Value-Based Rationales Improve Social Experience: A Multiagent Simulation Study

Sz-Ting Tzeng[a]**, Nirav Ajmeri**[b] **and Munindar P. Singh**[c]

[a]Umeå University, Umeå, Sweden
[b]University of Bristol, Bristol, UK
[c]North Carolina State University, Raleigh, NC, USA
ORCID (Sz-Ting Tzeng): https://orcid.org/0000-0001-9304-6566, ORCID (Nirav Ajmeri):
https://orcid.org/0000-0003-3627-097X, ORCID (Munindar P. Singh): https://orcid.org/0000-0003-3599-3893

Abstract. We propose *Exanna*, a framework to realize agents that incorporate values in decision making. An *Exanna* agent considers the values of itself and others when providing rationales for its actions and evaluating the rationales provided by others. Via multiagent simulation, we demonstrate that considering values in decision making and producing rationales, especially for norm-deviating actions, leads to (1) higher conflict resolution, (2) better social experience, (3) higher privacy, and (4) higher flexibility.

1 Introduction

A social norm states a shared standard of acceptable behavior in a society [31] and provides a basis for legitimate expectations regarding the behavior of others in the society. Each agent plays two roles: *actor* and *observer*. While exercising autonomy, an actor can deviate from the norms [26]. Such deviations may result in social conflicts and trigger positive or negative sanctions from observers. An acceptable rationale [33] can justify a deviation from a social norm.

Example 1 *Sharing a rationale. Alice wears a mask to the office and notices that Bella is not wearing a mask. Bella justifies her decision by stating that, first, the office has no mask mandate as the surrounding environment is safe. Second, she hates wearing a mask because wearing one gives her eczema. Alice agrees with Bella's view.*

A rationale provides the information to justify a decision [15]. In practice, rationales include additional information that others may be unable to observe, such as the actor's beliefs and preferences. Crafting a rationale remains an ongoing challenge. Rationales may be verbose, leading to information overload. Additionally, they might encompass private information that one may be hesitant to disclose, a concern particularly prevalent in healthcare settings.

Example 2 *Adapting a rationale. Bella and Alice share a concern for health. Despite Bella's aversion to wearing a mask because of eczema, given the safe environment, she feels it unnecessary to disclose her skin condition. Bella rationalizes her behavior of not wearing a mask by stating that the surrounding environment is secure and that a mask is unnecessary. Alice finds Bella's rationale acceptable.*

Values are motivational bases of one's behavior [23]. Reasoning about values is an essential capability to align agents with the values of their stakeholders [34, 35, 36], including providing and recognizing felicitous rationales for one's behaviors [18]. Deliberating over others' values can enhance persuasiveness and foster acceptance.

Instead of sharing all available and related information in a rationale, it is beneficial for an agent to share only the information that aligns well with self and others' values. Sharing such information preserves the privacy of the rationale provider and ensures that unnecessary information does not inundate the observer of a rationale.

Contribution and Findings Accordingly, this paper extends beyond existing research by considering values in decision making and rationale generation and evaluation. Our *Exanna* framework generates rationales that incorporate values and include only the information needed to justify a decision.

We evaluate *Exanna* via a multiagent simulation based on a pandemic scenario. We consider societies of agents with different kinds of rationales: Share-All, Share-Rules, and share value-aligned rules. With *Exanna*, we find that agents who consider value importance when giving rationales exhibit enhanced conflict resolution capabilities. Additionally, rationales aligned with values, albeit with less information provided, contribute to more favorable social experience.

Novelty Although prior research supports constructing explanations or making decisions based on values, this is the first study to investigate how values guide producing and using rationales for norm violation to support norm emergence and improve social experience.

Organization Section 2 discusses relevant related works. Section 3 details the *Exanna* framework. Section 4 describes a simulated pandemic scenario for evaluation. Section 5 demonstrates the results. Section 6 concludes with listing potential future directions.

2 Related Work

Research on agents interacting based on their rationales and modeling values is relevant to our approach.

Agents and Rationales Hind et al. [14] leverage existing supervised machine-learning techniques to generate rationales together with decisions without values involved and without exposing the inner details of the model. Whereas Hind et al. generate rationales based on the existing training set, *Exanna* generates rationales based on context and values.

Georgara et al. [11] show how to build rationales on why specific teams are formed. Specifically, Georgara et al. build rationales based on contrastive explanations and by exploring what-if scenarios. A causal attribution explains why a behavior occurs. We provide causal attribution of the selected action, precisely the premise, as rationale and withhold private information based on values.

Wang et al. [32] formulate rationales with the simplest subset of features that is sufficient as causal attribution for probabilistic solid guarantees on model behavior under observed data distribution. Contreras et al. [7] propose a mirror model and assume a high understandability from performing similar to an observer's mental simulation. They apply deep Q-network and saliency maps in rationale generation, highlighting related input features as rationales. These works reveal model features but not consider values.

Ajmeri et al. [2] propose Poros, a framework that shares full context as a rationale. Therefore, agents can adopt the perspectives of others and make corresponding decisions. However, Ajmeri et al. do not consider values. In *Exanna*, an actor *selectively* shares information based on its *values* and those of the observer.

Agents, Norms, and Values Tzeng et al. [29] define *social communication* (sanction, message, and hint), which besides actual reward or punishment, indicates normative information and potential outcomes. Unlike signaling others with normative information, rationales can enable information sharing and conflict resolution. Tzeng et al. [28] incorporate social value orientation (SVO) in decision making. Whereas values define what is important to agents, SVO describes the importance an agent places on its gain in relation to others. *Exanna* covers a broader range of motivations and behaviors.

Cranefield et al. [8] represent agents' plans to achieve goals as a goal-plan tree and expand the Belief-Desire-Intention language by annotating actions with the effects regarding values. Lera-Leri et al. [16] consider ethical principles (e.g., maximum utility and maximum fairness) for aggregating value systems, not just one value. Ajmeri et al. [3] aggregate users' value preferences to make ethically appropriate decisions. Besides making decisions based on aggregate value importance, *Exanna* agents generate rationales for their decisions with necessary information.

Agrawal et al.'s [1] agent learns norms as rules of optimal behaviors, but considers no values. *Exanna adaptively* shares learned rules as rationales that align with individuals' *values*.

Mosca and Such's [20] agent supports values in multiuser settings by considering the preferences and values of users. The agents justify solutions through contrastive explanations and positive answers. *Exanna* presents factors in alignment with values rather than presenting all that convey causal attribution.

Whereas other works construct explanations with values [33] or make decisions based on values [8], our focus is on building suitable rationales and investigating how value-aligned rationales shape decision making and influence social experience, especially privacy. Previous research considers only the causal links between the behaviors and each contextual factor but not the importance of the factors. Is it essential to include every single related factor? Numerous factors may come into play in real-world situations, yet people typically do not provide or need to provide exhaustive explanations to account for every one of these factors.

Ogunniye and Kökciyan [21] propose an ontology to represent the privacy domain. They introduce argumentation-based multiparty dialogues to reason about contextual norms, and resolve privacy conflicts. Di Scala and Yolum [10] propose an argumentation-based agent to achieve agreement on privacy among multiple users. Their agent provides feedback to the user, such as a summary or detailed advice on possible actions to improve performance. Whereas privacy is a right motivated by values, *Exanna* encompasses broader aspects of core values that serve as guiding principles for decisions and rationales. Additionally, rationales focus on explaining decisions while these works prioritize maintaining privacy among multiple users via argumentation. Aycı et al. [4] explain the privacy decisions (sharing content) using labels (private or public) that are assigned to topics predicted by machine learning. While their values (privacy) come from post-interpretation from humans, *Exanna* incorporates values in rationale construction.

Table 1 summarizes the above comparisons, emphasizing values and rationales. Whereas explanations imply clarity and satisfaction to the recipient, we emphasize justifying a decision based on underlying values without modeling the recipient's comprehension or acceptance of the rationale.

3 Method

We now describe the schematics and decision making in *Exanna* along with its rationale components.

3.1 Schematics of an Exanna Agent

Belief: An agent's view of the world, formed based on observations. b_t indicates the belief at time t. A belief is captured as a set of pairs of attributes and bindings.

