Seeing Through the Smoke: An Agent Architecture for Representing Health Protection Motivation Under Social Pressure

Veronika Kurchyna[1,2,3], Stephanie C. Rodermund[1,2], Ye Eun Bae [1],
Patrick Mertes[2], Philipp Flügger[2], Jan Ole Berndt[1] and Ingo J. Timm[1,2]

[1] *Cognitive Social Simulation, German Research Center for Artificial Intelligence, Behringstr. 31, 54296 Trier, Germany*

[2] *Business Informatics, Trier University, Universitaetsring 15, 54296 Trier, Germany*

[3] *Corresponding author*

Keywords: Agent Decision-Making, Social Simulation, Protection Motivation, Social Pressure

Abstract: Representing and emulating human decision-making processes in artificial intelligence systems is a challenging task. This is because both internal (such as attitude, perceived health or motivation) and external factors (such as the opinions of others) and their mutual interactions affect decision-making. Modelling agents capable of human-like behavior, including undesirable actions, is an interesting use case for designing different AI-systems when it comes to human-AI-interactions and similar scenarios. However, agent-based decision-models in this domain tend to reflect the complex interplay of these factors only to a limited extent. To overcome this, we enrich these approaches with an agent architecture inspired by theories from psychology and sociology. Using human health behavior, specifically smoking, as a case study, we propose an agent-based approach to combine social pressure within Protection Motivation Theory (PMT) to allow for a theory-based representation of potentially harmful behavior including both internal and external factors. Based on smoking in social settings, we present experiments to demonstrate the model's capability to simulate human health behavior and the mutual influences between the selected concepts. In this use case, the resulting model has shown that social pressure is a driving influence in the observable system dynamics.

1 Introduction

In recent years, AI applications have had an increasing impact on various aspects of people's professional and everday lives. Especially in health-related contexts there is a growing interest in designing interactive systems, e.g., in the field of assisted living or health care monitoring (Jovanovic et al., 2022; Qian et al., 2021). Unlike AI systems, however, human behavior is often irrational and guided by motivations or other internal processes, making it sometimes unpredictable and challenging for AI systems to respond appropriately to such behaviors. Agent-based social simulation (ABSS) has proven to be well-suited for examining the behaviors of individuals in response to social influences, individual needs, and dispositions (Davidsson, 2002; Squazzoni et al., 2014). Therefore, it can be used for investigating how cognitive and social factors interact and influence human behavior. ABSS is especially popular for its controlled environment and wide range of possibilities, which enables experiments that are difficult or impossible to implement in reality, and for testing mechanical characteristics of psychological and social science theories (Smith and Conrey, 2007). By simulating autonomous individuals, a complex system of various interactions emerges (Macy and Willer, 2002). For example, ABSS can be used to simulate a virus spreading among people (Tapp et al., 2022), family planning (Berndt et al., 2018), and crowd evacuation in the presence of fire (Wagner and Agrawal, 2014). Similarly, in health contexts, people's decisions are not only influenced by their own internal factors, such as attitude and health status, but also by external factors. Social pressure, for instance, refers to the expectations of the immediate social environment (Ajzen, 1991; Tesser, 1980) and it can induce individuals to conform to such expectations or resist them, depending on their dispositions.

By combining psychological theories explaining cognitive processes and concepts of social pressure, models can represent the mechanisms of real systems in an abstract manner. For instance, Protection Motivation Theory (PMT) is a widely used framework in psychology. It illuminates how individuals assess threats and make decisions about protective behaviors in situations involving health risks (Floyd et al., 2000; Hedayati et al., 2023), but also in other domains, for example information security (Mou et al., 2022) or pro-environmental behaviors (Kothe et al., 2019).

However, the nature of PMT is non-cyclical, not considering dynamic environments or the outcomes of one's actions. To address this issue, we use social pressure as a source of system change which creates a feedback-loop for agents. Smoking was chosen as use case, since both personal traits (internal factors; PMT) and social circumstances (external factors; social pressure) have an impact on the likelihood of the behavior (Xu et al., 2015). More precisely, individuals may be inclined to smoke more when perceiving a positive attitude towards smoking in their social environment and, conversely, are more likely to desist facing a negative attitude of others towards smoking; cf. e.g., (Ganley and Rosario, 2013).

