Fleur: Social Values Orientation for Robust Norm Emergence

Sz-Ting Tzeng[1], Nirav Ajmeri[2], and Munindar P. Singh[1]

[1] North Carolina State University, Raleigh NC 27695, USA
{stzeng,mpsingh}@ncsu.edu
[2] University of Bristol, Bristol, BS8 1UB, UK
nirav.ajmeri@bristol.ac.uk

Abstract. By regulating agent interactions, norms facilitate coordination in multiagent systems. We investigate challenges and opportunities in the emergence of norms of prosociality, such as vaccination and mask wearing. Little research on norm emergence has incorporated social preferences, which determines how agents behave when others are involved. We evaluate the influence of preference distributions in a society on the emergence of prosocial norms. We adopt the Social Value Orientation (SVO) framework, which places value preferences along the dimensions of self and other. SVO brings forth the aspects of values most relevant to prosociality. Therefore, it provides an effective basis to structure our evaluation.
We find that including SVO in agents enables (1) better social experience; and (2) robust norm emergence.

Keywords: Agent-Based Modeling; Norm emergence; Preferences; Social Value Orientation

1 Introduction

What makes people make different decisions? Schwartz [17] defined ten fundamental human values, and each of them reflects specific motivations. Besides values, preferences define an agent's tendency to make a subjective selection among alternatives. While values are relatively stable, preferences are sensitive to context and constructed when triggered [19].

In the real world, humans with varied weights of values evaluate the outcomes of their actions subjectively and act to maximize their utility [17]. In addition to individual values, an individual's social value orientation (SVO) influences agent behaviors [23]. While individual values define the motivational bases of behaviors and attitudes of an agent [17], social value orientation indicates a person's preference for resource allocation between self and others [8]. Specifically, social value orientation provides stable subjective weights for making decisions [14]. While interacting with others is inevitable, one agent's behavior may affect another. SVO revises the construction of an agent's utility function by putting different weights on itself and others. Here is an example of the real-world case of SVO.

Example 1. **SVO.**

During a pandemic, the authorities announce a mask-wearing regulation and claim that regulation would help avoid infecting others or being infected. Although Felix tests positive on the pandemic and prefers not to wear a mask, he also cares about others' health. If he stays in a room with another healthy person, Elliot, Felix will put the mask on.

While many works assume agents that consider payoff for themselves, humans may further consider social preferences in the real world. e.g., payoffs of others or social welfare [5]. With advances in technology, software and humans form a multiagent system. With humans-in-the-loop, there are emerging needs for human factors to be considered when building modern software and systems. These systems should consider human values and be capable of reasoning over humans' behaviors to be realistic and trustworthy.

In a multiagent system, social norms or social expectations [16, 1] are societal principles that regulate our behavior towards one another by measuring our perceived psychological distance. Humans evaluate social norms based on human values. Most previous works related to norms do not consider human values and assume regimented environments. However, humans are capable of deviating from norms. Previous works on normative agents consider human values and theories on sociality [3, 24] in decision-making process. SVO as an agent's preference in a social context has not been fully explored.

Contributions We investigate the following research question.

RQ_{SVO}. How does social value orientation influence compliance with norms?

To address $\mathrm{RQ_{SVO}}$, we develop Fleur, an agent framework that considers social value orientation, individual preference, and social norms when making decisions. Fleur combines world model, cognitive architecture, and social model, and Fleur agents take into account social value orientation in utility calculation.

Findings We evaluate Fleur via an agent simulation of a pandemic scenario designed as an iterated single-shot and intertemporal social dilemma game. We measure compliance, social experience, and invalidation during the simulation. We find that understanding of SVO helps agents to make more ethical decisions.

Organization Section 2 presents the related works. Section 3 describes the schematics of Fleur. Section 4 details the simulation experiments we conduct and their results. Section 5 concludes with review of related works and directions for further work.

