Incorporating Social Expectations into the Expectation Event Calculus

Nathan Lloyd[1] and Peter R. Lewis[1]

[1]Ontario Tech University, Canada
nathan.lloyd@ontariotechu.net

Abstract

This paper focuses on the internal representation of norms and the necessary concepts required for normative decision-making in computational agents. We leverage the social science literature to integrate currently absent prevalent concepts of social expectations within the Expectation Event Calculus (EEC) and, in doing so, extend the formalism. Through the adopted terminology of expectations, we distinguish between descriptive and social norms, enabling a more comprehensive description of individual and collective behavior conditional on social expectations. We introduce complementary abstractions for normative attributes to demonstrate and explain why such a distinction between expectations enables richer normative scenarios to be modeled, which has yet to be shown in the EEC. We demonstrate this extension through a binary-decision social scenario. First, through a single-agent implementation driven by hardwired narratives. Secondly, we demonstrate the extension through a selection of multi-agent scenarios that showcase a change in behavior conditional on expectations. As a minimal implementation of social expectations, we conclude the paper with themes and open challenges as avenues for further research.

Introduction

Norms are the "informal rules" (Bicchieri et al., 2022b) that guide and govern individual and collective behavior within society. They play a pivotal role in maintaining social order by providing a framework for acceptable and appropriate conduct, enabling individuals to understand what is *expected* of them within different situations. The social function that norms provide is central to discussions for motivating pro-social outcomes in *social dilemmas* Sengupta et al. (2021) and for achieving cooperation Fehr and Fischbacher (2004), among others. Although it is a common practice to discuss norms in terms of their functional role, especially so when they motivate desirable outcomes, this approach may overlook the essential prerequisites for norm emergence and dynamics. Namely, the importance of beliefs, expectations, shared knowledge, group structures, and relationships that are integral to their very existence. Rather than driving individual behavior to accomplish said outcomes, instead, we consider mechanism over function and discuss norms as operating on two distinct levels (Lapinski and Rimal, 2005, p.129). At an external level, they are propagated amongst group members through behavior and communication (Kincaid, 2004), i.e., sanctions. As well as at an individual level (internal), where they are perceived and interpreted. Norms, therefore, exist within the mind as cognitive artifacts and are externalized through behavior (Campenní et al., 2009), operating in some internal-external loop (Castelfranchi, 2016). As individuals manifest their perceived norms in their actions, their internal representations may change in response to the behavior of others. This loop underscores the interdependence of behavior and cognition, prompting an exploration into the representation of norms that artificial agents can contemplate. This paper extends the Expectation Event Calculus (Cranefield, 2014), a logical framework for reasoning about various expectations and norms. We do so to empower agents with the capacity to deliberate over a previously uncaptured distinction in the EEC between empirical and normative expectations, doing so in line with prevalent social theories of norms.

A Constructivist Theory of Norms

This section introduces where in the literature a distinction between descriptive norms and social norms arises and what factors cause this separation. Under Bicchieri's constructivist theory (Bicchieri, 2005), norms and, subsequently, normative behavior are thought to be contingent upon one's social expectations and conditional preferences. The first essential concept introduced in Bicchieri's conceptualization of social norms is social expectations. An umbrella term, social expectations describe one's *empirical expectations*, an expectation about others' behavior ("I expect they will do..."), and one's *normative expectations*, a belief about what others expect them to do ("I believe they think others ought to do..."). They are social as they describe what is believed about the prevalence and distribution of others' behaviors and beliefs (Wang, 2022), subject to change through experience.

Expectations form the basis of normative decisions under this theory but do not exclusively shape behavior. One's preferences are core to an individual's decision-making process, which we can describe as conditional or unconditional.

Conditional preferences describe an individual's inclination to act in a particular manner contingent upon what others think or do (their expectations). In contrast, unconditional preferences describe a preference to behave in a specific manner regardless of the actions or beliefs of others, signaling an "uninfluenced"[1]. We underscore the distinction to highlight the importance of expectations and their necessary quality of being interdependent. Social expectations allow us to describe two distinct categories of norms, which we will briefly introduce.

Descriptive Norms

A *Descriptive Norm* refers to a behavioral pattern an individual chooses to adopt, contingent upon others exhibiting that behavior (Bicchieri, 2005). A descriptive norm describes interdependent social behaviors contingent solely on empirical expectations, informally as "what people normally do", as in their typical behavior. Following this definition, descriptive norms can describe behaviors like imitation and coordination, as they are solely rooted in the observed or communicated behavior of others and an individual's preference to conform without consideration for normative expectations.

Social Norms

A *Social Norm* is defined as a rule of behavior that individuals prefer to conform to on the condition that they believe a sufficiently large number of their reference network[2] engages in that behavior and also holds a belief that others think that behavior ought to be engaged in (empirical and normative expectations respectively) (Bicchieri, 2016, p.35)[3]. Bicchieri's definition affirms the importance of preferences, social relationships, and expectations for social norms and, importantly, recognizes that preference for norm compliance is highly context-dependent. For example, if one is aware that people line up in a queue and people have a normative expectation that others ought to queue, we may describe a social norm.

Why distinguish Social Expectations?