Context: The factors that characterize the situation of an agent. Context is represented as a set of attribute-binding pairs. An example of context is as follows.

```
{Risk=None, Preference= ¬Wear,
    InteractWith=Colleague,
    OtherAgentType=Health,
    RiskFromAnother=High,
    Location=Office}
```

A context comprises public (e.g., an agent's location) and private (e.g., beliefs, preferences, and values) factors. Contextual factors may be associated with values (e.g., Risk relates to Health).

Goal: A set of states that an agent wants to achieve. The outcome of a goal after performing the selected actions is binary: achieved or not.

Action: A means to change the state in pursuit of one's goals and maximize associated payoffs.

Preference: A subjective inclination for an action over the alternatives.

Decision rule: A mapping between a premise (set of attribute-binding pairs) and a consequent (an action to be taken). An example rule is

```
{Risk=None, InteractWith=Colleague} =>
    ¬Wear
```

Norm: The expected behavior or the behavior of the majority in a group. When a majority applies the same decision rule, the rule becomes a norm. In *Exanna*, a norm uses the same if-then representation as a decision rule.

Sanction: A response to norm violation or satisfaction. A sanction can be a positive or negative reaction from one agent to another.

Payoff: The benefits an agent receives in a given state after taking an action. Payoffs involve intrinsic benefits (e.g., preferences) and extrinsic benefits (e.g., sanctions imposed by others).

Values: General motivations of agents. Specifically, values define what agents believe to be important, while goals are the desired

Table 1: Summary of comparisons with related work with respect to the application of values in decision making and in the generation and evaluation of rationales. Withhold means that not all factors are presented in the rationale.

	Rationale	Withhold	Values applied in		Rationale representation
			Decision	Rationale	
Cranefield et al. [8]	✗	✗	✓	✗	No rationales provided
Ajmeri et al. [3]	✗	✗	✓	✗	No rationales provided
Lera-Leri et al. [16]	✗	✗	✓	✗	No rationales provided
Tzeng et al. [28]	✗	✗	✓	✗	No rationales provided
Agrawal et al. [1]	✓	✗	✗	✗	Norm as causal attribution but no information hiding
Georgara et al. [11]	✓	✗	✗	✗	Original allocation and another solution with constraints
Contreras et al. [7]	✓	✗	✗	✗	Highlighted input features in deep Q-network but no information hiding
Wang et al. [32]	✓	✗	✗	✗	Prediction and a minimum subset of inputs but no information hiding
Hind et al. [14]	✓	✗	✗	✗	Texts predicted via supervised learning, along with the predicted action
Ajmeri et al. [2]	✓	✗	✗	✗	Full context
Mosca and Such [20]	✓	✗	✓	✓	Suggested action based on inputs from all users and possible outcome of the user's preference as causal attribution, but no information hiding
Winikoff et al. [33]	✓	✗	✗	✓	English mapping of traversed nodes from goal-tree relevant to the explanation
Exanna	✓	✓	✓	✓	Behavior rules (with information hiding) and alignment with values

states. Whereas goals are time-bounded and dynamic to context, values are long-lasting and stable and may transcend contexts [17]. A subset of values is applicable within a context [17], and each agent assigns an importance rating to each value.

Value importance: The importance of values in one context [23]. We store each value importance $V_{context}$ in a tuple where numbers add up to 1. v_i denotes the weight of one value in one value importance ($v_i \in V_{context}$) where $0 \leq v_i \leq 1$ and $\sum_{i=1}^{n} v_i = 1$. We treat each context as an attribute and store the corresponding $\langle V_{context} \rangle$ as its binding. For instance, an agent with value importances $V = \{V_{pandemic} = \{v_{health} = 0.6, v_{privacy} = 0.4\}; V_{normal} = \{v_{health} = 0.4, v_{privacy} = 0.6\}\}$ indicates that the agent values health over privacy during a pandemic but the opposite in a normal context.

3.2 Payoff Calculation with Values

Whereas preferences define the tendency of an individual to make a subjective selection among alternatives, values define the important things to an individual. Although both values and preferences are context-specific, values may transcend contexts [17]. Each agent stores values in a tuple where each value maintains a corresponding $M_{individual}$. Since agents do not make decisions with single values but with tradeoffs among multiple related values, we aggregate value importances when constructing a payoff [3, 22]. Below, f is the aggregated payoff with all corresponding values after selecting strategy Rx when the other player selects strategy Cy from $M_{individual}$.

$$f = \sum_{i}^{values} v_i \times r_{i,RxCy} \qquad (1)$$

We model interactions with payoffs f from the aggregation of $M_{individual}$.

3.3 Interaction and Decision Making

Interactions in *Exanna* are between an actor (rationale provider) and an observer (rationale observer), as shown in Figure 1. An actor selects an action based on its goal and beliefs and provides a rationale to the observer who witnesses its behavior. Upon receiving a rationale from the actor, the observer evaluates the rationale by making

an analogous decision. With a weighted sum of payoffs, we incorporate values in decision making where a substantial value casts a more significant effect on the final decision.
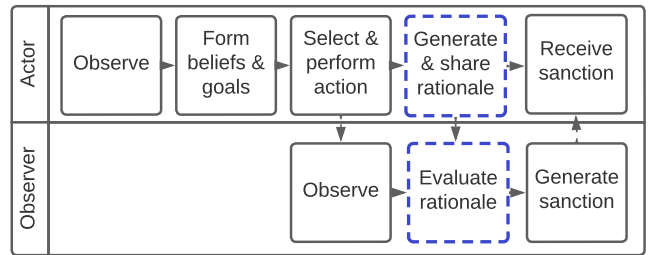
Figure 1: Interactions between *Exanna* agents.

Algorithm 1 gives the pseudo-code of an agent's decision-making loop. An agent forms beliefs b_t about the world based on its observations (Line 4). An agent's payoff is a weighted sum of payoffs corresponding to its values, goal achievement, and factors influencing the decision-making process, such as social circle or social environment. The Q function in Line 7 and reward in Line 8 refer to the payoff calculation in Section 3.2, which incorporate value importances and feedback from others. In Line 7, the agent selects the action that gives the best payoff for b_t. If the agent interacts with another agent, for its action the agent creates rationales based on b_t (including values of both parties) and the selected action (Line 11 with Algorithm 2) and sends those rationales. Other agents who observe the action and receive the rationales update their beliefs, evaluate the rationales (Algorithm 3) with their context, and give sanctions accordingly. Algorithm 4 defines a function to retrieve value importance from beliefs based on context. *Exanna* applies to scenarios that can be modeled as a partially observable Markov decision process and with clearly defined links between values and contextual factors.

3.4 Rationale Generation

Rationale generation in *Exanna* follows a *rule learning* process— a process of evolving rules from datasets or interactions. The basic form of a rule is *if premise then consequent*, where the consequent holds whenever the premise is true. We adapt XCS [5], which applies a genetic algorithm and reinforcement learning to evolve a set of

Algorithm 1 Decision making for an *Exanna* agent

1: Initialize agent (including value importances V and other mental states)
2: Initialize rule-value function Q
3: **for** $t \in \{1, \ldots, T\}$ **do**
4:    Form beliefs b_t based on perceived state
5:    Identify action space A
6:    $V_{actor}, V_{observer} = \text{GetValueImportance}(b_t)$
7:    With a probability ϵ select a random action $a_{actor} \in A$
      Otherwise select $a_{actor} = argmax_a Q(b_t, a, V_{actor})$
8:    Execute action a_{actor} and observe reward r_t
9:    **if** any observer agent *observer* exist **then**
10:       /* Generate rationales based on selected action and beliefs */
11:       $Rat_{actor} = \text{GenRationale}(b_t, a_{actor})$
12:       Send Rat_{actor} to *observer*
13:       Receive sanction $sanction_{observer}$ from *observer*
14:       Observe agent *observer*'s action $a_{observer}$
15:       **if** Receive rationales $Rat_{observer}$ from agent *observer* **then**
16:          Update beliefs b_t based on $Rat_{observer}$
17:       **end if**
18:       /* Generate sanctions based on beliefs and rationales */
19:       $sanction_{actor} = \text{EvalRationale}(Rat_{observer}, a_{observer}, b_t)$
20:    **end if**
21:    /* Agents learn from reward and sanction */
22:    learn($b_t, a_{actor}, r_t + sanction_{observer}, b_{t+1}$)
23: **end for**

---

rules or strategies based on payoffs or rewards produced by the proposed actions. Unlike other machine learning techniques, XCS generates a set of rules describing its decision. XCS process enables flexibility for the implementation of norms and supports interpretability by producing logical rules. An example rule of Example 2 is

```
{Risk=None, InteractWith=Colleague} => ¬Wear
```

The premise of a learned rule is a conjunction of attribute-binding pairs, e.g., {Risk=None, InteractWith=colleague}. Its consequent is an action to be taken when the premise holds—in the above example, ¬Wear. Each rule associates (1) a *fitness*, i.e., its suitability, (2) *numerosity*, i.e., the number of its instances in the rule set, (3) the expected *reward* if the rule applies, and (4) *prediction error*.