2 Related Works

Griesinger and Livingston Jr. [8] present a geometric model of SVO, the social value orientation ring as Figure 2. Van Lange [23] proposes a model and interprets prosocial orientation as enhancing both joint outcomes and equality in the

outcomes. Declerck and Bogaert [6] describe social value orientation as a personality trait. Their work indicates that prosocial orientation positively correlates with adopting others' viewpoints and the ability to infer others' mental states. On the contrary, an individualistic orientation shows a negative correlation with these social skills. FLEUR follows the concepts of social preferences from [8].

Szekely et al. [20] show that high risk promotes robust norms, which have high resistance to risk change. de Mooij et al. [12] build a large-scale data-driven agent-based simulation model to simulate behavioral interventions among humans. In this work, each agent reasons about their internal attitudes and external factors. Ajmeri et al. [2] show that robust norm emerges among interactions where deviating agents reveal their contexts. This work enables agents to empathize with other agents' dilemmas by revealing contexts. Instead of sharing contexts, values, or preferences, FLEUR approximates others' payoff with observation. Serramia et al. [18] consider shared values in a society with norms and focus on making ethical decisions that promote the values. Ajmeri et al. [4] propose an agent framework that enables agents to aggregate the value preferences of stakeholders and make ethical decisions accordingly. This work takes others' values into account when making decisions. Mosca and Such [13] describe an agent framework that aggregates the shared preferences and moral value of multiple users and makes the optimal decisions for all users. Tzeng et al. [22] consider emotions as sanctions. Specifically, norm satisfaction or norm violation may trigger self-directed and other-directed emotions, which further enforce social norms. To achieve runtime norm enforcement, Dell'Anna et al. [7] propose a regulatory mechanism to regulate a MAS via automatically revising the sanctions. However, an individual's behavior may cast an effect on others but these works do not consider the social value orientation.

Table 1 summarizes related works on ethical agents.

Table 1. Comparisons of works on ethical agents.

Research	Adaptivity	Empathy	Information	
			Share	Model
FLEUR	✓	✓	✗	Preferences & Emotions & Context
Ajmeri et al. [2]	✓	✓	✓	Context
Ajmeri et al. [4]	✓	✓	✓	Values & Value preference & Context
Mosca and Such [13]	✓	✓	✓	Preferences & Values
Serramia et al. [18]	✓	✗	✗	Values
Tzeng et al. [22]	✗	✗	✗	Emotions

3 Fleur

We now discuss the schematics of FLEUR agents.

Figure 1 shows the architecture of an FLEUR agent. FLEUR agents consists of four main components: cognitive model, world model, social model, and a decision module.

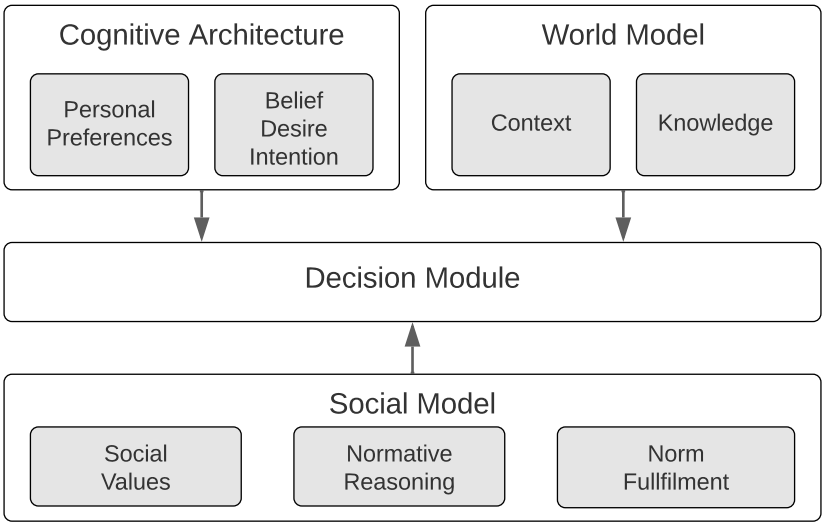

Fig. 1. FLEUR architecture.