Beyond differentiating between descriptive and social norms, delineating empirical and normative expectations aids in describing the conditions, processes, and attributes for normative behavior to occur, persist, and falter. For normative expectations, research has shown that mutually consistent normative expectations signal the probable existence of a social norm (Bicchieri and Chavez, 2010), fear of punishment from sanctions (that may accompany normative expectations) may motivate cooperative behavior (Fehr and Fischbacher, 2004), without sanctioning the influence of the norm is likely decreased (Bicchieri and Chavez, 2010). Moreover, empirical studies have provided insights into their interplay; when social expectations are inconsistent and in conflict, empirical expectations typically dominate (Bicchieri and Xiao, 2009; Kölle and Quercia, 2021).

Beyond the direct interplay between expectations, the distinction also warrants a deeper understanding of the conditions and mechanisms that form them and later serve in one's decision-making processes. Individuals may possess various attributes, like thresholds, that determine the point at which they consider the number of followers too small to be significant in motivating their conformity. Additionally, individuals may exhibit different thresholds for different norms (Bicchieri and Chavez, 2010), offering a means to introduce human-like heterogeneity into agent-based models. One's anticipation of sanction, norm sensitivity, risk perception, personal propensity to risk, reference network, perception of the current state, mood, and so on (Bicchieri, 2016; Wijermans et al., 2024, c.5), are all posited as factors that motivate decisions to conform, evade, or flout a norm. Similarly, a lack of knowledge or awareness of the norm or conditions it applies may explain some instances of non-compliance, such as when an individual from a population where queuing is not a social norm travels to one where it is, and vice versa; again underscoring the highly context-dependent nature of preferring to conform to prevailing norms (Bicchieri and Zhang, 2012).

An Example Scenario

To show the distinction between the underlying expectations, we will introduce an example scenario of archetypal individuals who show varying levels of sensitivity to normative and empirical expectations, one we will continue throughout the paper. We introduce this sensitivity as abstract attributes to demonstrate a tendency or preference to conform to one's empirical and normative expectations. In this example, two agents, Chris and Paul, have recently begun employment at a fictitious company and have both come to form expectations about their colleagues:

1. Every Friday, they observe their colleagues go to the bar after work; they form empirical expectations about their colleagues' behavior for Fridays to come.

2. No normative expectations have been formed by either Chris or Paul; no sanctions, communication, or other forms of behavior would indicate that others think they *ought* to (or not to) attend.

When facing new, uncertain, or ambiguous situations, individuals typically look to others to gather information; they often imitate or seek guidance rather than deliberate (Bicchieri, 2016, p.23). Paul is keen to fit in; thus, he is sensitive to his new colleagues' behavior and the empirical expectations he has formed. However, Chris *prefers* to go home

[1] Uninfluenced directly by expectations, but could be influenced by other means

[2] Individuals who matter to your decision-making process (Bicchieri, 2005)

[3] A more formal definition can be found here (Bicchieri, 2005, p.11)

and has avoided going to the bar on Fridays. In this example, we illicit how preferences play an essential role in conformity and how an agent's behavior is conditional on relevant expectations. Continuing the narrative, we may see that Chris's colleagues have noticed his reluctance to join in social activities and, in response, have begun to communicate that employees *should* attend the weekly event. Chris, Paul, and their colleagues may now have formed normative expectations, meeting the conditions for a social norm to exist. In this scenario, Chris may have shown no sensitivity to his empirical expectations but may be sensitive to normative expectations, particularly if sanctions (i.e., peer pressure) reinforce them. Whether Chris conforms and participates in social activities may hinge on the aforementioned attributes and mechanisms alongside his goals. Specifically, he perceives that adhering to social norms could propel his advancement within the company.

Social Expectations and Conditional Preferences

Thus far, we have described the role of social expectations in motivating descriptive and social norms and introduced some essential terminology for describing conditional normative behavior. Thus far, both forms of expectations have yet to be captured within the Expectation Event Calculus (Cranefield, 2014), presenting an opportunity to explore their dynamics in an established framework. We underscore the importance of this research in its ability to support the investigation and modeling of norm emergence and norm dynamics and to explore the conditions for richer normative behavior (Lloyd and Lewis, 2024): norm avoidance, alteration, creation, and destruction of norms as individuals and collectives.

Agents and Expectations

An expectation can be defined as a belief oriented toward the future, in which an individual has a vested interest and actively monitors (Castelfranchi, 2005); they are powerful beliefs that can originate from our experiences, observations, commitments, and institutional rules amongst other social contexts. As a belief about a future state, expectations "imply a subjective concern in the realization of a [future state that is expected to hold]" (Castelfranchi, 2005), a *concern* emphasizing the goal of determining whether the expectations are fulfilled through active monitoring. We draw comparisons between the language of "concern" and "interest" discussed by Coleman (Coleman, 1994), and later Fehr and Fischbacher (Fehr and Fischbacher, 2004), where an interest in the actions and beliefs of others creates the demand for a social norm.

