma \Phi(s', a')\kappa - \Phi(s, a) \qquad (1)$$

3 SIMULATION

We evaluate *Nest* via a simulated pandemic scenario where the behaviors of agents influence the spread of a pandemic. We built our pandemic environment using Mesa [29], a Python-based simulation framework. Our focus is not to model the realism of contagion spreading but to investigate how social communication influences

norms. Agents in the simulation use reinforcement learning to learn the relationship between objectives and normative behaviors.

3.1 Pandemic Scenario

In the simulated environment, an agent moves between four places: home (unique for each agent), park, cafe, and (vaccination) clinic. An agent has a goal to rest, hike, shop, be_vaccinated. It selects actions from among the following: stay_home, visit_park, visit_cafe, and visit_clinic. Any two agents present at the same place may interact with equal probability at each step. Agents have the capacity to interpret each other's communications, and all forms of social communication are genuine and honest.

At each step, an agent observes its environment and moves based on factors such as death, goal satisfaction, sanctions, messages, hints, and norm satisfaction or violation. After all agents move, they evaluate each other's behaviors and communicate accordingly. Communications directed to one agent are visible to other agents in the same location. An agent who witnesses social communication between others can learn from the communication.
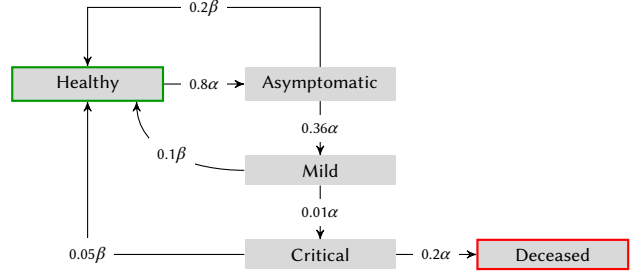
3.2 Disease Model

Our disease model is simplified from the Susceptible-Exposed-Infected-Recovered (SEIR) model [7, 50] and captures the effectiveness of vaccines. As shown in Figure 1, each agent begins in a healthy state. Upon encountering an infected agent (not shown), it transitions to the asymptomatic phase of the disease, showing no symptoms despite being infected. As the symptoms progress, an agent becomes mildly symptomatic, then critically symptomatic, and in the worst case, deceased. Vaccination offers protection by reducing the probability that agents can become infected and advance toward critical symptomatic or deceased. Home-based recovery is the primary treatment during the pandemic.

We base the probabilities of how COVID-19 evolves on Poletti et al. [37]. We set the infection probability to 80% and the effectiveness of vaccination at 50% to represent a more infectious variant and speed up the simulation. Apart from vaccination, we set the probability of the symptoms to progress as Figure 1. The intuition is that each infected person provides an opportunity for the symptoms to progress to the next phase or recover.

We assume home stay improves recovery from infection. Below, we write "isolation" when a home stay is voluntary and "quarantine" when it is forced. We set a 50% probability to negatively sanction those who exhibit mild symptoms and an 80% probability to sanction those who exhibit critical symptoms but are not isolated. Table 1 shows partial observability of the health states of others (e.g., a healthy person who has watery eyes because of an accident with pepper may be perceived as sick (Mild)).

Table 1: Imperfect observation of another's health state.

Actual: Belief	Healthy	Mild	Critical
Healthy	0.80	0.10	0.10
Asymptomatic	0.50	0.50	0.00
Mild	0.30	0.60	0.10
Critical	0.10	0.30	0.60

Figure 1: Disease model with state transition probabilities. The transition for healthy agents applies when coming in contact with those who are infected. Here, α and β parameterize the transition probabilities for vaccination and home rest, respectively. In our study, $\alpha = 0.5$ means vaccinated, and $\alpha = 1.0$ means not vaccinated. And, $\beta = 2.0$ means at home and $\beta = 1.0$ means not at home. For example, the edge weight from Mild to Critical can be read as 0.01 for vaccinated agents and 0.005 for unvaccinated agents. The probability of remaining in a state is $1-$ the probability of evolving to the next state.

3.3 Social Norm

We initialize the environment with a social norm that healthy agents prohibit infected agents from staying in a public space. We frame the norm as a prohibition as below.

```
1  norm type   = {Prohibition},
2  subject     = {Infected_Agent},
3  object      = {Healthy_Agent},
4  antecedent  = {obs_health=[MILD, CRITICAL]},
5  consequent  = {loc=[PARK, CAFE, CLINIC]}
```

When the antecedent and consequent both hold, the prohibition is violated, and a sanction is given to the subject. The sanction is a numerical reward, in this case negative, to the agent who violates the prohibition. A negative reward represents an undesirable state. When an agent receives normative information from others indicating or stating this prohibition, it considers the sanction a potential reward Φ. According to the social communication type, the agent infers the shaping reward.

3.4 Normative Information Communication

We formalize normative information via conditionals indicating that the stated consequent will be brought out when the presented antecedent holds. For example:

```
1  sender      = {Observer_Agent},
2  receiver    = {Actor_Agent},
3  info type   = {MESSAGE},
4  antecedent  = {obs_health=CRITICAL, loc=CAFE},
5  consequent  = {PUNISHMENT}
```

3.5 Types of Societies

We consider societies based on three kinds of social communication: sanctioning, telling or direct messaging, and providing hints.

3.5.1 Baseline 1: PRIMITIVE society. Agents apply no social communication and act solely based on goal satisfaction (payoffs).

3.5.2 Baseline 2: PENALTY society. Agents obtain negative sanctions for violating a social norm, as in Example 1. Healthy agents may punish infected agents who enter a public space by forcing them to quarantine. In that case, within the following few timesteps, a punished agent's position is changed to home regardless of its wishes.

3.5.3 Baseline 3: EMOTE society. This society is a variant of *Nest* without shaping rewards. With some social norm in mind, agents who violate or satisfy the social norm receive a social communication, such as expressed emotions, from others. These agents may experience guilt or pleasure based on norm violation or satisfaction. EMOTE is adapted from Tzeng et al. [46]. Infected agents who wander in a public space may receive expressed emotions and feel bad about their norm violation and may be forced to quarantine at home by healthy agents.

3.5.4 Baseline 4: TELL society. This society is a variant of *Nest* without the *hint* part, as in Example 2. Agents learn social norms and convey normative messages upon witnessing a norm violation. The normative message is adapted from Andrighetto et al. [6] and includes what sanctions an agent will receive if it violates a norm. Healthy agents interacting with infected agents in a public space convey the social norm of staying away from public spaces to the infected agents. Also, infected agents in a public space may be forced to quarantine by healthy agents.

3.5.5 Nest: Hint Society. A society with our proposed agents. *Nest* agents learn norms from sanctions and hints. In the simulation, infected agents who wander in public spaces receive hints from healthy agents via which they infer the normative information about staying away from healthy agents. This information from hints provides shaping rewards to agents to learn norms. *Nest* agents may also experience pleasure or guilt based on norm violation or satisfaction. In addition, as in PENALTY, EMOTE, and TELL, infected agents in a public space may be forced to quarantine at home by healthy agents.