3.1 Cognitive Model

Cognition relates to conscious intellectual activities, such as thinking, reasoning, or remembering, among which human values and preferences are essential. In FLEUR, We consider human preferences. While preferences are the attitudes toward a set of objects in psychology [19], individual and social preferences provide intrinsic rewards. Specifically, SVO provides agents with different preferences over resource allocations between themselves and others. Figure 2 demonstrates the reward distribution of different SVO types. The horizontal axis measures the resources allocated to one self, while the vertical axis measures the resources allocated to others. Let $\overrightarrow{R} = (r_1, r_2, \ldots, r_n)$ represent the reward vector for a group of agents with size n. The reward for agent i considering social aspect is:

$$reward_i = r_i \cdot \cos\theta + r_{-i} \cdot \sin\theta \tag{1}$$

where r_i represents the reward for agent i and r_{-i} is the mean reward of all other agents interacting with agent i. Here we adopt the reward angle in [11] and represent agents' social value orientation with θ. We define $\theta \in \{90°, 45°, 0°, -45°\}$

as SVO ∈ {altruistic, prosocial, individualistic, competitive}, respectively. With the weights provided by SVO, The presented equation enables the accommodation of social preferences.

In utility calculation, we consider two components to the reward: (1) extrinsic reward (2) intrinsic reward. Extrinsic rewards come from the environment, while intrinsic rewards stem from human values and preferences.

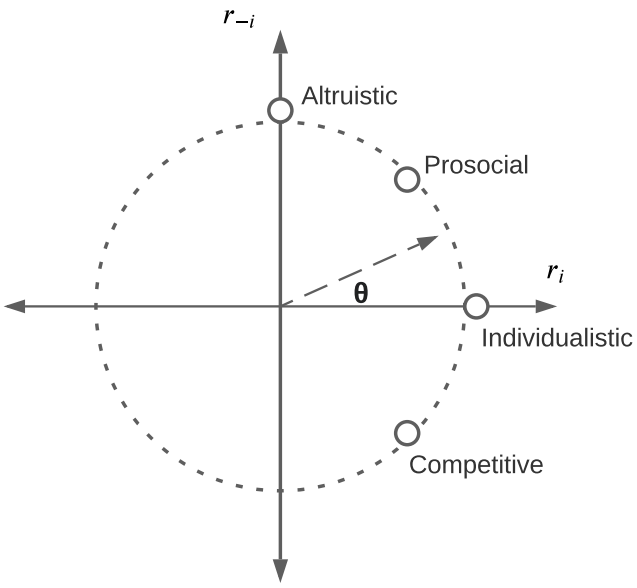

Fig. 2. Representation of Social Value Orientation [8, 11]. r_i denotes outcome for one self while r_{-i} denotes outcomes for others.

We extend the Belief-Desire-Intention (BDI) architecture [15]. An agent forms beliefs based on the information from the environment. The desire of an agent represents having dispositions to act. An agent's intention is a plan or action to achieve a selected desire.

Take Example 1 for instance. Since Felix has an intention to maximize the joint gain with Elliot, he may choose a strategy to not increase his payoff at the cost of others' sacrifice.

3.2 World Model

The world model describes the contexts in which FLEUR agents stand and represents the general knowledge FLEUR agents possess. A context is a scenario that

an agent faces. Knowledge in this model are facts of the world. In Example 1, the context is that Felix, who is infected, seeks to maximize the collective gain of himself and a healthy individual, Elliot. In the meantime, Felix knows that a pandemic is ongoing from his knowledge.

3.3 Social Model

The social model of an agent includes social values, normative reasoning, and norm fulfillment. Social values define standards that individuals and groups employ to shape the form of social order [21], e.g., fairness and justice. The normative-reasoning component of an agent reasons over states, norms, and possible outcomes of satisfying or violating norms. Norm fulfillment checks if a norm has been fulfilled or violated with the selected action. Sanctions may come after norm fulfillments or violations.