Expectation Event Calculus

Following an extensive review of computational expectations, Cranefield proposes the *Expectation Event Calculus* (EEC) (Cranefield, 2014), a logical tool for reasoning about

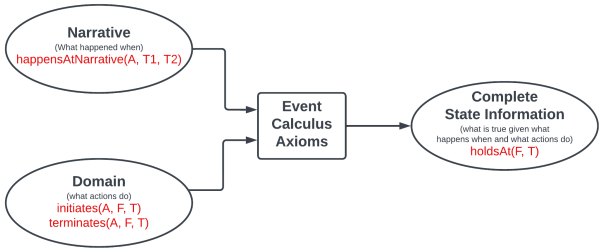

Figure 1: Overview of reasoning the Event calculus, adapted from (Shanahan, 2001; Sengupta et al., 2021).

expectations, commitments, and norms. As an extension of the *Event Calculus* (EC) (Kowalski and Sergot, 1986; Shanahan, 2001), EEC inherits EC's core principles, able to represent events with indirect or non-deterministic effects, continuous, concurrent, or compound actions, and reason about how these actions (often referred to as events) instigate some change in the world, given some initial state.

Event Calculus The Event Calculus depicts the state of the world at any given time through predicate facts; when these facts represent dynamic properties of the world, they are referred to as fluents, which are reified predicates. Reification enables fluents to be represented as terms to be passed into the logical axioms of EC, i.e., $holdsAt(on(cup, table), 1)$ reads as the cup is on the table at time 1. Events are also reified within the Event Calculus, such that $happensAt(grabs(wyn, cup), 2)$ may infer that $holdsAt(on(cup, table), 2)$ is false. Through the Event Calculus Axioms, one can query which fluents hold at a particular time and query the events that trigger the initiation and termination of fluents; i.e., picking up a cup from the table terminates the fluent of it being on the table at that time until some other event reestablishes that fluent at another.

More generally, the Event Calculus is posited as a method for *Common Sense Reasoning* (Mueller, 2014). Figure 1 illustrates the three core reasoning capabilities at the heart of the Event Calculus. The first, *temporal projection* or *prediction* (Mueller, 2014, p.7), starts with the narrative of events ("what happens when") and the domain ("what actions do") to reason about the result of an event ("what is true given what happens when and what actions do") (Shanahan, 2001). The Event Calculus also supports *abductive reasoning*, starting with the narrative and the result to reason about the events that lead from the initial to the final state. Finally, *postdiction* is reasoning about prior states with knowledge of the final state and what actions do. These capabilities underscore the robust and versatile nature of the Event Calculus, emphasizing its ability to handle diverse aspects of reasoning over events, actions, and outcomes.

Expectations In Bicchieri's characterization of social norms, expectations and conditionality are necessarily linked such that a norm can be described as a mapping from a particular context to a behavioral rule. Conditional expectation rules through the EEC (Cranefield, 2014) have been expressed as fluents: $exp\_rule(Cond, Exp)$, where the condition $Cond$ is some representation of past and present, and constraint Exp denotes some expectation regarding the future. If the condition of an expectation rule is true, the expectation becomes active. The additional semantics introduced into the Expectation Event Calculus (Cranefield, 2014) and its following implementations and extensions (Sengupta et al., 2021, 2023) enable the type of reasoning desirable for decision-making with both forms of expectations. Still, thus far, the EEC has focused heavily on institutional rules rather than individual cognition without this clear separation between empirical and normative expectations. Our contribution is to incorporate this distinction into the Expectation Event Calculus and, in doing so, extend the formalism. A significant focus in cognitive modeling revolves around understanding how individuals formulate situation-action rules of behavior (Clancey et al., 2005). As such, we adopt an individual cognitive approach to drive social behavior through social expectations within the Expectation Event Calculus. Cranefield's (Cranefield, 2014) proposal for the EEC has since been applied and adapted in subsequent work (Sengupta et al., 2021, 2023), which will be briefly explored in the following subsection.

EEC Applications Following its inception, the Expectation Event Calculus has been demonstrated as an effective tool for representing and reasoning about expectations and norms, thereby motivating cooperative behavior in social dilemmas (Sengupta et al., 2021). Similarly, in implementing Axelrod's *Metanorms Game* (Axelrod, 1986), Sengupta et al. (Sengupta et al., 2023) illustrates how expectations, explicit norms, and the incorporation of what-if reasoning (to compare alternative actions) enable the preservation of first-order norms, as defined in the Metanorms game, within society. Also introduced in subsequent work (Sengupta et al., 2023) are predicate naming conventions that help to remove any ambiguity in describing the external narrative events through a $happensAtNarritve$ predicate and internal inferred fulfillment or violation of events through a $happ$ predicate; providing further utility to the EEC.

Extension through Social Expectations

The research highlights two key areas that require attention. Firstly, current implementations focus primarily on expectations at the institutional level, overlooking the significance of understanding individual expectations among agents and how individual expectations lead to institutional ones. Secondly, there has yet to be any differentiation between types of expectations (empirical and normative) within EEC thus far, resulting in a challenge to differentiate between descriptive and social norms within agent-based models. Addressing these aspects will contribute to a more comprehensive and nuanced understanding of expectations in various contexts and provide a basis for distinguishing supporting attributes in normative decision-making.