This paper is structured as follows: Section 2 introduces concepts of social pressure, protective behaviors, and their implementation in ABSS to offer practical requirements for the conceptual agent architecture combining PMT and social pressure for health behaviors in general, which is introduced in Section 4. The implemented model is presented and analyzed in Section 5 using calibration and sensitivity analysis to demonstrate the concept for the use case of smoking. Then, the implications of the experiments are discussed as the results of the study. Finally, the findings are presented, focusing on potential avenues for future research to expand upon this approach in Section 6.

2 Background: Social Pressure and Self-Protection

We discuss concepts of social pressure as well as PMT with a special focus on smoking as use case. To situate this work in the context of current research, computational models applying these psychological mechanisms are presented.

Protective Behavior and Smoking: This paper addresses the use case of smoking, as a high proportion of deaths, especially in the northern hemisphere, is still attributable to the consequences of smoking (e.g. in 2019, this ratio was about 17% in Germany (Radtke, 2023)). One way of examining the possible factors that lead to an individual's decision to smoke is by using PMT to categorize different influencing factors.

Generally, those factors can be grouped into intrinsic and extrinsic factors. Intrinsic factors do not just include factors such as their own attitude towards smoking, but also more fundamental concepts such

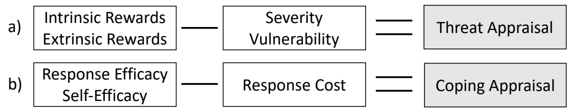

Figure 1: Threat (a) and Coping (b) Appraisal according to (Rogers, 1983).

as self-efficacy, which denotes the belief that one can successfully perform an action, e.g., quitting smoking (Rogers, 1983). Meanwhile, extrinsic factors include situational circumstances or the social environment's attitude towards a behavior (Xu et al., 2015).

According to PMT, individuals perform two cognitive processes: *threat appraisal* and *coping appraisal* (see Figure 1). These processes are triggered by information from internal or external sources and determine whether agents react adaptively (beneficial behavior) or maladaptively (harmful behavior) (Rogers, 1983).

In the use case of smoking, threat appraisal determines in how far adverse effects from harmful activities are acknowledged. Subjective perceptions of severity (e.g., how serious are the health consequences of smoking in general?) and vulnerability (how likely am *I* to suffer negative outcomes?) influence a person's willingness to react adaptively. High intrinsic rewards in the form of pleasure or satisfaction or extrinsic rewards, such as the feeling of belonging to a smoking group increase the probability of maladaptive behavior, in which a person will attempt to deny or downplay the risks of their choices.

Analogously, coping appraisal evaluates the recommended protective behavior (not smoking) and the agent's estimated ability to cope with and prevent the threat from occurring. This is determined by factors such as response efficacy (i.e., does it really bring health benefits to me if I quit smoking?) and self-efficacy (i.e., will I be able to quit smoking?). High efficacy increases the likelihood of adaptive responses, which is a healthy behavior (Rogers, 1983). Response costs, such as negative reactions from peers due to not smoking, may hinder adaptive responses.

Protection motivation can be expressed as either single or multiple actions, both one-time and repeated, or even inaction. That is, coping with a threat may require actively doing something, such as quitting smoking, or refraining, such as not starting to smoke (Rogers, 1983).

Social Pressure and Smoking: As mentioned previously, the attitude of peers impacts the appraisal of different strategies in the form of pressure. Social pressure can be defined as perceived normative pressure one receives from members of its immediate social environment (Tesser, 1980). When there is a conflict between the social network's opinion and the personal attitude towards a behavior, a person can experience *cognitive dissonance*, where different interests conflict with each other (Festinger, 1957). As a way to reduce this cognitive dissonance, a person typically has two options. On the one hand, they may change the comparison groups by seeking a new social environment with a preferably similar attitude. On the other hand, they yield to social pressure by adjusting their own disposition which can eventually lead to a change in behavior(Festinger, 1957; Tesser, 1980).

3 Models of Social Pressure and Protective Behavior

Models of social pressure are mainly expressed through normative multiagent systems, where the behavior of agents is rewarded for norm-compliance and sanctioned for deviation. An important distinction between types of models lies in the choice of mechanism to express pressure. On the one hand, explicit norms for behavioral regulation (see, e.g. (Castelfranchi et al., 1998)) are often distributed by communication between agents. On the other hand, perceived social pressure is more subtle, and is often tied to a threshold to determine when agent decisions are impacted by pressure (see, e.g. emotional autonomy). This leads to agents acting according to perceived social pressure, even in case of conflict with the agent's own attitude (Dechesne et al., 2013).