3.4.1 XCS for Rationale Generation Briefly

The key features of XCS are *Rule Discovery*, *Rule Subsumption*, and *Action Selection*. Rule discovery through the crossover and mutation processes involves introducing randomness to the antecedent by adding or removing factors, thereby generating rules that are more general or more specific. Given two rules, if the more general one exhibits lower predictive error within the given context, the algorithm retains it and discards the more specific one. When selecting an action, the algorithm selects the one with the best-aggregated fitness. Details on XCS are in Appendix A.

An example of a rationale for not wearing a mask is {Risk=None, Preference=¬Wear, InteractWith=colleague}. This rule means the mask is not needed when there is no infection risk and the actor prefers not to wear a mask while interacting with a colleague. Each agent keeps the rules it discovers and evolves those in a rule set for decision making.

3.4.2 Generating Value-Based Rationale

Not all factors of a rule generated by the rule learning process are suitable for inclusion in a rationale. In Example 2, sharing personal preferences is unnecessary when both agents value health. After generating the base rule, we post-process its factors using the values of the actor and observers. Thus, the agent who prefers the value of health adjusts its rationale for the colleague who also cares about health to a health-related causal attribution if it exists. For instance, no mask is required because there is no risk of infection when interacting with a colleague.

Algorithm 2 details the process of constructing rationale. An agent identifies its rules associated with beliefs b_t (Line 3) and the selected action (Line 4) and then aggregates all rules related to a rationale (Line 5). To minimize information exposure, an agent reveals private information only if it is associated with its values or those of others involved in the interaction (Line 8–9). For instance, if an agent who cares about freedom interacts with one who cares about freedom, it will exclude the infection risk from the environment in its rationales. For each rationale, an agent estimates privacy based on the least proportion of private factors included in the rationale (Line 13).

Algorithm 2 Rationale generation

Input: beliefs b_t, Action a
Output: Rationale Rat
Function: GenRationale

1: Identify private factors $P \subseteq b_t$
2: /* Generate associated rules with beliefs b_t */
3: Get match set ms with b_t
4: Generate action set from ms with a
5: Aggregate rules Rat associated with action set
6: $V_{actor}, V_{observer} = \text{GetValueImportance}(b_t)$
7: **for** factor in Rat **do**
8:    **if** factor $\in P$ and factor not related to V_{actor} and $V_{observer}$ **then**
9:       remove factor from Rat
10:   **end if**
11: **end for**
12: /* Compute privacy based on proportion of private factors */
13: privacy $= 1 - \frac{\#\text{private factors revealed in } Rat}{\#\text{private factors in } Rat}$

3.5 Rationale Evaluation

On receiving a rationale from the actor, the observer first updates its beliefs based on the rationale. Specifically, the observer updates the beliefs of unobservable information from the actor's context. In the rationale generation mask example, the observer updates its beliefs of the infection risk to "None". When evaluating a rationale, the observer makes an analogous decision based on the updated beliefs. If the observer's computed action matches the actor's observed action in that context, the observer accepts the actor's rationale.

Algorithm 3 defines how agents evaluate rationales. Initially, in Line 2, the observer updates its beliefs b_t based on the private context or beliefs included in the actor's rationale. Subsequently, the agent makes a decision analogous to the actor's context based on the updated beliefs. Specifically, with the provided rationale, an agent checks if any applicable rules align with its rule sets in Line 3. The agent identifies associated rules from b_t and adds them to applicable rules in Line 4. In Line 7, the agent calculates the fitness for each available action for each applicable rule and keeps the best action

for each rule. The agent accepts this rationale if any selected action matches the observed action.

Algorithm 3 Evaluating a rationale

Input: Rationales Rat, Observed action $a_{observer}$, Beliefs b_t
Output: Decision d
Function: EvalRationale

1: Initialize applicable rules ars
2: Update b_t with private information in Rat
3: Add triggered match set from Rat to applicable rules ars
4: Add triggered match set from b_t to applicable rules ars
5: **for** $rule$ in applicable rules ars **do**
6:    **for** act in possible actions **do**
7:       calculate fitness f_{act}
8:    **end for**
9:    Keep the act with best fitness
10: **end for**
11: **if** act contains $a_{observer}$ **then**
12:    Decision d = accept;
13: **else**
14:    Decision d = reject;
15: **end if**

Algorithm 4 Get value importance

Input: Beliefs b_t
Output: Value importance $V_{actor}, V_{observer}$
Function: GetValueImportance

1: $context = get\_context(b_t)$
2: /* Retrieve actor and observer's value importances from b_t based on $context$ */
3: $V_{actor}, V_{observer} = get\_value\_importance(b_t, context)$
4: **if** no $observer$ exist **then**
5:    $V_{observer} = null$
6: **end if**

---

4 Simulation

Real-world factors may be associated with underlying values. For instance, health state corresponds to health concerns. We assume the importance of factors according to their associated values.

We evaluate *Exanna* via a pandemic scenario based on Examples 1 and 2 and simulated using MASON [19]. Here agents move to various places, interact with other agents, decide to wear or not wear a mask, and provide a justification for their actions.

4.1 Scenario

The environment represents a multiagent society with several places and social circles. Our environment involves a finite population of 200 agents with different social circles. The environment includes one park, one hospital, five homes, five offices, and five parties. Agents move around and interact within these five places. Each agent is native to one home, one office, and one party. Agents in the same home, office, or party share the same family, colleague, or friend social circle. Each social circle has 40 agents. Time is represented in steps. Each agent moves to one place at each step and has a probability (50%) of interacting with one agent at the same place. Agents are more likely (75%) to move to places they are associated with when

they move to home, office, and party, i.e., an agent is more likely to visit their own home rather than someone else's home.

Each agent forms its goal based on its value importances. Specifically, each value in one context has a payoff matrix (Table 5 and 6); the weighted sum of the payoff determines the goal (desired states). An agent selecting an action that does not align with its goal is considered deviating from its goal.

In the simulated environment, when an agent encounters another agent at the same place, it chooses an action based on its goal—whether to wear a mask. In addition, the agent justifies its behavior based on its beliefs in that context. For instance, the agent gives a rationale—{Risk=None, InteractWith=Colleague}—while not wearing a mask. The beliefs of an agent include public and private factors. Each agent receives a payoff according to the interaction place for action selection, as in Table 2. Wearing a mask at a hospital during a pandemic is desirable. Place and value importances determine the payoff an agent gives to itself. An agent also gives sanctions as feedback to others based on their actions.

The sanctions are based on the social circle. Table 3 lists the sanctions associated with social circles. We run each simulation 10 times, with each run lasts 30,000 steps. We consider the values of freedom and health. In this setting, freedom refers to agents claiming their *free will* and adhering to their preferences.

Table 2: Actor's payoff based on the place. Numbers reflect general expectations of places.

Places	Wear	¬Wear
Home	−0.25	0.25
Office	0.25	−0.25
Party	−0.25	0.25
Park	−0.50	0.50
Hospital	0.50	−0.50

Table 3: Feedback from an observer based on social circle.