Implementing Social Expectations

This section introduces and demonstrates a minimal implementation and functionality within a social-scenario case study. To extend the EEC to include empirical and normative expectations, we introduce two additional, conditional expectation rules in the form of $emp\_exp\_rule(Cond, Exp)$ and $nor\_exp\_rule(Cond, Exp)$. While the expectations share a similar structure to one another and the former $exp\_rule(Cond, Exp)$, they are distinguished both by how they are informed and by the type of future-directed expectation (Exp) they contain, Emp_exp and Nor_exp respectively. To illustrate the distinction between empirical and normative expectations, we continue the social scenario example outlined previously and implement this extension through Prolog, following on from existing implementations of the EEC.

Agent Definition

For a binary decision scenario, we introduce a pair of fluent preferences in the form $pref(bar, Val)$ and $pref(home, Val)$, describing an agent's preference over an action or location by the values they contain. We introduce these preferences as fluents as it allows them to be called under relevant conditions, such that a preference to go to the bar or to go home would be activated when the relevant condition motivates their appearance, a step toward the dispositional account (Bicchieri, 2005; Lloyd and Lewis, 2024). The fluent preferences contain the location and a numerical value, so a preference may be read as 'I prefer to do X more than Y.'

We also introduce two 'normative' attributes, namely a $emp\_exp\_sens(Val)$ and a $nor\_exp\_sens(Val)$. These sensitivities are introduced as a general abstraction for sensitivity to empirical and normative expectation rules, such that one may be described as highly sensitive to the behavior of others but not what they believe others think should or ought to happen, and vice-versa. In the scenarios, preferences and sensitivities may hold values [0, 1, 2, 3]; arbitrary ordinal values highlight how a preference may dominate sensitivity and vice versa. As such, one may 'prefer to do X more than Y but is more sensitive to empirical expectations, one exists, and as such does Y,' i.e., $pref(bar, 0) < pref(home, 1) < emp\_exp\_sens(2)$. We utilize these abstract attributes to highlight the importance of representing both expectations and recognize that the literature posits similar attributes such as general sensitivity to norms, sensitivity to individual norms, and so on (Bicchieri,

2016), emphasizing a possible pathway to improve normative decision-making that aligns more closely with existing literature. Thus, an agent is defined as:

$$a_i =< prefBar, prefHome, empSens, norSens >$$

Conditionality on Social Groups

Next, we define the relationship between employees through an $employee(a_i, company)$ fluent, which acts as a condition for the expectation rules, read informally as 'if you are an employee; these expectations pertain to you, else they do not.' We lead this narrative with some assumptions about group membership; in the scenario we present, only one group exists, and membership is crisp, allowing the conditionality of expectations to be easily grounded upon membership to the fictional company, meeting the conditionality clause.

Expectation Formation

The behavior of others informs one's empirical expectations; as such, we define the initiation of an $emp\_exp\_rule(Cond, EmpExp)$ through a $process(T, State, Condition)$ predicate. As observations of others are recorded in the $happensAtNarrative()$ predicate, the process predicate takes the current time, physical state, and prevalent conditions, i.e., memberships or beliefs, to determine whether an expectation rule is initialized. As we have assumed the state of an agent is always at work when they decide, we use Prolog's anonymous variable in situ. Given the time, the agent can process all observed behaviors at that time step, ensure that the frequency of behaviors is greater than zero, and then check whether an expectation regarding that behavior already holds. If not, the agent forms an empirical expectation based on active conditions. Given the setup for the hardwired actors, a single empirical expectation rule will be generated when their behaviors are either all: go to the bar or go home, whereas, in a mixed setting, two empirical expectation rules will be generated, and in a setting where the behavior of other's is not observed, none are generated. Given the additional capabilities required to infer others' normative beliefs, we introduce the signal for a normative expectation through a broadcasted speech act, which agents are programmed to interpret. A speech act is defined through a $sa(a_i, NorExp)$ predicate, which contains the 'id' or name of the agent who gave the speech act and the normative expectation. Although a trivial implementation for signaling and interpreting a normative expectation, our use of a speech act symbolizes a request for behavior.

Single Agent Validation

Initial testing on a single agent was conducted to validate the extension presented in this work. This agent, which we have dubbed Joan, joins a fictitious company and observes their colleagues' behavior to form expectations. First, we introduce eight colleagues, who are effectively *hardwired* to act in a predetermined manner. Thus, we have a set of agents:

$$A = \{\text{joan}, \text{colleague}_1, \text{colleague}_2, \ldots, \text{colleague}_8\}$$