Besides explicit and perceived norms, there are also different ways rewards and punishments are included into the model. Typically, there is a distinction between direct sanctioning, such as monetary punishment, or indirect sanctioning in the form of emotional penalties. In such cases, decreased trust in nonconforming agents leads to reduced odds of cooperation, such as the conclusion of contracts (Nardin et al., 2016), in the future, sometimes even leading up to the exclusion of agents from the social network (Perreau de Pinninck et al., 2010).

Norms, rewards and punishments can be communicated (Savarimuthu et al., 2009) or only suspected and perceived through some kind of filter (Hashimoto and Egashira, 2001) just as the subjectively perceived pressure does not have to correspond to the actually exerted pressure or attitudes of the social network. Additionally, some individuals may have a stronger impact on social pressure than others, e.g., based on the strength of relationships or similarities to the person being influenced (Savarimuthu et al., 2008). It is important to note that changing one's attitude does not necessarily equal changing behavior, and will vary between different models (Sheeran and Webb, 2016).

As a result of these different mechanisms, agents can develop cognitive dissonance, as described in Section 2, where personal preferences no longer align with the environment. Typical dissolution strategies include changing attitude (Dechesne et al., 2013) towards adaption of other's opinions or exclusion from the social network (Perreau de Pinninck et al., 2010).

In the conceptual framework of PMT, the social network plays an important role by presenting an environmental source that can affect decisions, e.g., in the form of verbal persuasion or observational learning, such as in Hear et al. (Haer et al., 2016) in the context of flood risk management and the effect of different communication strategies and countermeasures based on communication (Badham and Gilbert, 2015). Mostly, PMT is used in connection with climate-related decisions, e.g., with regard to the implementation of preventive measures (see, e.g., (Wens et al., 2020; McEligot et al., 2019)) and the general vulnerability to the consequences of climate change (Krömker et al., 2008), or in health protective behavior in the context of infectious diseases ((Abdulkareem et al., 2018; Martin-Lapoirie et al., 2023)).

The impact of a social network on an agent's decision is often not considered at all (see, e.g., (Krömker et al., 2008; McEligot et al., 2019)) or is only regarded as an additional source of information (Abdulkareem et al., 2018). While works like (Haer et al., 2016; Badham and Gilbert, 2015) often include the social network as a factor in decision-making, both models lack a clear definition and the effects of social pressure with regard to the development of cognitive dissonance as well as the agent's responses to it.

Similarly to (Kurchyna et al., 2022), who examined another pair of cognitive and social theories for the use case of heart disease, this work investigates the interplay of cognitive dissonance, PMT, and the realism of resulting system dynamics.

4 A Conceptual Agent-Based Model of Protection Motivation and Social Pressure

Detached from the use case of smoking, this section introduces the conceptual model of health protection motivation under consideration of social pressure. Agents experience social pressure from their environment and determine whether to perform a health-related behavior (e.g., to smoke or not to smoke) based on the concepts of PMT (see Figure 2).

The environment consists of a set of agents A. Inspired by the approach presented by López y López et al. (López y Lopéz et al., 2002), the initialization of these agents is based on pre-defined archetypes characterized by a core desire. To reduce model complexity for demonstration purposes, this model focuses on desires that we assume to be relevant for most use cases: hedonism ($desire_h \in [0,1]$), security ($desire_s \in [0,1]$), and conformity ($desire_c \in [0,1]$) as defined by Schwartz (Schwartz, 1992). For example, those who care more to have fun have a strong desire for hedonism, those who are more careful when it comes to health and safety have a strong desire for security while those who favor social cohesion have a strong desire for conformity. The closer a value to 0, the weaker the desire an agent has while the closer value to 1, the stronger the desire is. Agents who desire pleasure for themselves the most (Heidari et al., 2020) may neglect their need for safety (Maddux and Rogers, 1983). Additionally, a desire to conform to expectations may interfere with the satisfaction of either desire (Tesser, 1980). The archetype determines the dominant desire of the agent, while the other two desires are present to a lesser degree and may influence the choice of an agent in edge cases. There is no explicit categorization of agents within the model, but is rather implied through their desires. However, for the ease of communications, we refer to the agents by their primary desire, the strongest one. Hence, the agent's environment can be defined as $A = \{Hedonists, Conformists, Self-Protectors\}$.