Social Circle	Observer move	
	Reject	Accept
Family	−1.00	1.00
Friend	−0.75	0.75
Coworker	−0.50	0.50
Stranger	−0.25	0.25

Table 4: Payoffs corresponding to a preference for wearing a mask.

		Agent 2	
		Wear	¬ Wear
Agent 1	Wear	1.00	1.00
	¬ Wear	−1.00	−1.00

Table 5: Payoffs corresponding to a preference for not wearing a mask.

		Agent 2	
		Wear	¬ Wear
Agent 1	Wear	−1.00	−1.00
	¬ Wear	1.00	1.00

Table 6: Payoffs for the value of health. The numbers reflect how safe an agent feels.

		Infection risk	
		No risk	High risk
Action	Wear	0.00	1.00
	¬ Wear	0.00	−1.00

4.2 Contextual Properties

Whereas agents have limited observations on the environment, the context includes the place (home, office, party, park, and hospital) where interactions occur, the relationship (family, friend, colleague, and stranger) with the observer, the subjective belief of infection risk of the environment, the personal preference on mask-wearing, and the types of observers. Due to the partial observations, agents act based on their beliefs and update the beliefs with given rationales.

4.3 Types of Societies

We define types of societies based on the rationale types. All societies include 50% of agents value health and 50% of agents value freedom. The value importances of agents are presented in Table 7. All agents optimize their behavior based on the weighted sum of payoffs from themselves and others.

Baseline 1: Share-All Society Agents share all information as rationales and are capable of assuming the viewpoints of others to facilitate decision making.

Baseline 2: Share-Rules Society Agents share their decision rules as rationales.

***Exanna*: Share Value-Aligned Rules Society** Agents share their decision rules along with selected information that aligns with the values of agents present as rationales.

Table 7: Value importances of agents in a pandemic setting. Each society has half Freedom-loving and half Health-freak agents.

Agents: Values	Freedom	Health
Freedom-loving	1.00	0.00
Health-freak	0.00	1.00

4.4 Evaluation

We run simulations with the Share-All, Share-Rules, and *Exanna* societies. We evaluate hypotheses on resolution, social experience, privacy, and flexibility using the following metrics.

$M_{Resolution}$	$\in [0, 100]$	Percentage of rationales accepted.
M_{Social}	$\in [−3, 3]$	Aggregated payoff that an agent receives for its behavior.
$M_{Privacy}$	$\in [0, 1]$	Proportion of private information retained during an interaction.
$M_{Flexibility}$	$\in [0, 1]$	Extent of deviation from an agent's own goal.

We conduct the independent t-test across pairs of societies. We measure effect size with Glass's [12] Δ since the societies have different standard deviations [13]. We adopt Cohen's [6] descriptors to interpret effect size: <0.2 (negligible), $[0.2,0.5]$ (small), $[0.5,0.8]$ (medium), and ≥0.8 (large).

5 Results

Table 8 summarizes our statistical analyses. *Exanna* offers higher social experience, conflict resolution, and flexibility, indicating that *Exanna* agents learn to act for the greater societal good. These results follow our intuition that value-aligned rationales are superior. However, if an agent prefers to keep certain information private, deviation from goals is expected. Results for the Share-All and Share-Rules societies indicate that increased information in a rationale may not be helpful.

Table 8: Results: Comparing mean ($\bar{X}$) and standard deviation (σ) of social experience, resolution, privacy, and flexibility in various societies. p is p-value from t-test.

|		Share-All	Share-Rules	*Exanna*
$M_{Resolution}$	$\bar{X}$	0.58	0.58	0.60
	σ	0.02	0.01	0.02
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
M_{Social}	$\bar{X}$	0.59	0.62	0.70
	σ	0.06	0.02	0.05
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
$M_{Privacy}$	$\bar{X}$	0.00	1.34e−5	0.25
	σ	0.00	2.11e−5	2.90e−3
	p	< 0.001	< 0.001	–
	Δ	∞	11 896.52	–
$M_{Flexibility}$	$\bar{X}$	0.10	0.09	0.12
	σ	0.01	0.01	0.03
	p	0.08	< 0.01	–
	Δ	1.59	2.89	–

$H_{Resolution}$ Figure 2 compares conflict resolution in various societies. *Exanna* offers better conflict resolution ($p < 0.001$; $\Delta > 0.8$, indicating a large effect) than other societies. Thus, we reject the null hypothesis corresponding to $H_{Resolution}$. We observe that, in scenarios where other agents do not accept the provided rationales, *Exanna* agents more flexibly deviate from their own goals to resolve conflicts. Our results demonstrate the dynamics of agent behaviors and the strategy of rationales.
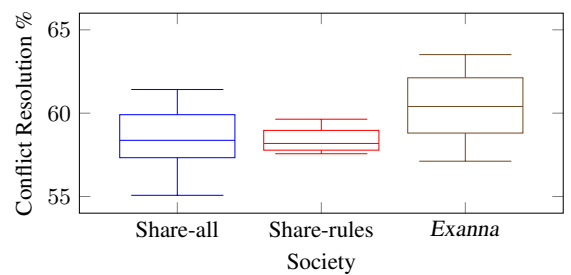
Figure 2: Comparing the resolution ($M_{Resolution}$) in various agent societies. The *Exanna* agent society has better resolution (Glass' $\Delta > 0.8$; $p < 0.001$) than the baseline societies.

$H_{Social\ Experience}$ For $H_{Social\ Experience}$, we measure the overall payoffs of agents in a society. An agent's payoff includes personal payoff from its action and feedback from its interaction. Figure 3 compares the social experience for Share-All, Share-Rules, and *Exanna* agent societies. We find that *Exanna* yields better social experience ($p < 0.001$; $\Delta > 0.8$, indicating a large effect)) than other societies. Specifically, *Exanna* agents receive better feedback from other

agents who receive their rationales. Thus, we reject the null hypothesis corresponding to $H_{Social\ Experience}$. On closer analysis, we observe that *Exanna* agents receive more negative sanctions than other societies initially but soon learn to deviate from their goals.
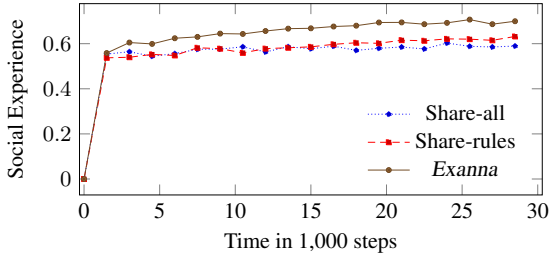

Figure 3: Comparing the social experience (M_{Social}) in various societies. *Exanna* agent society has better experience (Glass' $\Delta > 0.8$; $p < 0.001$) than other baselines.

H$_{Privacy}$ *Exanna* agents better retain their privacy ($p < 0.001$; $\Delta > 0.8$, indicating a large effect) compared to Share-All or Share-Rules agents. Thus, we reject the null hypothesis corresponding to H$_{Privacy}$.

Although both the Share-Rules and *Exanna* societies share learned rules as rationales, each *Exanna* agent aligns rationales to its values and those of the observers, and limits the private information shared to values that agents appraise. Our results show that a rationale stating causal attribution with minimum private information while aligning with individuals' values is sufficient to explain behaviors.

H$_{Flexibility}$ Figure 4 compares $M_{Flexibility}$ for Share-All, Share-Rules, and *Exanna* agent societies. We find that *Exanna* offers higher flexibility ($p < 0.01$; $\Delta > 0.8$) than the Share-Rules society. Although the mean flexibility in the Share-All society is lower than in the *Exanna* society, this difference is not significant ($p > 0.05$). Our results demonstrate that an agent can achieve a better social experience without sticking to only its goal.
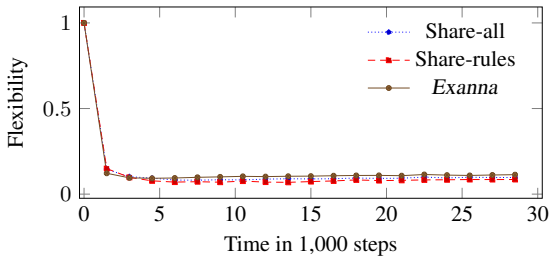
Figure 4: Comparing flexibility ($M_{Flexibility}$) in various agent societies. The *Exanna* society shows higher flexibility than baseline societies (Glass' $\Delta > 0.8$; $p < 0.05$ for Share-Rules but $p > 0.05$ for Share-All).