The behavior of the 'actor agents' is the first and second parameter of the scenarios as they effectively inform *Joan*'s empirical and normative expectations. We introduce four behavior variations where all hardwired agents go to the bar, go home, no behaviors are observed (no $happensAtNarrative(go(a_i, Loc), T)$), or an even split of behaviors exist. In addition to their behavior over the binary decision, we introduce signals for the normative expectation through a speech act predicate, which has an identical setup to behavior parameters; but is only broadcast at $T = 2$. Scripted narratives in combination with low-dimensional action spaces is a standard method to reduce complexity (Kovač et al., 2021), in this case, removing the need for agents to be situated and perceptive, effectively mimicking a social scenario with a single agent and enabling early validation before implementing a full multi-agent scenario. Whereas the 'actor agents' behavior is predetermined, Joan may act according to their preferences and expectations. Joan is aware that all their colleagues are the company's employees initially; however, Joan joins the company later; see Figure 2 for a scenario timeline. We utilize discrete time to evaluate how Joan's behavior changes through the formation and presence of expectation rules, given their sensitivity to them and preferred behavior/location. Joan's decisions and behaviors are recorded over 'five weeks', as described in Figure 2. In the first week, Joan joins the company, and then at $T = 2, 3$, Joan participates in the social scenario decision and observes the behavior of others. At $T = 4$, Joan leaves the company, effectively terminating the conditions for the expectations to hold, and finally makes the social scenario decision once more in $T = 5$. Each time step loosely represents a week for the sake of the narrative, where $happensAtNarrative(joins(joan, company), 1)$ would be read as Joan joins the company in the first week. Following this, each time step denotes the social scenario behavior, i.e., $happensAtNarrative(go(colleague_1, bar), 2)$ or $happensAtNarrative(go(colleague_1, home), 2)$.

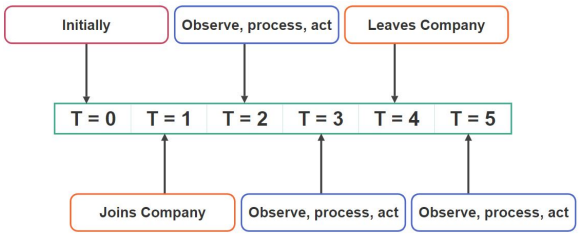

Figure 2: Timeline of the Narrative

Single Agent Scenarios

In each run instance, one value is selected from each scenario and agent property, as described in Table 1. An assumption is introduced that an agent's attributes cannot equal the value of another, such that sensitivity to an expectation could be read as a preference to conform to empirical or normative expectations; this amounts to a total of 384 combinations.

Table 1: Scenario properties

Scenario Properties	Values
Actor's Behavior	[AllHome, AllBar, 50/50, NoObs]
Actor Speech	[AllHome, AllBar, 50/50, NoObs]
empSens	[0,1,2,3]
norSens	[0,1,2,3]
prefHome	[0,1,2,3]
prefBar	[0,1,2,3]

Single Agent Results

Following the narrative of events outlined in Figure 2 and the 384 scenario combinations possible in Table 1, a subset of particularly noteworthy instances have been selected for discussion within the results section, the parameters of which are described in Table 2 and behaviors in these scenarios visualized in Figure 3.

Table 2: Demonstrated Single-Agent Scenarios

Scenario #	Description	Values
1	Prefs. in Solitude	[NoObs, NoObs, 0, 1, 3, 2]
2	Sens. to EmpExp	[AllBar, NoObs, 3, 0, 2, 1]
3	Sens. to NorExp	[NoObs, AllBar, 0, 3, 1, 2]
4	Conflicting EmpExp	[50/50, NoObs, 3, 2, 0, 1]

Scenario 1 captures Joan's behavior outside of a social context, where no observations of behavior or speech acts inform an expectation. Given a maximum value for $prefHome$, Joan follows their preferences unconditionally. Equally, if these attributes were to change such that Joan has a conditional preference to follow either expectation, the resulting behavior would default to one's preference. Scenario 2 captures Joan's behavior, in which the behavior of others is opposed to their preference for home ($empSens > prefHome > prefBar$). At the same time, they are more sensitive to the behavior of others and, as such, act against their preferences while the condition of that expectation holds; as shown in Figure 3 by a change in behavior once leaving the company. This is an example of interdependent behavior conditional on one's empirical, multilateral expectations. Scenario 3 captures Joan's behavior where the communicated normative expectation, through a speech act, is opposed to their preference of bar, and at the same time, are more sensitive to what they believe *others think should happen* ($norSens > prefBar > prefHome$). Similarly, this signals interdependent behavior that is, instead, conditional on one's normative, multilateral expectations. Scenario 4 captures Joan's behavior when there are conflicting (empirical) expectations. Joan prefers going to the bar more than home, and at the same time, they are more sensitive to the behavior of others ($empSens > prefBar > prefHome$) but have observed their colleagues exhibiting both behaviors. Given conflicting expectations, Joan again defaults to their own preference, as shown in Figure 3.

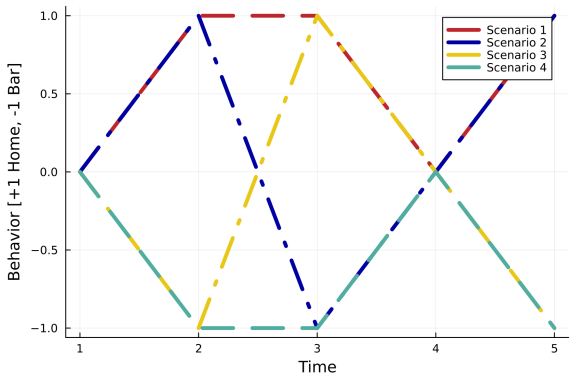

Figure 3: Single-Agent Behavior in described Scenarios

Discussion

Generally, we can consider Joan a conditional follower if their sensitivity to an expectation is higher than their preferences, while the expectation and its condition hold at the time of the decision, i.e., $holdsAt(employee(joan, company), T)$. If multiple conflicting expectations exist, Joan defaults to their maximum preference; however, in a more complex agent, we anticipate capabilities for social information seeking, particularly if the wrong decision is anticipated to be costly, necessitating further attributes and reasoning processes to calculate risk. Additionally, sensitivities to individual expectations and norms, perhaps in the form of $emp\_exp\_rule(Cond, Exp, Sens)$, as well as deontic language and grading as proposed in Malle et al. (2019), could enhance normative reasoning capabilities.