Each agent has a network N consisting of two disjoint groups: friends and locals. Friends ($F \subset A$) is a broad category which includes persons the agent has friendly relations with (like family members or friends). This set's size is fixed to $num\_friends \in \mathbf{N}$. Locals ($L \subset A$) contains people the agent encounters in its local environment. This may include colleagues, people on public transport or various service providers. Belonging to this set is determined by neighborhood in the 2D-space of the simulation. While the social network of friends remains persistent for the agent, random movement at each step leads to a frequent variation of the local environment. For more complex use cases, routines and the formation of long-term relationships between agents would replace the random movements.

Over the course of the simulation, agents experience situations in which social pressure is present and their actions are chosen based on PMT. This choice of action is influenced not only by the network of the agents, but also by their own attributes. Agents have an initial attitude towards a health behavior, based on their hedonism and security desires while conformity desire determines how likely they will follow majority opinion. This attitude att is mapped as a numerical value to an interval $[0,1]$. An agent's opinion towards a behaviour can range from strong aversion (0) to strong favor (1), introducing nuance that allows agents to choose a harmful behaviour, such as smoking, even if they are not in complete favour of it. This attitude is a proxy for the actual behavior and it allows tracking of changes over time. Due to the phenomenon of the intention-behavior gap (Sheeran and Webb, 2016), the translation of attitudes into actions requires further examination beyond the scope of a proof of concept. Another important attribute is the general $health\_status \in [0,1]$ of the agent, influencing the way agents assess their own vulnerability towards a threat. Here, the value 1 represents perfect health, while values towards 0 represent increasingly poor health such as chronic conditions and typical health risk factors. Due to the delay between action and effect, which makes it difficult to observe the actual causal relationship, the health status is kept static throughout the simulation.

In reality, factors such as economic status correlate with smoking and its frequency (Casetta et al., 2017). However, under the assumption that needs are a mediating factor between characteristics and observable actions, such factors are implied through both the agent's archetype and their initial attitude towards smoking. Depending on the purpose of the study, especially explainability and the tailoring of targeted interventions, implied characteristics may not suffice, and explicit implementation would be recommended in such a case.

The following paragraphs discuss the main components of the proposed concept, i.e., the social pressure component and the PMT components.

4.1 Social Pressure Component

Social pressure in this model is modeled by evaluating the perceived attitudes against the agent's own position. In this case, there is no filtering through subjective interpretation or other approaches from opinion dynamics, and instead, the actual attitude of the agent is being perceived as abstraction from how exactly this social pressure is formed and exerted, focusing on its effects.

Hence, agents consider the attitudes of remote peers ($\overline{att_f} \in [0,1]$) as well as those of locals in their current physical environment ($\overline{att_l} \in [0,1]$). While $\overline{att_f}$ takes the bidirectional influence $connection\_strength \in [0,1]$ of agents on each other into account, there is no such distinction for locals. As a result, friends are weighted differently and can have a larger or smaller influence compared to a local, unlike some alternative approaches in which random