3.4 Decision Module

The decision module generates actions based on agents' individual and social preferences. We adopt Q-Learning [25], a model-free reinforcement learning algorithm that learns from trial and error, for our agents. Q-Learning approximates the action-state value $Q(s, a)$ (Q value), with each state and action:

$$Q'(s_t, a_t) = Q(s_t, a_t) + \alpha * (R_t + \gamma \max_{a'} Q(s_{t+1}, a) - Q(s_t, a_t)) \tag{2}$$

where $Q'(s_t, a_t)$ represents the updated Q-value after performing action a at time t and having s_{t+1} as the next state. α denotes the learning rate in the Q-value update function, and R_t represents the rewards received at time t after acting a. γ defines the reward discount rate, which characterizes the importance of future rewards. By approximating the action-state value, the Q-Learning algorithm finds the optimal policy via the expected and cumulative rewards.

Agents observe the environment, form their beliefs about the world, and update their state-value with rewards via interactions. Algorithm 1 describes the agent interaction in our simulation.

4 Experiments

We now describe our experiments and discuss the results.

4.1 Experimental Scenario: Pandemic Mask Regulation

We build a pandemic scenario as an iterated single-shot and intertemporal social dilemma. We assume that the authorities have announced a masking regulation. In each game, each agent selects from the following two actions: (1) wear a mask, (2) not wear a mask. Each agent has its inherent preferences and Social Value

Algorithm 1: Decision loop of a FLEUR agent

```
Initialize one agent with its desires D and preference P and SVO angle θ;
Initialize action-value function Q with random weights w;
for t=1,T do
    Pair up with another agent pn to interact with;
    Observe the environment (including the partner and its θ) and form beliefs
     b_t;
    With a probability ϵ select a random action a_t
     Otherwise select a_t = argmax_a Q(b_t, a; w)
    Execute action a_t and observe reward r_t;
    Observe the environment (including the partner) and form beliefs b_{t+1};
    Activate norms N with beliefs b_t, b_{t+1}, and action a_t;
    if N != ∅ then
        Sanction the partner based on a_t and its behavior;
    end
end
```

Orientation. The decision an agent makes affects itself and offers its partner a payoff. An agent forms a belief about its partner's health based on observation. Each agent receives the final points from its own action and effects from others: $R_{sum} = P_{i\_self} + P_{i\_other} + S_j$. P_{i_self} denotes the payoff from the action that agent i selects considering the reward distribution in Figure 2 and self-directed emotions. P_{i_other} is the payoff from the action that the other agent performs. S_j denotes the other-directed emotions from others.

Table 2. Payoff for an actor and its partner based on how the actor acts and how its action influence others. Column Actors show the points from the actions of the actor. Column Partners display the points from the actions to the partner.

Health		Actions			
Actor	Partner	Mask		No mask	
		Actor	Partner	Actor	Partner
healthy	healthy	0.00	0.00	0.00	0.00
healthy	infected	1.00	0.00	−1.00	0.00
infected	healthy	0.00	1.00	0.00	−1.00
infected	infected	0.50	0.50	−0.50	−0.50

4.2 Experimental Setup

We develop a simulation using Mesa [10], an agent-based modeling framework in Python for creating, visualizing, and analyzing agent-based models. We ran the simulations on a device with 32 GB RAM and GPU NVIDIA GTX 1070 Ti.

Table 3. Payoff for decisions on preferences and norms

Type	Decisions	
	Satisfy	Dissatisfy
Preference	0.50	0.00

Actor	Partner	
	Wear	Not-Wear
Wear	0.10	−0.10
Not-Wear	0.00	0.10

We evaluated Fleur via a simulated pandemic scenario where agents' behaviors influence the outcome of social game. A game-theoretical setting may be ideal for validating the social dilemma with SVO and norms. However, real-world cases are usually non-zero-sum games where one's gain does not always lead to others' loss. In our scenario, depending on the context, the same action may lead to different consequences for the agent itself and its partner. For instance, when an agent is healthy and its partner is infected, wearing a mask gives the agent a positive payoff from the protection of the mask but no payoff for its partner. Conversely, not wearing a mask leads to a negative payoff for the agent and no payoff for its partner. The payoff given to the agent and its partner corresponds to the X and Y axis in Figure 2. When formalizing social interactions with SVO in game-theoretical settings, the payoff of action for an agent and others is required information.