Multi-Agent Simulation

We continue testing with multi-agent simulations, utilizing the same social scenario and expanding the set of decision-making agents from 1 to 30, removing hardwired actors and randomly assigning the aforementioned attributes under the same conditions and assumptions. We chose to exhibit these findings through synchronized scheduling, incorporating a buffer in behavior to ensure that observations and processing take place for each agent before any action is

taken by themselves or another. We implement the MAS through Julia's Agents package (Datseris et al., 2022) and utilize the abstract model and agent type. We initialize agents with their own SWIPL connection (through Julia's $open(command, "r+")$) and initialize their knowledge base to contain initial information about the state of the world, i.e., employee membership. Again, we demonstrate these agents through scenarios of length five and run each scenario ten times to account for randomized agent attributes.

In the MAS, we introduce two new fields: the first denotes an agent's ID, and the second represents a perceptive probability, which may also act as a proxy for social proximity (Bicchieri et al., 2022a)—this 'vision' attribute is simply a probability threshold that must be passed to perceive, and memorize, the behavior of others. By introducing this attribute, we may compare agents with complete information with those without, demonstrating that if individuals see others flout the underlying expectations, their willingness to comply may decrease. Conversely, this enables us to demonstrate how "unobserved" transgressions may maintain conditional followers. In the Single-Agent Simulation, complete knowledge was assumed. As such, we aim to relax this here with 2%, 5%, 10%, 25%, 50%, and 100% homogeneous vision thresholds, each run ten times with 30 agents.

MAS Scenario Specifications

We construct six scenarios based on the above-mentioned methodology to illustrate collective behavioral change over time. These scenarios are described below:

1. Initially, an empirical and normative expectation of going to the bar.
2. Initially, there is an empirical expectation of going to the bar but a normative expectation to go home.
3. Initially, there is a normative expectation of going to the bar but an empirical expectation to go home.
4. Initially, an empirical expectation of going to the bar.
5. Initially, a normative expectation of going to the bar.
6. No initial expectations.

MAS Scenarios Results

Each run utilizes a different random seed for attribute initialization; however, to promote repeatability and comparability, the random seed is applied between scenario and vision changes, such that the first run of scenario 1 and scenario 2 have agents with equal attributes. Table 3 outlines the mean and median values the agents take across all ten runs, which is essential to demonstrate as it shows that generally, agents prefer to go home the most.

How observation changes behavior We compare vision attributes across scenario 1 to illustrate how social proximity or the capacity to witness others' behavior impacts individual and collective behavior. As demonstrated by the collection of all runs in Figure 4, when individuals initially share aligned empirical and normative expectations, what we may loosely describe as a social norm, the more perceptive an agent is of others' transgressions the quicker that norm erodes through time.

Table 3: Agent Attribute Summary Statistics

Agent Attributes	Mean	Median
empSens	1.4	2
norSens	1.56667	1
prefBar	1.3	1
prefHome	1.73333	2

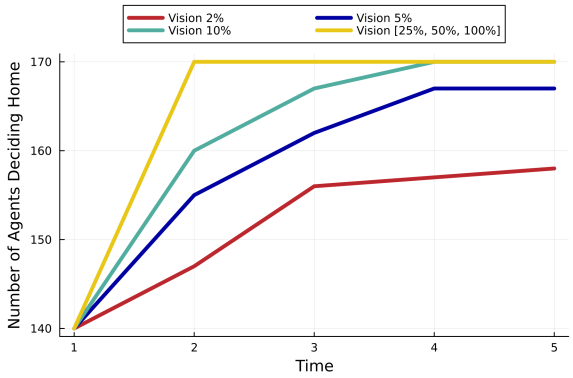

Figure 4: How Social Proximity Changes Behavior

It's noteworthy that initially, 30 agents behaved in opposition to their preferences, and depending on how 'close' or perceptive these agents are, ultimately determines the speed at which individuals default to their preferred behavior due to perceiving other's transgression, which, as shown in Table 3, is more commonly to 'go home.'

Comparing Scenarios Figure 5 compares scenarios one through six when vision is 5%, displaying the frequency of agents going home across every run. In scenarios 1, 2, and 4, there is initially an empirical expectation of going to the bar. Like what is observed in Figure 4, initially, 30 agents act against their preferences, going to the bar instead of home. The variation between these scenarios comes from the absence of, or inclusion of supporting or conflicting normative expectations. As demonstrated in Figure 5, scenario 1 is comparatively further away from the expected 170 frequency, as a normative expectation to go to the bar prevents individuals from choosing their preferred action, and, given that agents perceive fewer transgressions and may be more sensitive to normative expectations, individuals continue to adhere. For those who would prefer to go home, scenario 2 has a supporting normative expectation. Thus, we observe a steady rise in the transgression of the empirical expectations, as observing others transgress weakens the grip of the initial

expectation, which also describes the trend in the behavior of scenario 4.