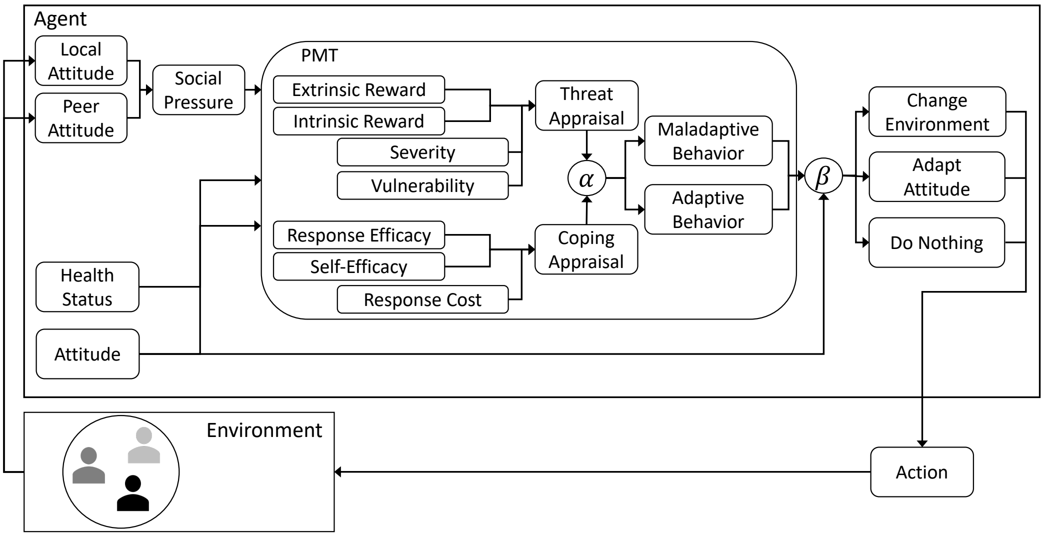

Figure 2: Agent-based model of PMT under social pressure.

encounters always have less impact than friends (Li et al., 2022), permitting situations of accute peer pressure exerted by local contacts. The attribute reflecting *connection_strength* has an impact on the perceived attitude of surrounding agents, and in certain cases this can be perceived to be stronger than it actually is and vice versa. For instance, $\overline{att_f}$ is determined as follows (Equation 1):

$$\overline{att_f} = \frac{\sum_{i=1}^{|F|} att_i * connection\_strength_i}{|F|} \tag{1}$$

Social pressure results from the sum of both average attitudes of locals and friends weighted according to the number of connections of each type. The resulting social $pressure \in [0,1]$ is thus determined by the distance between the agent's own attitude att and the values for $\overline{att_f}$ and $\overline{att_l}$ which are included proportionally to the size of the whole network N (see Equation 2).

$$pressure = |att - \frac{\overline{att_f} * |F| + \overline{att_l} * |L|}{|N|}| \tag{2}$$

4.2 Protection Motivation Components

As described in Section 4, agents can choose between two behavior modes: *maladaptive* or *adaptive* behavior. The decision for either is made based on factors such as perceived social pressure, health status and personal attitudes. To decide for one of two behavior modes, the agent first computes the threat appraisal ($appraisal_{threat}$) and coping appraisal ($appraisal_{coping}$). A $behavior \in [m,a]$ is maladaptive (m), if $appraisal_{threat} > appraisal_{coping}$ and adaptive (a) if $appraisal_{coping} > appraisal_{threat}$ (see Figure 2, Function α).

$Appraisal_{threat}$ is based on rewards that can be obtained from performing a non-beneficial behavior (intrinsic i and extrinsic e) in the presence of individually perceived severity ($severity_p \in [0,1]$) and vulnerability ($vul \in [0,1]$). Based on the theoretical equation in Figure 1, threat appraisal of agents is calculated as follows:

$$\begin{aligned} appraisal_{threat} = &\frac{1}{2} * ((reward_i + reward_e) - \\ &(severity_p + vul)) \end{aligned} \tag{3}$$

In this use case, $reward_i$ is the own attitude representing the pleasure from smoking, while $reward_e$ is the social pressure towards smoking — conforming to this expectation is perceived as reward. $Severity_p$ is moderated by individual perceptions (Badham and Gilbert, 2014) and defined as

$$severity_p = severity_o + ((1 - desire_c) * (desire_s - severity_o)) \tag{4}$$

The objective severity ($severity_o$) is moderated by $desire_s$ and $desire_c$ to express the individual levels of anxiety (Badham and Gilbert, 2014). Vul is the assessment of personal applicability of a threat, and thus individual health (Rogers, 1983) influences its computation as follows:

$$vul = 1 - (health + (desire_s - health) * health) \tag{5}$$

$Appraisal_{coping}$ examines the self-efficacy (*efficacy$_s$*) and the response efficacy (*efficacy$_r$*) of an action while evaluating the response cost ($cost_{res}$) of the action. If $appraisal_{threat}$ is high, quitting non-beneficial behavior is perceived as more effective, which leads to an adaptive response. It is defined as follows:

$$appraisal_{coping} = \frac{efficacy_r + efficacy_s}{2} - cost_{res} \quad (6)$$

where $efficacy_r$ is defined as

$$efficacy_r = att * \frac{vul + severity_p}{2} \quad (7)$$

with perceived attitude (att_p) defined as

$$att\_p = \frac{att_f * |F| + att_l * |L|}{|N|} \quad (8)$$

and $cost_{res}$ defined as

$$cost_{res} = \frac{att + (desire_c * att\_p)}{2} \quad (9)$$

signifying both losing out on pleasure as well as defying social norms. If a person has a highly favourable attitude towards a non-beneficial behavior, desisting is difficult. Likewise, a strong attitude (in either direction) in the environment att_p is hard to resist, with both cases increasing response cost $cost_{res}$ (Huang and Wen, 2014).