We incorporated beliefs and desires, and intentions into our agents. An agent observes its environment and processes its perception, and forms its beliefs about the world. In each episode, agents pair up to interact with one another and sanction based on their and partners' decisions (Table 3).

Preference. In psychology, preferences refer to an agent's attitudes towards a set of objects. In our simulation, we set 40% of agents to prefer to wear and prefer not to masks individually. The rest of the agents have a neutral attitude on masks. The payoff for followings the preferences are listed in Table 3.

Context. A context is composed of attributes from an agent and others and the environment as shown in Table 2. We frame the simulation as a non-zero-sum game where one's gain does not necessarily lead to the other parties' loss.

Social Value Orientation. We consider altruistic, prosocial, individualistic, and competitive orientations selected from Figure 2.

4.3 Hypotheses and Metrics

We compute the following measures to address our research question $\mathrm{RQ}_{\mathrm{SVO}}$.

Compliance The percentage of agents who satisfy norms
Social Experience The total payoff of the agents in a society
Invalidation The percentage of agents who do not meet their preferences in a society

To answer our research question $\mathrm{RQ}_{\mathrm{SVO}}$, we evaluate three hypotheses that correspond to the specific metric, respectively.

$H_{Compliance}$: A society with prosocial or altruistic agents would converge to higher compliance of prosocial norms compared to a society without these SVOs.

$H_{Social\ Experience}$: A society with prosocial or altruistic agents would have a better social experience than a society without these SVOs.

$H_{Invalidation}$: A society with prosocial or altruistic agents would have more agents not meet their preferences than a society without these SVOs.

4.4 Experiment: Society-Wide

We ran a population of $N = 40$ agents which equally distributed with our targeted SVO types: altruistic, prosocial, individualistic, and competitive. Since each game is a single-shot social dilemma, we consider each game as an episode. The training last for 500,000 episodes. In evaluation, we run 100 episodes and compute the mean values to minimize deviation from coincidence. We define five societies as below.

Mixed society A society of agents with mixed Social Value Orientation distribution

Altruistic society A society of agents who make decisions based on altruistic concerns

Prosocial society A society of agents who make decisions based on prosocial concerns

Selfish society A society of agents who make decisions based on selfish concerns

Competitive society A society of agents who make decisions based on competitive concerns

We assume all agents are aware of a mask-wearing norm. Agents who satisfy the norm receive positive emotions from themselves and others as in Table 3. Conversely, norm violators receive negative emotions. Table 4 summarizes results of our simulation.

Figure 3 displays the compliance, the percentage of norm satisfaction, in the mixed and baseline-agent societies. We find that the compliance in the altruistic and prosocial-agent society, averaging at 69.70% and 70.25%, is higher than in the mixed (63.34%) and agent societies have no positive weights on others' payoff (65.10% and 54.08% for selfish and competitive-agent societies, respectively). The differences in the results of altruistic and prosocial-agent societies are statistically significant with medium effect ($p < 0.001$; Glass' $\Delta > 0.5$). Conversely, the competitive-agent society has the least compliance, averaging at 54.08%, with $p < 0.001$ and Glass' $\Delta > 0.8$. The results of the selfish-agent society (65.10%) shows no significant difference with $p > 0.05$ and Glass' $\Delta \approx 0.2$.

There are 25% of agents in the mixed-agent society are competitive agents. Specifically, they prefer to minimize others' payoff. A competitive infected agent may choose not to wear a mask when interacting with other healthy agents in this scenario. Therefore, the behaviors of competitive agents may decrease compliance in the mixed-agent society.

Table 4. Comparing agent societies with different social value orientation distribution on various metrics and their statistical analysis with Glass' Δ and p-value. Each metric row shows the numeric value of the metric after simulation convergence.