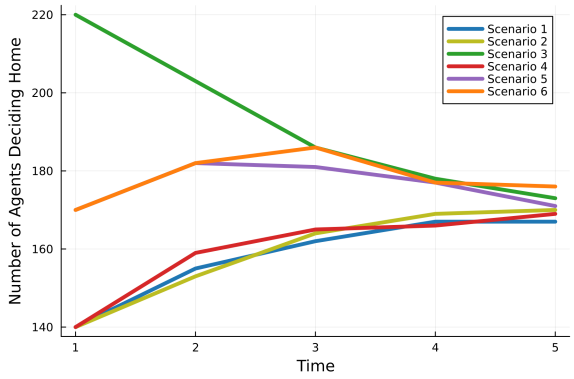

Figure 5: Comparing Scenarios at 5% Vision

In Scenario 6, agents prioritize their self-interest, given the absence of social expectations. However, as they observe the behavior of others, empirical expectations emerge, initially tilting behavior towards returning home and swaying those who would prefer to go to the bar; this is made clearer in Figure 6. Yet, over time, agents observe the behavior of those who go to the bar, and the weight of the empirical expectation to go home decreases, and agents' behavior gradually aligns with their preferences. Scenario 3 demon-

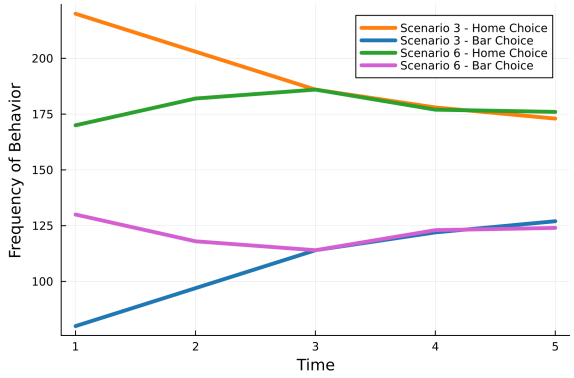

Figure 6: Empirical Expectation Formation at 5% Vision

strates the importance of the sensitivity attributes. As shown in Table 3, the median value for one's sensitivity to empirical expectations ranks higher than their normative sensitivity; as such, across the ten runs, more agents will cede to their empirical expectations over their normative expectations if they are in conflict, and, still rank higher than their own preferences. Naturally, as behaviors are observed by the agents and an opposing empirical expectation is formed, behavior reduces to following preferences. Finally, scenarios 5 and 6 demonstrate how the formation of empirical expectations modifies individual behavior under conditions of incomplete knowledge, knowing that behavior should settle at approximately 170 of the agents preferring home and conversely 130 at bar; the scenarios illustrate how preexisting and new empirical expectations through observed behavior, influence individual social decisions. Over time, the collective begins converging toward individual preferences due to mixed social signals.

Conclusion and Future Work

This paper presents an extension of the Expectation Event Calculus to endow agents with the capability to represent and reason over conditional empirical and normative expectation rules. We demonstrate this extension through a binary choice social scenario, with a single-agent demonstration under hardwired narratives and multi-agent simulations demonstrating behavioral trends over time. Through these simulations, we demonstrate how an individual's sensitivity to social expectations, given their perception of them, may lead to normative decisions beyond conformity and, in doing so, exhibit the potential for modeling richer social behaviors. While we demonstrate the important difference between the two forms of expectations for normative decision-making, we recognize the multiple research avenues available to continue this work. Namely, generalizing the formation of expectations to enable a broader method of associating behavior with social expectations. We anticipate this to be more straightforward for empirical expectations, but normative expectations may necessitate Theory of Mind capabilities (Lloyd and Lewis, 2024) to reason about the beliefs and goals of others. Additionally, future work may enrich normative decision-making through deontic language, weighting normative expectations on what is obligatory, permissible, or prohibited and grading these accordingly (Malle et al., 2019). In this paper, we used expectations as a form of social level aggregation, binding them to the employee membership, relaxing this, and implementing a contextual reference network to the decision-making process, which will likely enable enhanced social scenario modeling. As disembodied, space-less agents, convergence toward one's preferences was expected over time. While convergence may occur if the norms and supporting expectations remain stable, it is not guaranteed given dynamic environments and changing populations (migration and generational replacement) (Heidari et al., 2020); richer agent-based modeling may relax this constraint. In this study, agents lacked inherent motivation to communicate normative expectations, monitor others' behavior, and enforce sanctions. While not always essential for motivating norm compliance, sanctioning does play a pivotal role in regulating social behavior. Finally, future work may utilize this formalization to enable reflection on norms as they exist in society and the mind (Lewis and Sarkadi, 2024) to go beyond simply compliance and conformity.

References

Axelrod, R. (1986). An evolutionary approach to norms. *American political science review*, 80(4):1095–1111.

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Bicchieri, C. (2016). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford University Press.