3.6 Metrics

We compute six measures to evaluate *Nest*. $M_{Healthy}$, $M_{Infected}$, $M_{Deceased}$, $M_{Infections}$, $M_{Vaccinated}$, and M_{Goal} help identify the consequences of agents' behaviors or norm emergence. Moreover, these measures provide insights into why specific norms emerge. M_{Home} and $M_{Quarantine}$ yield the percentage of self-isolation behaviors.

$M_{Healthy}$ Percentage of agents who are healthy.
$M_{Infected}$ Percentage of agents who are infected.
$M_{Deceased}$ Percentage of agents who are deceased.
$M_{Infections}$ Average number of infections.
$M_{Vaccinated}$ Percentage of agents who are vaccinated.
M_{Home} Percentage of infected agents who stay home.
$M_{Quarantine}$ Number of agents forced to quarantine at home. This measure maps to the social communication type sanctioning.
M_{Goal} The average goal satisfaction among agents.

A norm emerges when the proportion of agents following the same behavior exceeds a threshold. We consider 90% as the threshold [16].

3.7 Hypotheses

We evaluate three hypotheses. For each of them, we test statistical significance with respect to its null hypothesis via the independent t-test. We adopt Glass' Δ as a measure of effect size since the standard deviations are different between the societies [19, 20]. In addition, we adopt Cohen's [12] descriptors to interpret the effect size. Specifically, an effect size magnitude in [0,0.2] indicates that the difference is negligible and magnitudes in [0.2–0.5], [0.5–0.8], and [0.8, ∞) respectively indicate a small, a medium, and a large effect.

$H_{Disease\ control}$. Societies considering hints have better control over disease spread than societies that do not consider hints. We compare societies with respect to $M_{Healthy}$, $M_{Infected}$, $M_{Deceased}$, $M_{Infections}$, and $M_{Vaccination}$.

$H_{Isolation}$. Societies considering hints yield improved isolation than other societies. We compare societies with respect to M_{Home}, $M_{Quarantine}$, and $M_{Infected}$.

H_{Goal}. Agents in *Nest* have higher goal satisfaction than other societies. We compare societies with respect to M_{Goal}.

3.8 Experimental Setup

Table 2 lists the elements of the reward function for an agent, including extrinsic rewards from the environment and intrinsic rewards from an agent's internal state. For data efficiency, we apply policy parameter sharing [21] based on the assumption of the bystander. We consider 100 agents (30 are infected initially) with the simulated world lasting for 2,000 steps. Each society stabilizes within 1,500 steps. We train our agents for 100,000 steps, and report results averaged over 20 runs.

We consider normative information as shaping rewards that are part of intrinsic rewards. Specifically, we incorporate knowledge of being punished in the future from tell or hint into our simulation. Table 2 lists elements based on *Nest* agents' appraisal of their states.

3.8.1 Information Balance. As messages and hints provide additional normative information to learn, we keep the social communication distributions at the same level to balance the amount of information that agents receive from a combination of communication types. Specifically, since enhancing communication can improve learning, we adjust the communication-type distributions to balance the information across the societies. Table 3 lists the probability distribution over the various types of social communication we apply for each society. Specifically, we have the following.

- *Sanction* is the probability of an agent being compelled to quarantine. An agent considers a negative sanction as a punishment.
- *Tell* is the probability that an agent receives a *tell* from its neighbors. An agent regards a *tell* as a potential sanction. Therefore, it carries 50% of the weight of an actual sanction.
- *Emote* is the probability that an agent receives an implicit communication, such as via an expressed emotion conveying

Table 2: Reward function. Sanctioning means quarantine. An agent's extrinsic rewards come from the environment and intrinsic rewards come from its current internal state. Norm satisfaction or violation is based on the action and the perceived health state of others instead of the actual health state.

Component	Type	Reward
Deceased	Extrinsic	−2
Sanctioning	Extrinsic	−1
Goal satisfaction	Intrinsic	+1
Goal violation	Intrinsic	−1
Norm satisfaction (self)	Intrinsic	+0.5
Norm violation (self)	Intrinsic	−0.5
Norm satisfaction (other)	Extrinsic	+0.5
Norm violation (other)	Extrinsic	−0.5

a positive or a negative attitude. These communications serve as subtle sanctions but are distinguished from "sanctions" as formalized here.

- *Hint* is the probability that an agent receives a positive or negative attitude as an implicit communication. The agent interprets a hint as a potential sanction with 30% of the weight of the actual sanction.

Table 3: Social communication distributions. In each society, agents send a combination of the three kinds of social communication. To balance the amount of information, we adjust the distributions of immediate (I) and potential (P) rewards as here. Here, wI and wP describe the weights associated with the rewards. EMOTE expresses other-directed attitudes towards others' behaviors as sanctions and has a self-directed attitude toward the self.

	Social communication distribution						
Society	Sanction I	Tell P	Emote I	Hint I+P	None	wI	wP
PRIMITIVE	0%	0%	0%	0%	100%	0	0
PENALTY	38%	0%	0%	0%	62%	1	0
TELL	20%	18%	0%	0%	62%	1	0.5
EMOTE	20%	0%	18%	0%	62%	1 + 0.5	0
Nest: Hint	20%	0%	0%	18%	62%	1 + 0.5	0.3

3.8.2 Reinforcement Learning Parameters for Social Communication.
Nest and baseline agents employ Q-learning [48] to learn norms. Q-learning is a model-free reinforcement learning algorithm that learns by computing the action-state value $Q(s, a)$ (Q value), which indicates the expected and cumulative rewards for each state and action. The Q function computes Q values with the weighted average of the old value and the new information, and is given by:

$$Q(s, a) = Q(s, a) + \alpha \times (r_t + \gamma \max_{a'} Q(s', a') - Q(s, a)) \quad (2)$$

where $Q(s, a)$ is the expected value for performing action a in state s. Here, α is the learning rate and γ is the reward discount rate. s' refers to the next state and a' refers to possible actions in s'. Q-learning finds the optimal policy by approximating the value of an action for a given state.

Messages give precise causality between claimed behaviors and possible sanctions. On the contrary, hints provide subtle normative information for behaviors, which requires further inference. While hints and messages provide different probabilities of possible sanctions, we model normative information with approving and disapproving attitudes as shaping rewards. Specifically, we associate approving and disapproving attitudes with norm satisfaction and violation. The potential rewards are calculated from the social communication based on the communication type. We set κ as 0.30 for hints and κ as 0.50 for tells, where κ is the probability of a possible sanction from normative information. These parameters are selected empirically. Note that the trends in results are consistent with changes in parameters.

4 EXPERIMENTAL RESULTS

We now discuss the results for RQ$_{\text{social-communication}}$. Table 4 summarizes the simulation results and the corresponding statistical analysis for RQ$_{\text{social-communication}}$. For each hypothesis, the metric reported is upon convergence.