Emerged Norm   A norm emerges when the proportion of agents adhering to a particular behavior surpasses a threshold. We consider 90% as the threshold [9]. We observe that *Exanna* promotes more general norms than the Share-All and Share-Rules Societies. For instance, the following norms emerged only in *Exanna*.[1]

```
{preference = ¬Wear, InteractWith =
    Colleague, location=OFFICE} => Wear
```

```
{OberverAgentType = FREEDOM, InteractWith =
    Colleague, location=HOSPITAL} => Wear
```

---
[1] Additional emerged norms are listed in Appendix C.

6   Conclusions, Limitations, and Directions

Responsible autonomy requires that agents represent, reason with, and communicate values in their rationales. We demonstrate via a multiagent study how we could create agents who incorporate values in decision making and in rationale generation and evaluation. Value-aligned rationales offer better social experience and higher conflict resolution. Whereas value-aligned rationales withhold partial information, agents learn to deviate from their goals to protect their privacy. Specifically, agents who receive rejections from others become more flexible to improve cooperation.

Assumptions and Limitations   We make simplifying assumptions. First, agents can identify other agents' types, which indicate their values. We limit our simulation to two values (health and freedom) to demonstrate how value importances shape behaviors. A real-world scenario may include more intertwined values. Investigating the impact of these values on decision making and exploring approaches to elicit value importances necessitates further research. In addition, our numerical representation captures the importance of each value, but we do not target the complex contextual features in this work. We focus on providing insights and methodologies to understand and assess complex situations and generate informed rationales. We create and model simplified abstractions of intricate behaviors, enabling the analysis of complex situations.

By modeling the context as attribute-binding pairs, our approach is adaptable to various settings. However, our approach could suffer from state space explosion like other rule-based learning approaches. Despite considering single actions for simplicity, we focus on studying the effects of value-based explanations.

Our work contributes to developing value-aligned, trustworthy AI by showing how value-driven rationales help resolve conflicts arising from norm deviation and thus affect the emergence of norms. Our work complements approaches focused on eliciting values from stakeholders [17, 18] and figuring out what are adequate reasons for norm deviation [26] and for achieving trustworthy AI [25].

We assume the importance of contextual factors according to their associated values. For instance, someone prioritizing health may present factors correlating to health benefits or drawbacks. Although our scenario is simple, the rationales are dynamically constructed based on rule learning.

Future Directions   First, incorporating the cost of privacy loss is crucial. For instance, sharing sensitive information and sharing interests with chatbots impose different costs. Thus, modeling such costs is essential to capture how agents decide. Second, empowering agents to make informed choices regarding disclosure. For instance, when an agent acting on behalf of a stakeholder engages with agents other than the healthcare provider and the stakeholder, restricting the sharing of sensitive information may be desirable. Third, integrating rationales into the decision-making process, not just using them as supplementary information [29]. Having rationales as part of the decisions may increase the flexibility of an agent. Fourth, build an ontology to associate information with values, which we model as factors. An ontology helps to model varied factors or concepts and their intertwined relationships. While *Exanna* enables value-driven rationales and focuses on the decisions of a single agent, one future direction is to promote values and norms in a multiagent system [24].

Reproducibility   The codebase for our simulation is publicly available [27]. The appendices provide additional details, including hyperparameters for reproducibility, the complete set of emerged norms, and supplementary evaluations.

Acknowledgements

References

[1] R. Agrawal, N. Ajmeri, and M. P. Singh. Socially intelligent genetic agents for the emergence of explicit norms. *Proc. 31st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 10–14, Vienna, July 2022. IJCAI. doi: 10.24963/ijcai.2022/2.

[2] N. Ajmeri, H. Guo, P. K. Murukannaiah, and M. P. Singh. Robust norm emergence by revealing and reasoning about context: Socially intelligent agents for enhancing privacy. *Proc. 27th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 28–34, Stockholm, July 2018. IJCAI. doi: 10.24963/ijcai.2018/4.

[3] N. Ajmeri, H. Guo, P. K. Murukannaiah, and M. P. Singh. Elessar: Ethics in norm-aware agents. *Proc. 19th International Conference on Autonomous Agents and Multiagent Systems, (AAMAS)*, pages 16–24, Auckland, May 2020. IFAAMAS. doi: 10.5555/3398761.3398769.

[4] G. Aycı, A. Özgür, M. Şensoy, and P. Yolum. Can we explain privacy? *IEEE Internet Computing*, 27(4):75–80, 2023. doi: 10.1109/MIC.2023. 3270768.

[5] M. V. Butz and S. W. Wilson. An algorithmic description of XCS. *Proc. 3rd International Workshop on Learning Classifier Systems*, volume 1996 of *LNCS*, pages 253–272, Paris, France, 2000. Springer. doi: 10.1007/3-540-44640-0_15.

[6] J. Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 2nd edition, 1988. doi: 10.4324/9780203771587.

[7] V. Contreras, M. Schumacher, and D. Calvaresi. Integration of local and global features explanation with global rules extraction and generation tools. In *Explainable and Transparent AI and Multi-Agent Systems (EXTRAAMAS)*, volume 13283 of *LNCS*, pages 19–37, Virtual Conference, 2022. Springer. doi: 10.1007/978-3-031-15565-9_2.

[8] S. Cranefield, M. Winikoff, V. Dignum, and F. Dignum. No pizza for you: Value-based plan selection in BDI agents. *Proc. 26th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 178–184, Melbourne, 2017. IJCAI. doi: 10.24963/ijcai.2017/26.

[9] J. Delgado. Emergence of social conventions in complex networks. *Artificial Intelligence*, 141(1–2):171–185, 2002. doi: 10.1016/S0004-3702(02)00262-X.

[10] D. Di Scala and P. Yolum. Paccart: Reinforcing trust in multiuser privacy agreement systems. *Proc. 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 2787–2789, London, 2023. IFAAMAS. doi: 10.5555/3545946.3599078.

[11] A. Georgara, J. A. Rodriguez Aguilar, and C. Sierra. Building contrastive explanations for multi-agent team formation. *Proc. 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 516–524, Auckland, 2022. IFAAMAS. doi: 10.5555/3535850.3535909.

[12] G. V. Glass. Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10):3–8, 1976. doi: 10.3102/0013189X005010003.

[13] R. J. Grissom and J. J. Kim. *Effect Sizes for Research: Univariate and Multivariate Applications*. Routledge, Abingdon-on-Thames, 2012. doi: 10.4324/9780203803233.

[14] M. Hind, D. Wei, M. Campbell, N. C. Codella, A. Dhurandhar, A. Mojsilović, K. Natesan Ramamurthy, and K. R. Varshney. TED: Teaching AI to explain its decisions. *Proc. 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 123–129, Honolulu, 2019. ACM. doi: 10.1145/3306618.3314273.

[15] P. Langley. Explainable, normative, and justified agency. *Proc. 33rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 9775–9779, Honolulu, 2019. AAAI Press. doi: 10.1609/aaai.v33i01.33019775.

[16] R. Lera-Leri, F. Bistaffa, M. Serramia, M. Lopez-Sanchez, and J. Rodriguez-Aguilar. Towards pluralistic value alignment: Aggregating value systems through ℓp-regression. *Proc. 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 780–788, Auckland, 2022. IFAAMAS. doi: 10.5555/3535850.3535938.

[17] E. Liscio, M. van der Meer, L. C. Siebert, C. M. Jonker, N. Mouter, and P. K. Murukannaiah. Axies: Identifying and evaluating context-specific values. *Proc. 20th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 799–808, London, 2021. IFAAMAS. doi: 10.5555/3463952.3464048.