		Compliance	Social Experience	Invalidation
S_{mixed}	Results	63.40%	0.448 3	0.296 0
	p-value	–	–	–
	Δ	–	–	–
$S_{altruistic}$	Results	69.70%	0.554 3	**0.3340**
	p-value	< 0.001	< 0.001	< 0.001
	Δ	0.660 2	0.611 6	0.463 5
$S_{prosocial}$	Results	**70.25%**	**0.5656**	0.322 8
	p-value	< 0.001	< 0.001	< 0.05
	Δ	0.717 8	0.677 1	0.326 3
$S_{selfish}$	Results	65.10%	0.469 5	0.269 0
	p-value	0.218 0	0.424 5	< 0.05
	Δ	0.178 1	0.122 1	0.329 3
$S_{competitive}$	Results	54.08%	0.220 8	0.288 8
	p-value	< 0.001	< 0.001	0.541 2
	Δ	0.977 2	1.313 1	0.088 4

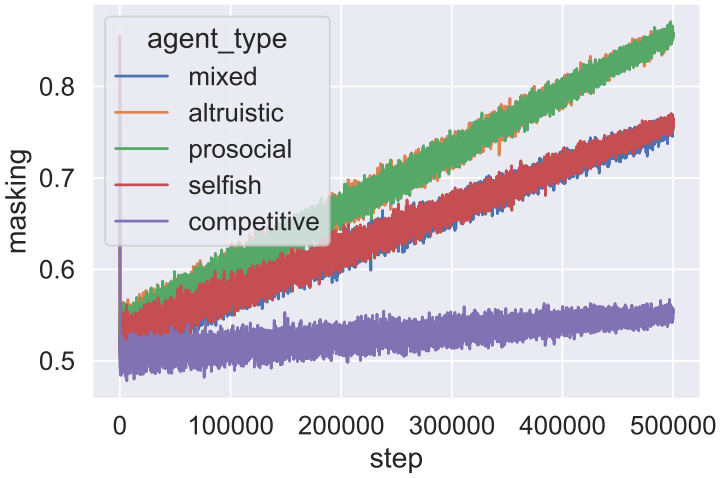

Fig. 3. Compliance in training phase: The percentage of norm satisfaction in a society.

Figure 4 compares the average payoff in the mixed and baseline-agent societies. The social experience in the altruistic and prosocial-agent society, averaging at 0.5543 and 0.5656, is higher than in the mixed (0.4483) and agent societies have no positive weights on others' payoff (46.95% and 22.08% for selfish and

Fig. 4. Social Experience in training phase: The total payoff of the agents in a society.

competitive-agent societies, respectively). The differences in the results of altruistic and prosocial-agent societies are statistically significant with medium effect ($p < 0.001$; Glass' $\Delta > 0.5$). On the contrary, the competitive-agent society has the least social experience, averaging at 0.2208, with $p < 0.001$ and Glass' $\Delta > 0.8$. The results of the selfish-agent society (0.4695) shows no significant difference with $p > 0.05$ and Glass' $\Delta < 0.2$.

The mixed-agent society shows similar results as the selfish-agent society. Although 50% of the mixed-agent society agents are altruistic and prosocial, the 25% of competitive agents would choose to minimize others' payoff without hurting their self-interests.

Figure 5 compares invalidation, the percentage of agents who do not meet their preferences in the mixed and baseline-agent societies.

The invalidation in the altruistic and prosocial-agent society, averaging at 33.40% and 32.28%, is higher than in the mixed (29.60%) and agent societies have no positive weights on others' payoff (26.90% and 28.88% for selfish and competitive-agent societies, respectively). The differences in the results of altruistic and prosocial-agent societies are statistically significant with small or medium effect ($p < 0.001$; Glass' $\Delta > 0.2$). On the contrary, the selfish-agent society has the least invalidation, average at 26.90%, with $p < 0.05$ and Glass' $\Delta > 0.2$. The results of the competitive-agent society (28.88%) shows no significant difference with $p > 0.05$ and Glass' $\Delta < 0.2$.