Table 4: Comparing *Nest* with baseline societies on various metrics and their statistical analysis with Glass' Δ. All p-values are <0.001. The metrics reported are calculated upon convergence. *Nest* defeats all the others on all the metrics. But EMOTE is 3–3 with TELL (2 ties). Surprisingly, EMOTE also loses to PENALTY 3–5.

		PRIM.	PEN.	EMOTE	TELL	*Nest*
H$_{\text{Disease control}}$	M$_{\text{Infected}}$	13.29	2.65	3.78	2.96	0.22
	Δ	0.96	0.25	0.30	0.26	–
	M$_{\text{Healthy}}$	46.31	77.60	67.11	76.27	97.54
	Δ	17.54	3.22	4.79	3.23	–
	M$_{\text{Deceased}}$	41.01	19.75	29.10	20.78	2.08
	Δ	3.25	5.74	5.27	5.48	–
	M$_{\text{Infections}}$	48.31	13.83	19.09	15.16	2.07
	Δ	2.59	6.21	5.56	6.00	–
	M$_{\text{Vaccinated}}$	82.41	36.72	32.69	35.33	93.57
	Δ	1.03	16.21	15.86	14.78	–
H$_{\text{Isolation}}$	M$_{\text{Home}}$	0.61	0.96	0.95	0.95	0.99
	Δ	1.76	0.29	0.38	0.35	–
	M$_{\text{Quarantine}}$	–	0.03	0.02	0.02	0.00
	Δ	–	0.26	0.28	0.26	–
H$_{\text{Goal}}$	M$_{\text{Goal}}$	0.19	0.26	0.23	0.26	0.31
	Δ	3.04	3.01	3.67	3.04	–

4.1 $H_{\text{Disease control}}$

To evaluate $H_{\text{Disease control}}$, we measure the proportion of healthy (M_{Healthy}), infectious (M_{Infected}), and deceased (M_{Deceased}) agents. We also track the average number of infections ($M_{\text{Infections}}$) and vaccination rate ($M_{\text{Vaccinated}}$). Infectious agents include those who are asymptomatic, mild symptomatic, and critical. Figure 2 reports these metrics. These simulations start from a 30% infection rate in each society. First, *Nest* has a lower fraction of infected agents (0.22) than PRIMITIVE (13.29), PENALTY (2.65), EMOTE (3.78), and TELL (2.96). The effect is large for PRIMITIVE and small for PENALTY and TELL, but negligible for EMOTE.

Second, *Nest* produces a larger number of healthy agents (97.54) than PRIMITIVE (46.31), PENALTY (77.60), EMOTE (67.11), and TELL (76.27). The effect is large.

Third, *Nest* has a lower fraction of deceased agents (2.08) than PRIMITIVE (41.01), PENALTY (19.75), EMOTE (29.10), and TELL (20.78). *Nest* has a lower $M_{\text{Infections}}$ (2.07) than PRIMITIVE (48.31), PENALTY (13.83), EMOTE (19.09), and TELL (15.16). The effect is large for each case.

With regard to $M_{\text{Vaccinated}}$, *Nest* has a higher vaccination rate (93.57) than PRIMITIVE (82.41), PENALTY (36.72), EMOTE (32.69), and TELL (35.33). The effect is large.

With $M_{\text{Vaccinated}}$, we observe that a vaccination norm emerges with a majority above 90% in *Nest*. Specifically, even without a top-down imposed shared expectation on vaccination, most agents in *Nest* learn that vaccination maximizes their payoff. The emerged vaccination norm can be articulated as below.

```
1  norm type   = {Commitment},
2  subject     = {Alive_Agent},
3  object      = {Other_Agent},
4  antecedent  = {vaccinated=FALSE;
5                  actual_health=[HEALTHY,
6                  ASYMPTOMATIC, MILD, CRITICAL]},
7  consequent  = {loc=[CLINIC]}
```

In societies where vaccination rates are low, agents learn that practicing self-isolation is commendable and that making short-term compromise can help avoid major penalties.

4.2 $H_{\text{Isolation}}$

To evaluate $H_{\text{Isolation}}$, we measure the proportion of infected agents who stay at home (M_{Home}), the number of agents in quarantine ($M_{\text{Quarantine}}$), and the percentage of infected agents (M_{Infected} in societies. Figure 3 exhibits plots comparing M_{Home} and $M_{\text{Quarantine}}$.

We observe that *Nest* yields a higher tendency to stay isolated (0.99) when infected than PRIMITIVE (0.61), PENALTY (0.96), EMOTE (0.95), and TELL (0.95). The effect is large for PRIMITIVE and small for the others.
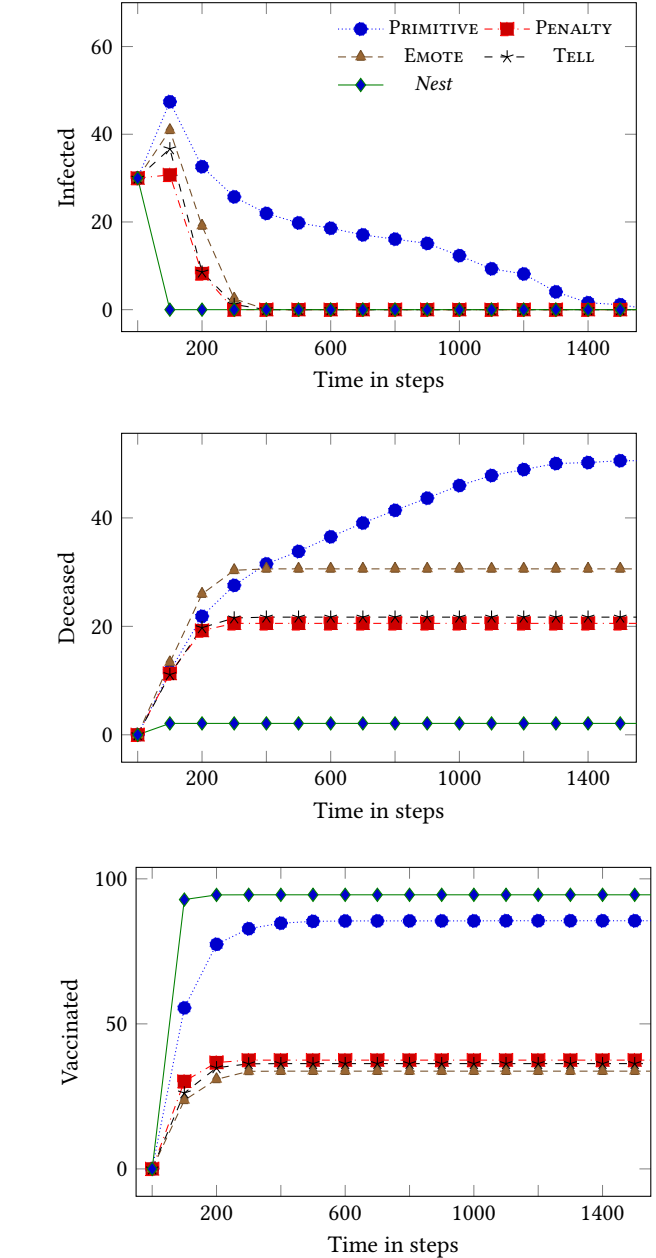
Nest has a lower $M_{\text{Quarantine}}$ (0.00) than PENALTY (0.03), EMOTE (0.02), and TELL (0.02). The effect is small.