[18] E. Liscio, R. Lera-Leri, F. Bistaffa, R. I. Dobbe, C. M. Jonker, M. Lopez-Sanchez, J. A. Rodriguez-Aguilar, and P. K. Murukannaiah. Value inference in sociotechnical systems. *Proc. 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1774–1780, London, 2023. IFAAMAS. doi: 10.5555/3545946. 3598838.

[19] S. Luke, C. Cioffi-Revilla, L. Panait, K. Sullivan, and G. Balan. MASON: A multiagent simulation environment. *Simulation: Transactions of the Society for Modeling and Simulation International*, 81(7):517–527, July 2005.

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

[21] G. Ogunniye and N. Kökciyan. Contextual integrity for argumentation-based privacy reasoning. *Proc. 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 2253–2261, London, 2023. IFAAMAS. doi: 10.5555/3545946.3598903.

[22] N. Osman and M. d'Inverno. A computational framework of human values. *Proc. 23rd Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1531–1539, Auckland, May 2024. IFAAMAS. doi: 10.5555/3635637.3663013.

[23] S. H. Schwartz. An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, 2(1), 2012. doi: 10.9707/2307-0919.1116.

[24] M. Serramia, M. Rodriguez-Soto, M. Lopez-Sanchez, J. A. Rodriguez-Aguilar, F. Bistaffa, P. Boddington, M. Wooldridge, and C. Ansotegui. Encoding ethics to compute value-aligned norms. *Minds and Machines*, pages 1–30, 2023. doi: 10.1007/s11023-023-09649-7.

[25] A. M. Singh and M. P. Singh. Wasabi: A conceptual model for trustworthy artificial intelligence. *IEEE Computer*, 56(2):20–28, Feb. 2023. doi: 10.1109/MC.2022.3212022.

[26] A. M. Singh and M. P. Singh. Norm deviation in multiagent systems: A foundation for responsible autonomy. *Proc. 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 289–297, Macau, Aug. 2023. IJCAI. doi: 10.24963/ijcai.2023/33.

[27] S.-T. Tzeng. Code for "Value-Based Rationales Improve Social Experience: A Multiagent Simulation Study", 2024. URL https://doi.org/10. 5281/zenodo.12801893.

[28] S.-T. Tzeng, N. Ajmeri, and M. P. Singh. Fleur: Social values orientation for robust norm emergence. *Proc. 15th International Workshop on Coordination, Organizations, Institutions, Norms, and Ethics for Governance of Multi-Agent Systems (COINE)*, volume 13549 of *Lecture Notes in Computer Science*, pages 185–200, Auckland, May 2022. Springer. doi: 10.1007/978-3-031-20845-4_12.

[29] S.-T. Tzeng, N. Ajmeri, and M. P. Singh. Norm enforcement with a soft touch: Faster emergence, happier agents. *Proc. 23rd International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1837–1846, 2024. doi: 10.5555/3635637.3663046.

[30] R. J. Urbanowicz and W. N. Browne. *Introduction to Learning Classifier Systems*. Springer Briefs in Intelligent Systems. Springer, New York, 2017. doi: 10.1007/978-3-662-55007-6.

[31] G. H. Von Wright. *Norm and Action: A Logical Enquiry*. International Library of Philosophy and Scientific Method. Humanities Press, New York, 1963.

[32] E. Wang, P. Khosravi, and G. V. d. Broeck. Probabilistic sufficient explanations. *Proc. 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 3082–3088, Montreal, Aug. 2021. IJCAI. doi: 10.24963/ijcai.2021/424.

[33] M. Winikoff, G. Sidorenko, V. Dignum, and F. Dignum. Why bad coffee? Explaining BDI agent behaviour with valuings. *Artificial Intelligence*, 300:103554, 2021. doi: 10.1016/j.artint.2021.103554.

[34] J. Woodgate and N. Ajmeri. Macro ethics for governing equitable sociotechnical systems. *Proc. 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1824–1828, Online, May 2022. IFAAMAS. doi: 10.5555/3535850.3536118. Blue Sky Ideas Track.

[35] J. Woodgate and N. Ajmeri. Macro ethics principles for responsible AI systems: Taxonomy and directions. *ACM Computing Surveys*, 56(11): 1–37, July 2024. doi: 10.1145/3672394.

[36] V. Yazdanpanah, E. H. Gerding, S. Stein, M. Dastani, C. M. Jonker, and T. J. Norman. Reasoning about responsibility in autonomous systems: Challenges and opportunities. *AI & Society: Journal of Knowledge, Culture and Communication*, 38(4):1453–1464, Aug. 2023. doi: 10. 1007/s00146-022-01607-8.

A Appendix: Procedures of XCS

The overall process of XCS includes the following sub-processes.

- Matching: A process that matches the current context and all rules/classifiers to generate a match set. For instance, in our running example, the match set for Bella may include (1) {Risk = Low} ⇒ Wear [fitness = 0.3], (2) {Risk = Low} ⇒ ¬Wear [fitness = 0.7], (3) {OtherAgentType = Health} ⇒ Wear [fitness = 0.8], and (4){OtherAgentType = Health} ⇒ ¬Wear [fitness = 0.2]. The fitness is based on the accuracy of each rule's reward prediction.

- Covering: A process that guarantees diversity via adding a random classifier whose conditions match the current context. For instance, adding {Risk = Low, Relationship = Friend} ⇒ ¬Wear to the rule set.

- Action selection: XCS selects actions with pure exploration or pure exploitation with ϵ greedy. If not in exploration mode, this process returns the action with the highest fitness-weighted aggregation of reward.

$$\text{fitness}_a = \sum_{i}^{\text{rule}} \text{fitness}_i \times \text{numerosity}_i \times \text{predicted\_reward}_i \quad (2)$$

  where a $\in A$ and A is the action space. Rules represent all rules applied to the context and for action a. With the above example and formula, the agent would choose not to wear a mask due to $\text{fitness}_{\neg\text{Wear}} > \text{fitness}_{\text{wear}}$.

- Formation of action set: It includes all classifiers that propose the chosen action based on the match set. For instance, {Risk = Low} ⇒ ¬ Wear, {OtherAgentType = Health} ⇒ ¬Wear, and {Risk = Low, Relationship = Friend} ⇒ ¬Wear.

- Updating classifier parameters [30]: An agent updates the rule parameters (e.g., accuracy and fitness) based on the received payoff. The following equation updates the predicted reward, where p is the predicted reward, β is the learning rate, and r is the received reward.

$$p \leftarrow p + \beta(r - p) \quad (3)$$

  The prediction error ε is updated with the following equation.

$$\varepsilon \leftarrow \varepsilon + \beta(|r - p| - \varepsilon) \quad (4)$$

  The fitness of a rule is based on its accuracy, which is inversely proportional to the prediction error. We update the accuracy $kappa$ with the following formula.

$$\kappa = \begin{cases} 1 & \text{if } \varepsilon < \varepsilon_0 \\ \alpha\left(\frac{\varepsilon}{\varepsilon_0}\right)^{-\nu} & \text{otherwise,} \end{cases} \quad (5)$$

  where α is the scaling factor that raises a non-accurate rule to be close to an accurate rule. ε_0 is the threshold of prediction error below which the prediction error of a rule is assumed to be zero. ν defines how accuracy is related to prediction error and aims to help differentiate similar classifiers. For fitness calculation, we next calculate the relative accuracy κ' of each rule.

$$\kappa' = \frac{\kappa}{\sum_{cl \in [A]} \kappa_{cl}} \quad (6)$$

  where $[A]$ represents the corresponding action set. Finally, the fitness update of a rule is as follows.

$$F \leftarrow F + \beta(\kappa' - F) \quad (7)$$

  where F is the fitness of a rule.

- Subsumption: A process that replaces offspring rules with more general parent rules if it exists. Otherwise, save the offspring rules. Specifically, a more general rule yields a minor prediction error. For instance, if rule {Risk = Low} ⇒ ¬Wear has less prediction error than rule {Risk = Low, Relationship = Friend} ⇒ ¬Wear, the former rule would replace the later rule and increases the numerosity.

- Deletion: Each action set has the same maximum number of rules. XCS removes the low-fitness rules.