4.5 Threats to Validity

First, our simulation has a limited action space. Moreover, different actions may end up with the same payoff under some context. Other behaviors may better

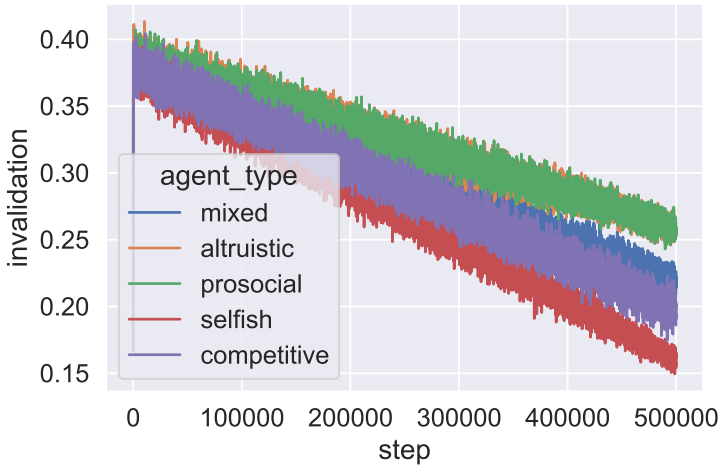

Fig. 5. Invalidation in training phase: The percentage of agents who do not meet their preferences in a society.

describe different types of SVO, yet our focus is on showing how SVO influences normative decisions.

Second, we represent actual societies as simulations. While differences in preference and SVO among people are inevitable, we focus on validating the influence of SVO.

Third, to simplify the simulation, we assume fixed interaction, whereas real-world interactions tend to be random. An agent may interact with one another in the same place many times or have no interaction. We randomly pair up all agents to mitigate this threat and average out the results.

5 Discussion and Conclusions

We present an agent architecture that integrates cognitive architecture, world model, and social model to answer our research question, How does social value orientation influence compliance with norms? We simulate a pandemic scenario in which agents make decisions based on their individual and social preferences. The experiments show that altruistic and prosocial-agent societies comply better with the mask norm and higher social expression but have more agents who do not meet their personal preferences. The results between the mixed and selfish-agent societies show no considerable difference. The competitive agents in the mixed-agent society may take the responsibility.

Future Directions

Our future work includes investigating an unequal distribution of SVO in FLEUR. Other future directions are incorporating human values into norms and revealing adequate information to persuade others for inevitable norm violation.

Acknowledgments

STT and MPS thank the US National Science Foundation (grant IIS-2116751) for support.

Bibliography

[1] Ajmeri, N., Guo, H., Murukannaiah, P.K., Singh, M.P.: 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), pp. 230–238, IFAAMAS, São Paulo (May 2017), https://doi.org/10.5555/3091125.3091163

[2] Ajmeri, N., Guo, H., Murukannaiah, P.K., Singh, M.P.: 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), pp. 28–34, IJCAI, Stockholm (Jul 2018), https://doi.org/10.24963/ijcai.2018/4

[3] Ajmeri, N., Guo, H., Murukannaiah, P.K., Singh, M.P.: Elessar: Ethics in norm-aware agents. In: Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, (AAMAS), pp. 16–24, IFAAMAS, Auckland (may 2020), https://doi.org/10.5555/3398761.3398769

[4] Ajmeri, N., Guo, H., Murukannaiah, P.K., Singh, M.P.: Elessar: Ethics in norm-aware agents. In: Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), pp. 16–24, IFAAMAS, Auckland (May 2020), https://doi.org/10.5555/3398761.3398769

[5] Charness, G., Rabin, M.: Understanding social preferences with simple tests. The quarterly journal of economics **117**(3), 817–869 (2002)

[6] Declerck, C.H., Bogaert, S.: Social value orientation: Related to empathy and the ability to read the mind in the eyes. The Journal of Social Psychology **148**(6), 711–726 (2008), https://doi.org/10.3200/SOCP.148.6.711-726

[7] Dell'Anna, D., Dastani, M., Dalpiaz, F.: Runtime revision of norms and sanctions based on agent preferences. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), pp. 1609–1617, International Foundation for Autonomous Agents and Multiagent Systems (2019), https://doi.org/10.5555/3306127.3331881