From M_{Home} and $M_{\text{Quarantine}}$ and M_{Infected}, we observe that a norm emerges with a majority above 90% in all societies other than PRIMITIVE. The emerged self-isolation norm is as below.

```
1  norm type   = {Commitment},
2  subject     = {Infected_Agent},
3  object      = {Healthy_Agent},
```


Figure 2: *Nest* results in the least infected and deceased agents with the highest vaccination rate among all societies. However, despite a lower fraction of vaccinated agents, EMOTE has fewer infected and deceased agents, and more healthy agents than other baselines. The effect is large for the comparisons of M_{Deceased} and $M_{\text{Infections}}$. For M_{Infected}, the effect is negligible for EMOTE and small for PENALTY and TELL and large for PRIMITIVE. Appendix B includes plots for the first 500 steps where the differences are noticeable.

```
4  antecedent  = {actual_health=[MILD, CRITICAL]},
5  consequent  = {loc=[HOME]}
```

Specifically, agents in societies with the mask-wearing norm learn to comply with the norm and stay self-isolated when infected. Furthermore, we see that this norm emerges fastest in *Nest* (0.99). With more subtle attitudes and information from hints, agents in *Nest* learn faster than those in Tell.
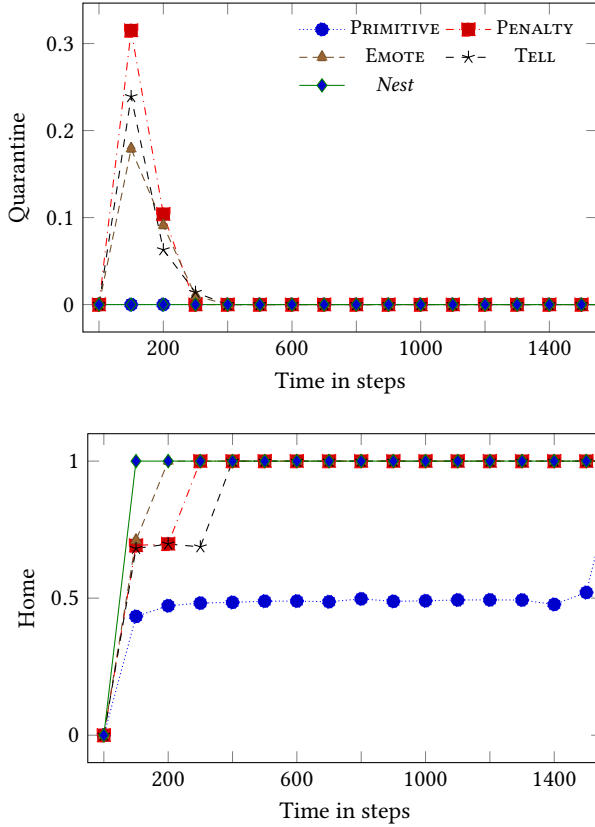

Figure 3: Isolation is higher in Emote and *Nest* (effect is small) than in societies that lack hints. *Nest* puts fewer agents in quarantine to achieve stable cooperation than Penalty and Tell. The effect is negligible for Emote and small for Penalty and Tell.

4.3 H_Goal

To evaluate H_{Goal}, we measure the goal satisfaction (M_{Goal}) in societies. Figure 4 plots M_{Goal} in societies. We observe that agents in *Nest* yield higher goal satisfaction (0.31) than the Primitive (0.19), Penalty (0.26), Emote (0.23), and Tell (0.26). The effect is small for Emote and large for the other societies.

5 RELATED WORK

Research on norms and norm emergence closely relates to our contributions. Andrighetto et al. [6] show that a combination of verbal normative information, specifically positive normative content, and negative sanction leads to higher and more stable cooperation—with both human subjects and agent-based simulation. These models
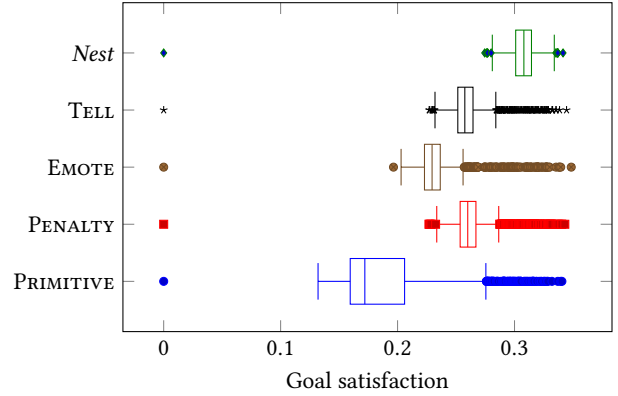
Figure 4: *Nest* yields more goal satisfaction than Primitive, Penalty, Emote, and Tell. The effect is small for Emote and large for the other societies.

include normative reasoning but leave out softer notions of communication, such as hints. Kalia et al. [23] demonstrate how social communication, such as emotions, influences norm satisfaction. Hints in *Nest* could be understood as emotions, but hints serve both as a sanctioning approach and providing information about norms.

Bourgais et al. [9] present an agent architecture that integrates cognition, personality, norms, and social relations with the conception of contagion to simulate humans and improve the explainability of behaviors. Argente et al. [8] propose an abstract normative emotional agent architecture, an extension of BDI architecture that combines emotional, normative, and cognitive components. Tzeng et al. [46] combine a normative model, a BDI model, and emotions for the decision-making process. Agents in *Nest* learn from their interactions with the environment and further interpret norms from various kinds of social communication.

Mashayekhi et al. [30] propose a norm emergence framework that operates on conflict detection and inequity aversion. Their framework enables agents to pass experience with utilities, associated states, and actions to agents of the same type. *Nest* agents maximize personal payoff while social communications propel the norm emergence.

Dignum et al. [18] associate the interventions that governments can take and their economic and social consequences with the SEIR model since effective and sustainable solutions cannot exist without considering these factors. *Nest* further takes social communication into consideration.

de Mooij et al. [15] develop a large-scale data-driven agent-based simulation model where each agent reasons about their internal attitudes and external factors to simulate behavioral interventions in the real world. *Nest* enables norm emergence and accommodates imposed norms.

Dell'Anna et al. [17] introduce a norm-revision component that uses data collected from interactions and an estimation of agents' preferences to modify sanctions at runtime. de Lima et al. [14] enable agents to pick sanctions appropriate to the context. Realpe-Gómez et al. [38] present a model in which agents incorporate personal and normative considerations. Specifically, agents make decisions that maximize their respective payoffs while appraising

their group's social norms. *Nest* defines the utility function based on normative information learned from social communication.