Bicchieri, C. and Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, 23(2):161–178.

Bicchieri, C., Dimant, E., Gächter, S., and Nosenzo, D. (2022a). Social proximity and the erosion of norm compliance. *Games and Economic Behavior*, 132:59–72.

Bicchieri, C., Dimant, E., Gelfand, M., and Sonderegger, S. (2022b). Social norms and behavior change: The interdisciplinary research frontier.

Bicchieri, C. and Xiao, E. (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2):191–208.

Bicchieri, C. and Zhang, J. (2012). An embarrassment of riches: Modeling social preferences in ultimatum games. *Elsevier*.

Campenní, M., Andrighetto, G., Cecconi, F., and Conte, R. (2009). Normal= normative? the role of intelligent agents in norm innovation. *Mind & Society*, 8:153–172.

Castelfranchi, C. (2005). Mind as an anticipatory device: For a theory of expectations. In *International Symposium on Brain, Vision, and Artificial Intelligence*, pages 258–276. Springer.

Castelfranchi, C. (2016). A cognitive framing for norm change. In *Coordination, Organizations, Institutions, and Norms in Agent Systems XI: COIN 2015 International Workshops, COIN@ AAMAS, Istanbul, Turkey, May 4, 2015, COIN@ IJCAI, Buenos Aires, Argentina, July 26, 2015, Revised Selected Papers*, pages 22–41. Springer.

Clancey, W. J., Sierhuis, M., Damer, B., and Brodsky, B. (2005). Cognitive modeling of social behaviors. *Cognition and multi-agent interaction: from cognitive modeling to social simulation*, pages 151–184.

Coleman, J. S. (1994). *Foundations of social theory*. Harvard university press.

Cranefield, S. (2014). Agents and expectations. In *Coordination, Organizations, Institutions, and Norms in Agent Systems IX: COIN 2013 International Workshops, COIN@ AAMAS, St. Paul, MN, USA, May 6, 2013, COIN@ PRIMA, Dunedin, New Zealand, December 3, 2013, Revised Selected Papers 16*, pages 234–255. Springer.

Datseris, G., Vahdati, A. R., and DuBois, T. C. (2022). Agents.jl: a performant and feature-full agent-based modeling software of minimal code complexity. *SIMULATION*, 0(0):003754972110688.

Fehr, E. and Fischbacher, U. (2004). Social norms and human cooperation. *Trends in cognitive sciences*, 8(4):185–190.

Heidari, S., Wijermans, N., and Dignum, F. (2020). Agents with dynamic social norms. In *Multi-Agent-Based Simulation XX: 20th International Workshop, MABS 2019, Montreal, QC, Canada, May 13, 2019, Revised Selected Papers 20*, pages 112–124. Springer.

Kincaid, D. L. (2004). From innovation to social norm: Bounded normative influence. *Journal of health communication*, 9(S1):37–57.

Kölle, F. and Quercia, S. (2021). The influence of empirical and normative expectations on cooperation. *Journal of Economic Behavior & Organization*, 190:691–703.

Kovač, G., Portelas, R., Hofmann, K., and Oudeyer, P.-Y. (2021). Socialai: Benchmarking socio-cognitive abilities in deep reinforcement learning agents. *arXiv preprint arXiv:2107.00956*.

Kowalski, R. and Sergot, M. (1986). A logic-based calculus of events. *New generation computing*, 4:67–95.

Lapinski, M. K. and Rimal, R. N. (2005). An explication of social norms. *Communication theory*, 15(2):127–147.

Lewis, P. R. and Sarkadi, Ş. (2024). Reflective artificial intelligence. *Minds and Machines*, 34(2):1–30.

Lloyd, N. and Lewis, P. R. (2024). Towards reflective normative agents. In *Proceedings of the 18th Social Simulation Conference, European Social Simulation Association*. In-press.

Malle, B. F., Bello, P., and Scheutz, M. (2019). Requirements for an artificial agent with norm competence. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '19, page 21–27, New York, NY, USA. Association for Computing Machinery.

Mueller, E. T. (2014). *Commonsense reasoning: an event calculus based approach*. Morgan Kaufmann.

Sengupta, A., Cranefield, S., and Pitt, J. (2021). Solving social dilemmas by reasoning about expectations. In *International Workshop on Coordination, Organizations, Institutions, Norms, and Ethics for Governance of Multi-Agent Systems*, pages 143–159. Springer.

Sengupta, A., Cranefield, S., and Pitt, J. (2023). Generalising axelrod's metanorms game through the use of explicit domain-specific norms. In *International Workshop on Coordination, Organizations, Institutions, Norms, and Ethics for Governance of Multi-Agent Systems*, pages 21–36. Springer.

Shanahan, M. (2001). The event calculus explained. In *Artificial intelligence today: Recent trends and developments*, pages 409–430. Springer.

Wang, C. (2022). *Normative expectations and subjective beliefs: an incentivised experimental study*. PhD thesis, University College Cork.

Wijermans, N., Vriens, E., and Andrighetto, G. (2024). Heterogeneous agent decision-making – an empirically informed approach to behavioural types. In *Proceedings of the 18th Social Simulation Conference, European Social Simulation Association*. In-Press.