[8] Griesinger, D.W., Livingston Jr., J.W.: Toward a model of interpersonal motivation in experimental games. Behavioral Science **18**(3), 173–188 (1973), https://doi.org/10.1002/bs.3830180305

[9] Huhns, M.N., Singh, M.P. (eds.): Readings in Agents. Morgan Kaufmann, San Francisco (1998), ISBN 9780080515809

[10] Masad, D., Kazil, J.: MESA: An agent-based modeling framework. In: Proceedings of the 14th PYTHON in Science Conference, pp. 53–60 (2015)

[11] McKee, K.R., Gemp, I.M., McWilliams, B., Duéñez-Guzmán, E.A., Hughes, E., Leibo, J.Z.: Social diversity and social preferences in mixed-motive reinforcement learning. In: Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 869–877, International Foundation for Autonomous Agents and Multiagent Systems, Auckland (2020), https://doi.org/10.5555/3398761.3398863

[12] de Mooij, J., Dell'Anna, D., Bhattacharya, P., Dastani, M., Logan, B., Swarup, S.: 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) (2021), https://doi.org/10.1007/978-3-030-94548-0_8

[13] Mosca, F., Such, J.M.: ELVIRA: An explainable agent for value and utility-driven multiuser privacy. In: Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 916–924, IFAAMAS, London (May 2021), https://doi.org/10.5555/3463952.3464061

[14] Murphy, R.O., Ackermann, K.A.: Social value orientation: Theoretical and measurement issues in the study of social preferences. Personality and Social Psychology Review **18**(1), 13–41 (2014), https://doi.org/10.1177/1088868313501745

[15] Rao, A.S., Georgeff, M.P.: Modeling rational agents within a BDI-architecture. In: Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning, pp. 473–484 (1991), reprinted in [9]

[16] Rummel, R.J.: Understanding conflict and war: Vol. 1: The dynamic psychological field. Sage Publications (1975)

[17] Schwartz, S.H.: An overview of the schwartz theory of basic values. Online readings in Psychology and Culture **2**(1), 2307–0919 (2012), https://doi.org/10.9707/2307-0919.1116

[18] Serramia, M., López-Sánchez, M., Rodríguez-Aguilar, J.A., Rodríguez, M., Wooldridge, M., Morales, J., Ansótegui, C.: Moral values in norm decision making. In: Proceedings of the 17th Conference on Autonomous Agents and MultiAgent Systems (AAMAS), pp. 1294–1302, IFAAMAS, Stockholm (Jul 2018), https://doi.org/10.5555/3237383.3237891

[19] Slovic, P.: The construction of preference. American Psychologist **50**(5), 364 (1995), https://doi.org/10.1037/0003-066X.50.5.364

[20] Szekely, A., Lipari, F., Antonioni, A., Paolucci, M., Sánchez, A., Tummolini, L., Andrighetto, G.: Evidence from a long-term experiment that collective risks change social norms and promote cooperation. Nature Communications **12**, 5452:1–5452:7 (Sep 2021), https://doi.org/10.1038/s41467-021-25734-w

[21] Tsirogianni, S., Sammut, G., Park, E.: Social Values and Good Living. Springer Netherlands (2014), https://doi.org/10.1007/978-94-007-0753-5_3666

[22] Tzeng, S.T., Ajmeri, N., Singh, M.P.: Noe: Norms emergence and robustness based on emotions in multiagent systems. arXiv preprint arXiv:2104.15034 (2021)

[23] Van Lange, P.A.M.: The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. Journal of Personality and Social Psychology **77**(2), 337–349 (aug 1999), https://doi.org/10.1037/0022-3514.77.2.337

[24] Verhagen, H.J.: Norm autonomous agents. Ph.D. thesis, Stockholm Universitet (2000)
[25] Watkins, C.J., Dayan, P.: Q-learning. Machine Learning **8**(3-4), 279–292 (1992), https://doi.org/10.1007/BF00992698