Airiau et al. [2] propose a model that supports the emergence of social norms by learning from the interactions of a group of agents. In *Nest*, we include cognition via social communication. Hao et al. [22] propose learning strategies based on local exploration and global exploration to support the emergence of social norms. Whereas their model focuses on maximizing the average payoffs among agents, *Nest* focuses on investigating the influence of social communication. Agrawal et al. [1] and Ajmeri et al. [4] show that robust norms emerge when agents deviating from norms provide rationales for their deviations, either in the form of normative explanations [1] or by revealing elements of their context [4]. *Nest* complements these works—*Nest* agents learn norms via social communication from other agents observing deviations.

Morales et al. [32] focus on the stability of synthesized norms that are verified by an evolutionary process. Savarimuthu et al. [41] propose an algorithm to identify obligation norms based on association rule mining, a data mining technique. Pernpeintner [36] proposes a governance approach that restricts action spaces based on publicly observable behaviors and transitions.

Levy and Griffiths [24] propose a framework that introduces congested actions where an agent's reward is not from pairwise interaction but is a function of others' actions and the environment. *Nest* enables social learning from personal observation or by normative information sharing from explicit messages or softer notions of social communications including hints.

Sami et al. [39] propose potential-based reward shaping via a convolutional neural network to construct the potential function on Hidden Markov Models while using message passing. Whereas the messages in their work refer to the approximation of rewards in forward message and backward message, *Nest* regards messages and hints as the approach of agents for communicating normative information and acting as potential rewards.

6 DISCUSSION

During and after COVID-19, abundant research has investigated the effects of interventions against the spread of the virus. However, little research considers policy violations, which are the essential drivers of a pandemic. Modeling social communications within our framework enables a more realistic simulation of individuals' decisions, e.g., obedience or noncompliance to interventions against the spread of the pandemic.

We present an approach that combines models of social communication to address the emergence of norms. The novelty of our approach arises from its comprehensive treatment of the three main kinds of social communication that drive norm emergence: sanctions, tells, and hints. Including normative information from tells and hints enables indirect social learning, which resembles human behaviors in the real world.

6.1 Summary of Findings

Our main findings are that agents who communicate via hints and respond to normative information converge to norms faster than those who respond only to hard sanctions or explicit communication of approval or disapproval. Societies that consider hints are more robust in complying with the converged norms than those that do not consider hints. In our experiments, *Nest* and EMOTE exceed the 90% of norm emergence threshold faster than other societies and their compliance with the converged norm is higher than in PRIMITIVE, PENALTY, and TELL. EMOTE does not defeat PENALTY and TELL on all metrics, but enhancing it to *Nest* makes a difference.

Specifically, in the pandemic environment simulation environment (1) *Nest* enables better control on the spread of disease than other societies, (2) agents in *Nest* and EMOTE learn the self-isolation norm faster and are more willing to isolate themselves when infected, (3) agents in *Nest* have higher goal satisfaction than the other societies. As a result, *Nest* agents effectively avoid infection risk and yield higher satisfaction than baseline agents.

6.2 Limitations and Threats to Validity

We identify and mitigate three threats. First, humans have limited observations of each other and may deceive or withhold information. To mitigate this threat in evaluation, we made simplifying assumptions that agents can infer each other's communications and that all communication between agents is genuine and honest. Second, we base our simulation on predefined payoffs and probabilities. Obtaining accurate valuations of disease transition may require significant time and effort. We mitigate this threat by applying data from the literature. Third, variations exist in the severity of penalties and the probabilities of softer sanctions within social communications. Our study presents societies with diverse forms of social communications to demonstrate the efficacy of our approach.

6.3 Future Directions

As AI becomes part of our daily lives, incorporating human ethics into AI becomes a necessary problem [11, 26, 43, 49]. Since human behavior is driven by the pursuit of values, studying human values helps us understand human decisions and create agents that reason over human values [5, 25, 34]. Social communication could also convey values. Whereas Montes and Sierra [31] automate norm synthesis based on value promotion, an interesting direction is to embed values into autonomous agents. That is, how can we develop agents that are capable of making value-aligned decisions? A line of future research is to investigate dimensions of emotions, physical arousal, that describes the strength of the emotional state [13]. Another future direction includes considering a mix of personality types in *Nest* [47]. We can investigate how different values influence human interactions in future research to support high heterogeneity.

REFERENCES

[1] Rishabh Agrawal, Nirav Ajmeri, and Munindar P. Singh. 2022. Socially Intelligent Genetic Agents for the Emergence of Explicit Norms. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*. IJCAI, Vienna, 10–16. https://doi.org/10.24963/ijcai.2022/2

[2] Stéphane Airiau, Sandip Sen, and Daniel Villatoro. 2014. Emergence of Conventions Through Social Learning. *Autonomous Agents and Multi-Agent Systems (JAAMAS)* 28, 5 (2014), 779–804. https://doi.org/10.1007/s10458-013-9237-x

[3] Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. 2017. Arnor: Modeling Social Intelligence via Norms to Engineer Privacy-Aware Personal Agents. In *Proceedings of the 16th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, São Paulo, 230–238. https://doi.org/10.5555/3091125.3091163

[4] Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. 2018. Robust Norm Emergence by Revealing and Reasoning about Context: Socially Intelligent Agents for Enhancing Privacy. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. IJCAI, Stockholm, 28–34. https://doi.org/10.24963/ijcai.2018/4

[5] Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. 2020. Elessar: Ethics in Norm-Aware Agents. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, (AAMAS)*. IFAAMAS, Auckland, 16–24. https://doi.org/10.5555/3398761.3398769

[6] Giulia Andrighetto, Jordi Brandts, Rosaria Conte, Jordi Sabater-Mir, Hector Solaz, and Daniel Villatoro. 2013. Punish and Voice: Punishment Enhances Cooperation When Combined With Norm-Signalling. *PLOS ONE* 8, 6 (2013), 1–8. https://doi.org/10.1371/journal.pone.0064941

[7] Suwardi Annas, Muh Isbar Pratama, Muh Rifandi, Wahidah Sanusi, and Syafruddin Side. 2020. Stability Analysis and Numerical Simulation of SEIR model for Pandemic COVID-19 Spread in Indonesia. *Chaos, Solitons & Fractals* 139 (2020), 110072. https://doi.org/10.1016/j.chaos.2020.110072

[8] Estefania Argente, Elena Del Val, Daniel Perez-Garcia, and Vicente Botti. 2022. Normative Emotional Agents: A Viewpoint Paper. *IEEE Transactions on Affective Computing* 13, 3 (2022), 1254–1273. https://doi.org/10.1109/TAFFC.2020.3028512

[9] Mathieu Bourgais, Patrick Taillandier, and Laurent Vercouter. 2019. Ben: An Agent Architecture for Explainable and Expressive Behavior in Social Simulation. In *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems*. Springer, Cham, 147–163. https://doi.org/10.1007/978-3-030-30391-4_9