B Appendix: Hyperparameters for Reproducibility

Table 9 lists the hyperparameters we set for our simulations. We opt for the default configurations as in the literature [30] since our emphasis is on studying the impact of values in rationales and not in fine-tuning the learning process. *Learning rate* refers to reinforcement learning and determines the proportion an agent learns from recent experiences. Adjusting the *don't care probability* alters the probability of including the perceived factors in the generated rule. Increasing the *accuracy threshold* reduces the number of rules to maintain. *Fitness exponent* and *fitness falloff* determine the accuracy of the rules. *Genetic algorithm threshold* controls the probability of rule exploration. *Mutation probability* and *crossover probability* determine how often the mutation and crossover happen. *Experience thresholds* for *deletion* and *subsumption* guarantees a specific number of times a rule has to be applied before being deleted or subsumed.

The codebase for our simulation is publicly available [27].

Table 9: Hyperparameters for our settings.

Parameter	Value
Population size	200.00
Learning rate	0.10
Don't care probability	0.30
Accuracy threshold	0.01
Fitness exponent	5.00
Genetic algorithm threshold	25.00
Mutation probability	0.40
Crossover probability	0.80
Experience threshold for deletion	20.00
Experience threshold for subsumption	20.00
Fitness falloff	0.10

C Appendix: Detailed Results

Table 10 summarizes the statistical analysis of our simulations, including additional results for actor payoffs and observer payoffs and flexibility across different agent types. Actor payoff and observer payoff comprise social experiences.

These results demonstrate that first, with the same context, agents with different value importances can still evolve to different behaviors. Second, more adaptive norms can emerge among agents with different values.

H_Social Experience Social experience includes actor payoff and observer payoff. Figures 5 and 6 plot the payoffs of the actors who select actions, explain their behaviors, and receive feedback from observers in the Share-All, Share-Rules, and *Exanna* agent societies. Figures 7 and 8 compare the payoff from the observer who reacts to the actor's behavior in the Share-All, Share-Rules, and *Exanna* agent societies. The freedom-loving agents within *Exanna* society

Table 10: Results: Comparing mean ($\bar{X}$) and standard deviation (σ) social experience, resolution, privacy, and flexibility in various societies and agent types. p is p-value from t-test. M_{Social} has two subclasses, actor payoff and observer payoff.

		Share-All	Share-Rules	*Exanna*
$M_{Resolution}$	$\bar{X}$	0.58	0.58	0.60
	σ	0.02	0.01	0.02
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
M_{Social}	$\bar{X}$	0.59	0.62	0.70
	σ	0.06	0.02	0.05
	p	< 0.001	< 0.001	–
	Δ	1.80	3.11	–
Payoff Actor for health-freak	$\bar{X}$	0.63	0.63	0.63
	σ	0.03	0.03	0.03
	p	< 0.001	< 0.001	–
	Δ	10.40	9.62	–
Payoff Actor for freedom-loving	$\bar{X}$	0.96	0.96	0.94
	σ	0.03	0.03	0.03
	p	< 0.001	< 0.001	–
	Δ	0.60	0.64	–
Observer Payoff for health-freak	$\bar{X}$	−0.05	−0.07	0.02
	σ	0.09	0.09	0.09
	p	< 0.001	< 0.001	–
	Δ	2.71	2.39	–
Observer Payoff for freedom-loving	$\bar{X}$	−0.38	−0.36	−0.28
	σ	0.09	0.09	0.10
	p	< 0.001	< 0.001	–
	Δ	1.10	0.81	–
$M_{Privacy}$	$\bar{X}$	0.00	1.34e−5	0.25
	σ	0.00	2.11e−5	2.90e−3
	p	< 0.001	< 0.001	–
	Δ	∞	11 896.52	–
$M_{Flexibility}$	$\bar{X}$	0.10	0.09	0.12
	σ	0.01	0.01	0.03
	p	0.08	< 0.01	–
	Δ	1.59	2.89	–
Flexibility for health-freak	$\bar{X}$	0.14	0.12	0.13
	σ	0.06	0.07	0.06
	p	< 0.001	< 0.001	–
	Δ	0.85	0.44	–
Flexibility for freedom-loving	$\bar{X}$	0.06	0.06	0.09
	σ	0.03	0.03	0.03
	p	< 0.001	< 0.001	–
	Δ	0.90	0.87	–

encounter more adverse feedback than other societies initially. However, some of them quickly adapt and begin to divert from their original goals. As a result of the behavioral change made by freedom-loving agents, there has been an enhancement in the feedback re-
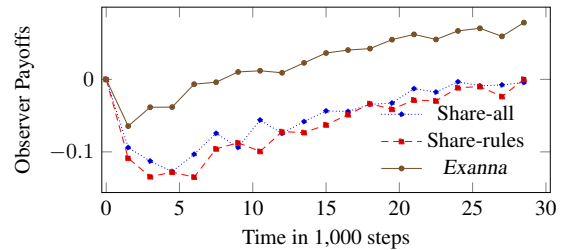
ceived by health-freak agents.
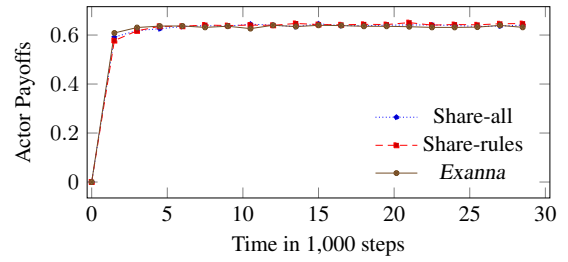
Figure 5: Comparing the actor payoff for health-freak agents in various agent societies. Actors are agents who act and receive feedback from others. Health-freak agents in each society have similar actor payoffs.
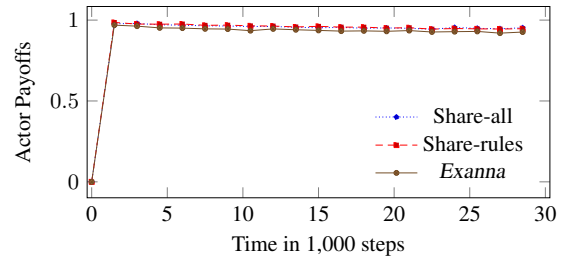
Figure 6: Comparing the actor payoff for freedom-loving agents in various agent societies. Actors are agents who act and receive feedback from others. The freedom-loving agents in *Exanna* society have lower actor payoffs (Glass' $\Delta > 0.5$; $p < 0.001$) than the baseline societies.


Figure 7: Comparing the observer payoff for health-freak agents in various societies. Observers give feedback based on observed behaviors and received rationales. The health-freak agents in *Exanna* society have better observer payoffs (Glass' $\Delta > 0.8$; $p < 0.001$) than the baseline societies.
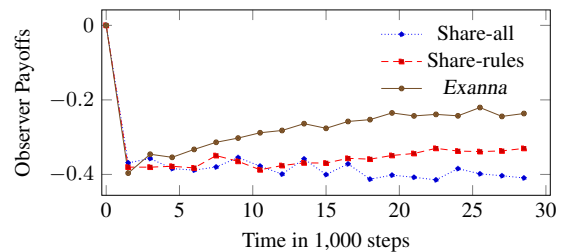

Figure 8: Comparing the observer payoff for freedom-loving agents in various societies. Observers give feedback based on observed behaviors and received rationales. The freedom-loving agents in *Exanna* society have better observer payoffs (Glass' $\Delta > 0.8$; $p < 0.001$) than the baseline societies.