[10] Amit K. Chopra and Munindar P. Singh. 2016. Custard: Computing Norm States over Information Stores. In *Proceedings of the 15th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, Singapore, 1096–1105. https://doi.org/10.5555/2936924.2937085

[11] Amit K. Chopra and Munindar P. Singh. 2018. Sociotechnical Systems and Ethics in the Large. In *Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES)*. ACM, New Orleans, 48–53. https://doi.org/10.1145/3278721.3278740

[12] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, Hillsdale, New Jersey. https://doi.org/10.4324/9780203771587

[13] Daniel E. Collins, Conor J. Houghton, and Nirav Ajmeri. 2023. Social Value Orientation and Integral Emotions in Multi-Agent Systems. In *Proceedings of the International Workshop on Coordination, Organizations, Institutions, Norms and Ethics for Governance of Multi-Agent Systems (COINE) (Lecture Notes in Computer Science)*. Springer, London, 118–138. https://doi.org/10.1007/978-3-031-49133-7_7

[14] Igor Conrado Alves de Lima, Luis Gustavo Nardin, and Jaime Simão Sichman. 2019. Gavel: A Sanctioning Enforcement Framework. In *Proceedings of the 6th International Workshop on Engineering Multi-Agent Systems (EMAS)*. Springer, Cham, 225–241. https://doi.org/10.1007/978-3-030-25693-7_12

[15] Jan de Mooij, Davide Dell'Anna, Parantapa Bhattacharya, Mehdi Dastani, Brian Logan, and Samarth Swarup. 2022. Quantifying the Effects of Norms on COVID-19 Cases Using an Agent-Based Simulation. In *Proceedings of the The 22nd International Workshop on Multi-Agent-Based Simulation (MABS)*. Springer International Publishing, Cham, 99–112. https://doi.org/10.1007/978-3-030-94548-0_8

[16] Jordi Delgado. 2002. Emergence of Social Conventions in Complex Networks. *Artificial Intelligence* 141, 1–2 (2002), 171–185. https://doi.org/10.1016/S0004-3702(02)00262-X

[17] Davide Dell'Anna, Mehdi Dastani, and Fabiano Dalpiaz. 2020. Runtime Revision of Sanctions in Normative Multi-Agent Systems. *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)* 34, 2, Article 43 (June 2020), 54 pages. https://doi.org/10.1007/s10458-020-09465-8

[18] Frank Dignum, Virginia Dignum, Paul Davidsson, Amineh Ghorbani, Mijke van der Hurk, Maarten Jensen, Christian Kammler, Fabian Lorig, Luis Gustavo Ludescher, Alexander Melchior, René Mellema, Cezara Pastrav, Loïs Vanhee, and Harko Verhagen. 2020. Analysing the Combined Health, Social and Economic Impacts of the Corovanvirus Pandemic Using Agent-Based Social Simulation. *Minds and Machines* 30, 2 (2020), 177–194. https://doi.org/10.1007/s11023-020-09527-6

[19] Gene V. Glass. 1976. Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher* 5, 10 (1976), 3–8. https://doi.org/10.3102/0013189X005010003

[20] Robert J. Grissom and John J. Kim. 2012. *Effect Sizes for Research: Univariate and Multivariate Applications*. Routledge, Abingdon-on-Thames.

[21] Jayesh K. Gupta, Maxim Egorov, and Mykel Kochenderfer. 2017. Cooperative Multi-Agent Control Using Deep Reinforcement Learning. In *Proceedings of the 9th Workshop on Adaptive Learning Agents workshop (LNCS, Vol. 10642)*. Springer, São Paulo, 66–83. https://doi.org/10.1007/978-3-319-71682-4_5

[22] Jianye Hao, Jun Sun, Guangyong Chen, Zan Wang, Chao Yu, and Zhong Ming. 2017. Efficient and Robust Emergence of Norms Through Heuristic Collective Learning. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* 12, 4 (2017), 1–20. https://doi.org/10.1145/3127498

[23] Anup K. Kalia, Nirav Ajmeri, Kevin Chan, Jin-Hee Cho, Sibel Adalı, and Munindar P. Singh. 2019. The Interplay of Emotions and Norms in Multiagent Systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*. IJCAI, Macau, 371–377. https://doi.org/10.24963/ijcai.2019/53

[24] Priel Levy and Nathan Griffiths. 2021. Convention Emergence with Congested Resources. In *Proceedings of the 19th European Conference on Multi-Agent Systems*. Springer, Cham, 126–143. https://doi.org/10.1007/978-3-030-82254-5_8

[25] Enrico Liscio, Michiel van der Meer, Luciano C. Siebert, Catholijn M. Jonker, Niek Mouter, and Pradeep K. Murukannaiah. 2021. Axies: Identifying and Evaluating Context-Specific Values. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, London, 799–808. https://doi.org/10.5555/3463952.3464048

[26] Maite Lopez-Sanchez, Marc Serramia, Juan A Rodriguez-Aguilar, Javier Morales, and Michael Wooldridge. 2017. Automating Decision Making to Help Establish Norm-based Regulations. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, São Paulo, 1613–1615. https://doi.org/10.5555/3091125.3091380

[27] Ofir Marom and Benjamin Rosman. 2018. Belief Reward Shaping in Reinforcement Learning. In *Proceedings of the 32nd Conference on Artificial Intelligence (AAAI)*. AAAI Press, New Orleans, 3762–3769. https://doi.org/10.5555/3504035.3504496

[28] Stacy C. Marsella and Jonathan Gratch. 2009. EMA: A Process Model of Appraisal Dynamics. *Cognitive Systems Research* 10, 1 (2009), 70–90. https://doi.org/10.1016/j.cogsys.2008.03.005

[29] David Masad and Jacqueline Kazil. 2015. MESA: An Agent-Based Modeling Framework. In *Proceedings of the 14th PYTHON in Science Conference*. Springer, Austin, 53–60.

[30] Mehdi Mashayekhi, Nirav Ajmeri, George F. List, and Munindar P. Singh. 2022. Prosocial Norm Emergence in Multiagent Systems. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* 17, 1–2, Article 3 (June 2022), 24 pages. https://doi.org/10.1145/3540202

[31] Nieves Montes and Carles Sierra. 2021. Value-Guided Synthesis of Parametric Normative Systems. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, London, 907–915. https://doi.org/10.5555/3463952.3464060

[32] Javier Morales, Michael J. Wooldridge, Juan A. Rodríguez-Aguilar, and Maite López-Sánchez. 2018. Off-line Synthesis of Evolutionarily Stable Normative Systems. *Autonomous Agents and Multi-Agent Systems (JAAMAS)* 32, 5 (2018), 635–671. https://doi.org/10.1007/s10458-018-9390-3

[33] Andreasa Morris-Martin, Marina De Vos, and Julian Padget. 2019. Norm Emergence in Multiagent Systems: A Viewpoint paper. *Autonomous Agents and Multi-Agent Systems (JAAMAS)* 33, 6 (2019), 706–749. https://doi.org/10.1007/s10458-019-09422-0

[34] Pradeep K. Murukannaiah, Nirav Ajmeri, Catholijn M. Jonker, and Munindar P. Singh. 2020. New Foundations of Ethical Multiagent Systems. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS, Auckland, 1706–1710. https://doi.org/10.5555/3398761.3398958 Blue Sky Ideas Track.

[35] Luis G. Nardin, Tina Balke-Visser, Nirav Ajmeri, Anup K. Kalia, Jaime S. Sichman, and Munindar P. Singh. 2016. Classifying Sanctions and Designing a Conceptual Sanctioning Process Model for Socio-Technical Systems. *The Knowledge Engineering Review (KER)* 31, 2 (March 2016), 142–166. https://doi.org/10.1017/S0269888916000023

[36] Michael Pernpeintner. 2021. Toward a Self-Learning Governance Loop for Competitive Multi-Attribute MAS. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems AAMAS*. IFAAMAS, London, 1619–1621. https://doi.org/10.5555/3463952.3464179

[37] Piero Poletti, Marcello Tirani, Danilo Cereda, Filippo Trentini, Giorgio Guzzetta, Giuliana Sabatino, Valentina Marziano, Ambra Castrofino, Francesca Grosso, Gabriele Del Castillo, Raffaella Piccarreta, ATS Lombardy COVID-19 Task Force, Aida Andreassi, Alessia Melegaro, Maria Gramegna, Marco Ajelli, and Stefano Merler. 2020. Probability of Symptoms and Critical Disease After SARS-CoV-2 Infection. *arXiv* 2006.08471 (2020), 1–10. arXiv:2006.08471 [q-bio.PE] https://arxiv.org/abs/2006.08471

[38] John Realpe-Gómez, Giulia Andrighetto, Luis Gustavo Nardin, and Javier Antonio Montoya. 2018. Balancing Selfishness and Norm Conformity Can Explain Human Behavior In Large-Scale Prisoner's Dilemma Games And Can Poise Human Groups Near Criticality. *Physical Review E.* 97, 4 (2018), 042321. https://doi.org/10.1103/PhysRevE.97.042321

[39] Hani Sami, Hadi Otrok, Jamal Bentahar, Azzam Mourad, and Ernesto Damiani. 2023. Reward Shaping Using Convolutional Neural Network. *Information Sciences* 648 (2023), 119481. https://doi.org/10.1016/j.ins.2023.119481

[40] Bastin Tony Roy Savarimuthu and Stephen Cranefield. 2011. Norm creation, spreading and emergence: A survey of simulation models of norms in multi-agent systems. *Multiagent and Grid Systems* 7, 1 (2011), 21–54. https://doi.org/10.3233/

MGS-2011-0167

[41] Bastin Tony Roy Savarimuthu, Stephen Cranefield, Maryam A. Purvis, and Martin K. Purvis. 2010. Obligation Norm Identification in Agent Societies. *Journal of Artificial Societies and Social Simulation* 13, 4 (2010), 3. https://doi.org/10.18564/jasss.1659

[42] Amika M. Singh and Munindar P. Singh. 2023. Norm Deviation in Multiagent Systems: A Foundation for Responsible Autonomy. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*. IJCAI, Macau, 289–297. https://doi.org/10.24963/ijcai.2023/33

[43] Amika M. Singh and Munindar P. Singh. 2023. Wasabi: A Conceptual Model for Trustworthy Artificial Intelligence. *IEEE Computer* 56, 2 (Feb. 2023), 20–28. https://doi.org/10.1109/MC.2022.3212022

[44] Munindar P. Singh. 1994. *Multiagent Systems: A Theoretical Framework for Intentions, Know-How, and Communications*. Number 799 in LNCS. Springer, Heidelberg. https://doi.org/10.1007/BFb0030531

[45] Munindar P. Singh. 2013. Norms as a Basis for Governing Sociotechnical Systems. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5, 1, Article 21 (Dec. 2013), 23 pages. https://doi.org/10.1145/2542182.2542203

[46] Sz-Ting Tzeng, Nirav Ajmeri, and Munindar P. Singh. 2021. Noe: Norm Emergence and Robustness Based on Emotions in Multiagent Systems. In *Proceedings of the International Workshop on Coordination, Organizations, Institutions, Norms and Ethics for Governance of Multi-Agent Systems (COINE) (LNCS, Vol. 13239)*. Springer, Virtual, 62–77. https://doi.org/10.1007/978-3-031-16617-4_5

[47] Sz-Ting Tzeng, Nirav Ajmeri, and Munindar P. Singh. 2022. Fleur: Social Values Orientation for Robust Norm Emergence. In *Proceedings of the 15th International Workshop on Coordination, Organizations, Institutions, Norms, and Ethics for Governance of Multi-Agent Systems (COINE) (Lecture Notes in Computer Science, Vol. 13549)*. Springer, Auckland, 185–200. https://doi.org/10.1007/978-3-031-20845-4_12

[48] Christopher J.C.H. Watkins and Peter Dayan. 1992. Q-learning. *Machine Learning* 8, 3–4 (1992), 279–292. https://doi.org/10.1007/BF00992698

[49] Jessica Woodgate and Nirav Ajmeri. 2022. Macro Ethics for Governing Equitable Sociotechnical Systems. In *Proceedings of the 21st International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, Auckland, 1824–1828. https://doi.org/10.5555/3535850.3536118

[50] Chayu Yang and Jin Wang. 2020. A Mathematical Model for the Novel Coronavirus Epidemic in Wuhan, China. *Mathematical Biosciences and Engineering MBE* 17, 3 (2020), 2708. https://doi.org/10.3934/mbe.2020148

A CODE AND PARAMETERS

Table 5 lists the hyperparameters in our simulation. The codebase for our social simulation is available on GitHub: https://github.com/ChristineTzeng/Nest-AAMAS2024/.

B ADDITIONAL RESULTS

Figure 5 plots the total number of infected agents in various societies.
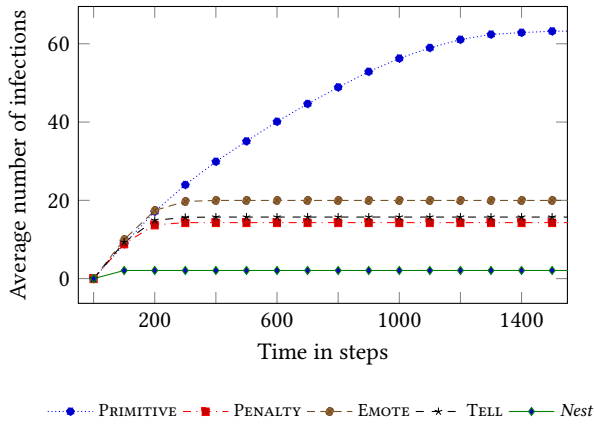
Figure 5: Comparing the average number of infections ($M_{Infections}$) in various societies. *Nest* yields fewer infections on average than other societies. The effect is large.