H$_{\text{Flexibility}}$ We compare agents' flexibility of goals as the metric of evaluating H$_{\text{Flexibility}}$. Figures 9 and 10 compare M$_{\text{Flexibility}}$ for health-freak and freedom-loving agents in the Share-All, Share-Rules, and *Exanna* agent societies. Referring to Figure 8, the freedom-loving agents compromise on goals, thereby enhancing flexibility and enriching social experience.
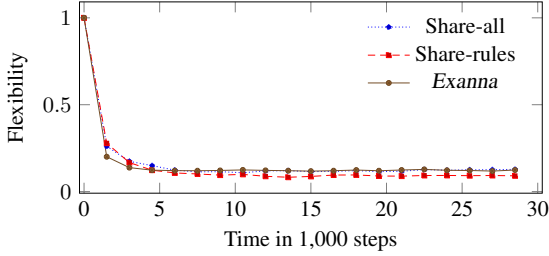
Figure 9: Comparing the flexibility for health-freak agents in various agent societies. The health-freak agents in *Exanna* society exhibit higher flexibility (Glass' $\Delta > 0.8$; $p < 0.001$) compared to the baseline societies.
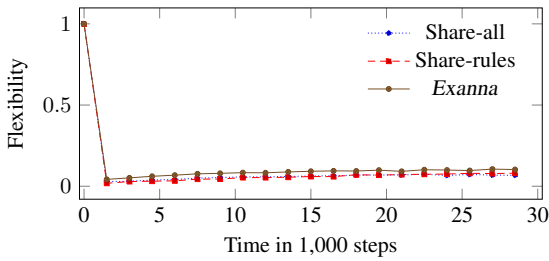
Figure 10: Comparing the flexibility for freedom-loving agents in various agent societies. The freedom-loving agents in *Exanna* society has higher flexibility (Glass' $\Delta > 0.8$; $p < 0.001$) than the baseline societies.

Complete Set of Emerged Norms

Table 11 lists the norms that emerge in the simulations. An emerged norm is a rule adopted by more than 90% of agents in one society.

D Appendix: Additional Experiments with Varying Value Importances

Agents may not always place extreme importance to one value over other. For instance, an agent can appreciate freedom (0.3) but cares more about health (0.7). To investigate further, we conduct additional experiments with a different set of value importances for freedom-loving and health-freak agents. Table 12 lists the value importances the freedom-loving and health-freak agents place in the additional scenario. We run each simulation five times, with other settings identical to the original simulations.

Table 13 shows the detailed results for agent societies with value importances as in Table 12.

H$_{\text{Resolution}}$ Figure 11 compares conflict resolution in various societies. *Exanna* offers better conflict resolution than other societies.

H$_{\text{Social Experience}}$ Figure 12 compares the social experience for Share-All, Share-Rules, and *Exanna* agent societies. A social experience includes personal payoff from its action and feedback from its interaction. *Exanna* yields better social experience ($p < 0.01$; $\Delta > 0.8$,
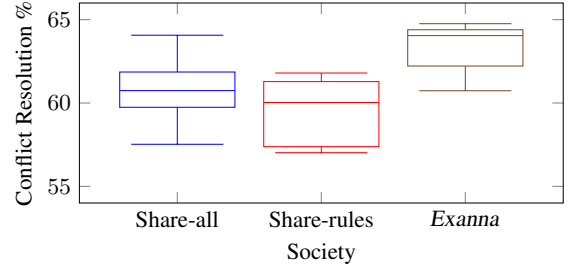
Figure 11: Comparing the resolution (M$_{\text{Resolution}}$) in various agent societies with value importances as in Table 12. The *Exanna* agent society has better resolution than the baseline societies.

indicating a large effect) than other societies. However, the mixed values lead to a worse social experience. With mixed values, an action now includes different concerns, which sometimes may contradict each other. For example, an agent may appreciate freedom and prefer not to wear a mask, but in the meantime, health concerns decrease its payoff on that decision.
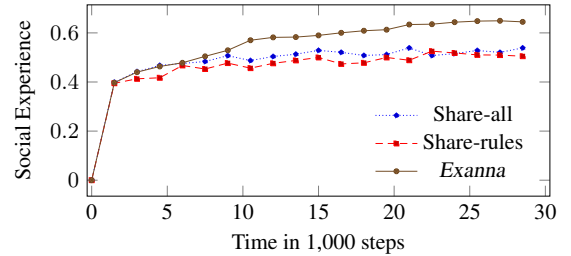
Figure 12: Comparing the social experience (M$_{\text{Social}}$) in various societies with value importances as in Table 12. *Exanna* agent society has better experience (Glass' $\Delta > 0.8$; $p < 0.001$) than other baselines.

H$_{\text{Privacy}}$ *Exanna* agents better retain their privacy ($p < 0.001$; $\Delta > 0.8$, indicating a large effect) compared to Share-All or Share-Rules agents.

H$_{\text{Flexibility}}$ Figure 13 compares M$_{\text{Flexibility}}$ for Share-All, Share-Rules, and *Exanna* agent societies. The mixed values lead to higher flexibility in agent societies than in the main simulations. Specifically, the mixed values narrow the numerical gap between each action, decreasing the threshold for agents to change their minds.
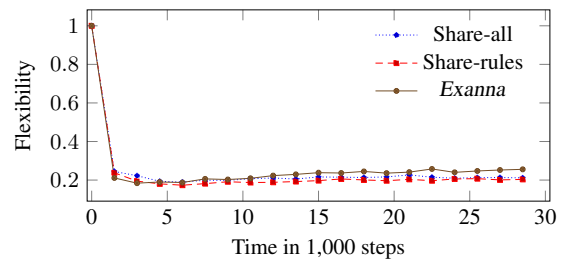
Figure 13: Comparing flexibility (M$_{\text{Flexibility}}$) in various agent societies with value importances as in Table 12. The *Exanna* society shows higher flexibility than baseline societies (Glass' $\Delta > 0.8$; $p < 0.05$ for Share-Rules society but $p > 0.05$ for Share-All society).

Table 11: Emerged norms in agent societies. Common means the norms emerge in each agent society.

Society	Norm				Consequence
	Premise				
Common	Risk	=	NONE;		WEAR
	preference	=	¬WEAR;		
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
	Risk	=	NONE;		WEAR
	preference	=	¬WEAR;		
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	HOSPITAL		
	Risk	=	RISK;		WEAR
	preference	=	¬WEAR;		
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
	Risk	=	RISK;		WEAR
	preference	=	¬WEAR;		
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	HOSPITAL		
Share-All	Risk	=	NONE;		WEAR
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
Share-Rules	preference	=	¬WEAR;		WEAR
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
Exanna	preference	=	¬WEAR;		WEAR
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
	preference	=	¬WEAR;		WEAR
	InteractWith	=	COLLEAGUE;		
	location	=	HOSPITAL		
	preference	=	¬WEAR;		WEAR
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
	preference	=	¬WEAR;		WEAR
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	HOSPITAL		
	OberverAgentType	=	HEALTH;		WEAR
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
	OberverAgentType	=	HEALTH;		WEAR
	InteractWith	=	COLLEAGUE;		
	location	=	HOSPITAL		
	OberverAgentType	=	FREEDOM;		WEAR
	InteractWith	=	COLLEAGUE;		
	location	=	HOSPITAL		
	Risk	=	RISK;		WEAR
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		
	Risk	=	NONE;		WEAR
	OberverAgentType	=	HEALTH;		
	InteractWith	=	COLLEAGUE;		
	location	=	OFFICE		

Table 12: Value importances of agents in the additional scenario. Each society has half Freedom-loving and half Health-freak agents.

Agents: Values	Freedom	Health
Freedom-loving	0.7	0.3
Health-freak	0.3	0.7

Table 13: Results: Comparing mean ($\bar{X}$) and standard deviation (σ) of social experience, resolution, privacy, and flexibility in various societies in the additional scenario. p is p-value from t-test.

		Share-All	Share-Rules	*Exanna*
$M_{Resolution}$	$\bar{X}$	60.79	59.50	63.23
	σ	2.43	2.20	1.70
	p	0.1	< 0.05	–
	Δ	1.00	1.69	–
M_{Social}	$\bar{X}$	0.53	0.52	0.65
	σ	0.06	0.05	0.03
	p	< 0.01	< 0.01	–
	Δ	2.24	2.57	–
$M_{Privacy}$	$\bar{X}$	0.00	5.78e−5	0.25
	σ	0.00	1.04e−4	2.60e−3
	p	< 0.001	< 0.001	–
	Δ	∞	2414.62	–
$M_{Flexibility}$	$\bar{X}$	0.22	0.20	0.25
	σ	0.03	0.02	0.02
	p	< 0.05	< 0.01	–
	Δ	1.29	2.02	–