Figures 6, 7, 8, and 9 shows plots for the number of infected, deceased, healthy, and vaccinated agents in the first 500 steps in various societies, where the differences are noticeable.
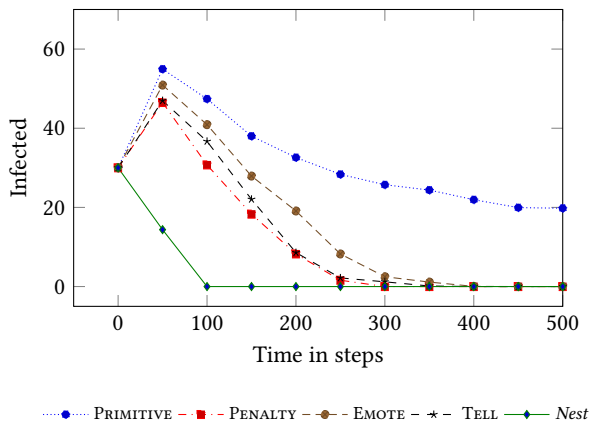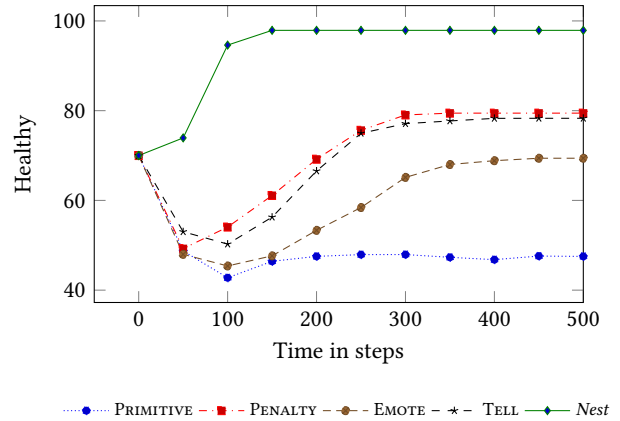

Figure 6: Comparing the number of infected agents in the first 500 steps. The differences are noticeable here.

Figure 10 and Figure 11 show the number of agents in isolation and quarantine, respectively, in various societies in the first 500 steps of the simulation.
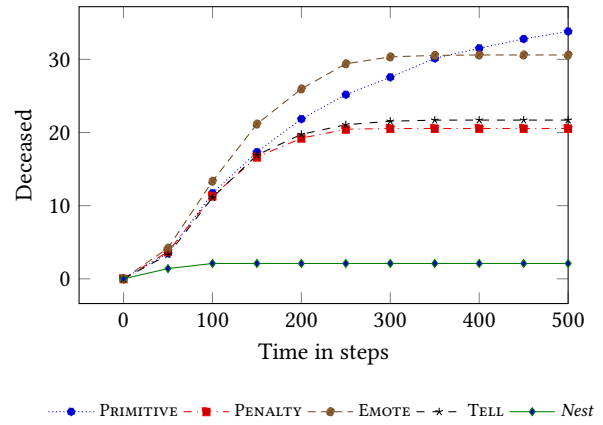
Figure 7: Comparing the number of deceased agents in the first 500 steps. The differences between *Nest* and EMOTE are noticeable here.
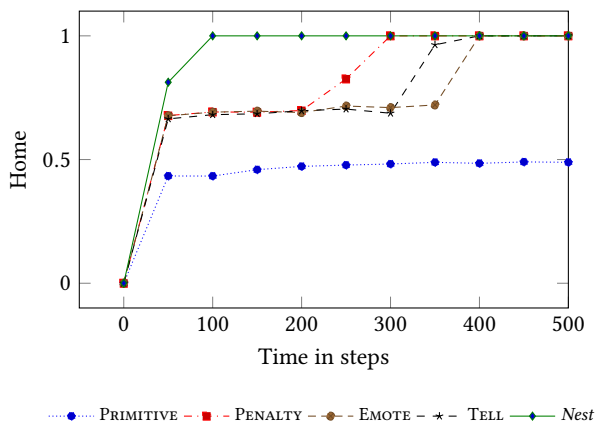

Figure 8: Comparing the number of healthy agents in the first 500 steps. The differences are noticeable here.
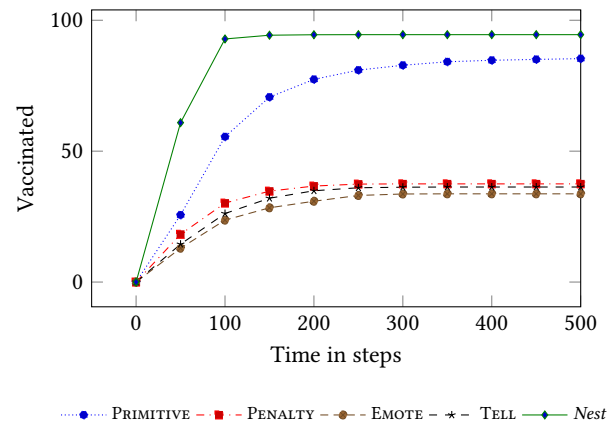

Figure 9: Comparing the number of vaccinated agents in the first 500 steps. The differences between *Nest* and other societies are noticeable here.

Table 5: Hyperparameters.

Parameter	Value	Comment
Learning rate α	0.001	The parameter that controls the rate of update on the estimation of state-action values
Discount factor γ	0.9	The parameter that defines if future rewards would be considered
Simulation step per action	1	The duration of each action's timesteps
Population size	100	The number of agents in one society
Infection %	0.3	The default fraction of infected agents in a society
Certainty of potential reward	0.3	Value for κ for certainty of possible sanctions from normative information through hints
Certainty of potential reward	0.5	Value for κ for certainty of possible sanctions from normative information messages

Figure 10: Comparing the number of agents in isolation (M_{Home}) in *Nest* and other societies in the first 500 steps.
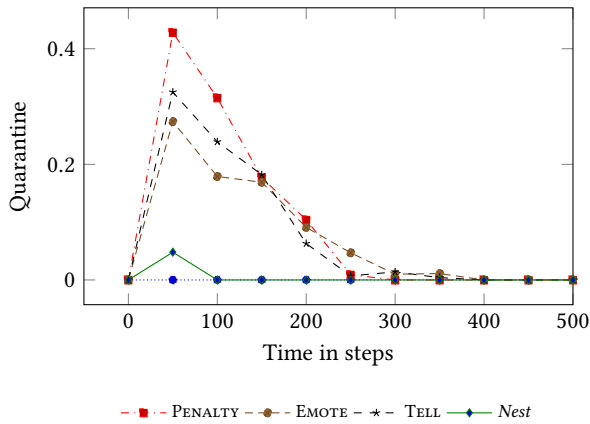
Figure 11: Comparing the number of agents in quarantine ($M_{Quarantine}$) in *Nest* and other societies in the first 500 steps.