$Efficacy_s$ depends on both own attitude and the opinions of the environment, as defined in Equation 10.

$$efficacy_s = \begin{cases} 1 - att & \text{if } att > 0.5 \land att\_p \leq 0.5 \\ 1 - att\_p & \text{if } att \leq 0.5 \land att\_p > 0.5 \\ 1 - \frac{att + att\_p}{2} & \text{else} \end{cases} \quad (10)$$

Here, we set 0.5 as the transition point where an agent's attitude toward a certain behavior switches from *not favorable* to *favorable* or vice versa.

In each step of the simulation, all agents make a decision to act and impact on their immediate surrounding agents and friends. Each agent calculates the threat and coping appraisals and acts according to their decision. There are three potential actions as defined in Figure 3, Function β, based on the cognitive dissonance theory by Festinger (Festinger, 1957) earlier introduced in Section 2. :

1. *Create new consonant cognitions:* An agent cuts ties with friends that exert too much pressure on it and finds new friends as well as moves on to another location to switch their surrounding agents (*change environment*). The user-defined variable $contact\_termination \in [0, 1]$ determines at which degree of difference the entire local network and friends with large attitude differences are replaced to preserve the set $num\_friends$.
2. *Change cognition:* An agent bows to the social pressure and adjusts its own desires and attitudes towards the group mean (*adapt attitude*). By adapting att, $desire_s$ and $desire_h$, the agent reduces the distance between its internal states and its social network. The amplitude of changes is moderated by a user-determined change rate ($c\_r \in [0, 1]$). As a result, cognitive dissonance is reduced.
3. *Reduce the importance of dissonant cognitions:* An agent does nothing if its own attitude is strong enough to resist external influence or if *pressure* is low due to similar attitudes (*do nothing*).

In Figure 3, both the actions *change environment* as well as *adapt attitude* are chosen by the agent if its attitude att is either specifically high or low (below or above a determined threshold thd_a and 1-thd_a) and the social network differs from that attitude in a way, that it exceeds the value of c_r. The agent decides for a major change in its network only if $min\_stay \in [0, 10]$ steps have passed since previous changes to prevent a constant exchange of the agent's social network. If this is true, there are two cases, in which the agent might adapt its network structure: In the first case (a), the agent has a strong attitude towards the respective behavior and where $appraisal_{threat}$ is stronger than $appraisal_{coping}$. The behavior thus is defined as maladaptive, i.e., the agent performs unhealthy behaviours. Here, agents with a negative attitude towards the harmful behavior are excluded from the agents sphere of influence to reduce cognitive dissonance. The second case (b) has a contrasting effect. The agent has a negative attitude towards the behavior. However, because the majority of its network has a positive attitude and the agent's $appraisal_{coping}$ is greater than its $appraisal_{threat}$, the agent shows an adaptive behavior

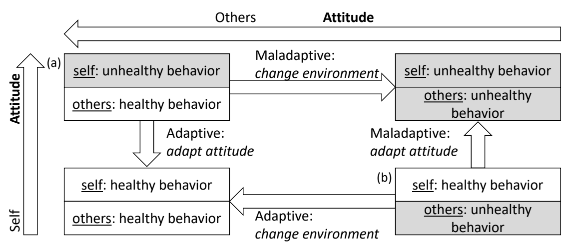

Figure 3: Decision for maladaptive or adaptive action depends on the level of attitude (healthy or unhelathy behavior) and the social pressure of the network (corresponds to Figure 2, Function β).

and excludes friends and locals that might influence it towards the harmful behavior. *Adapt attitude* makes a similar distinction. If the agent's *att* is rather favorable towards smoking and the social network exerts much pressure while $appraisal_{coping}$ is stronger than the $appraisal_{threat}$, leading to the adaptive behavior, the agent adapts its attitude and desires to reduce cognitive dissonance to its network. In the second case, the agent's $appraisal_{threat}$ is higher, which leads to the maladaptive behavior. The agent adapts its *att* towards unhealthy behavior to be more positive-minded towards it. In contrast to the option of adapting the network of friends and locals, *adapt attitude* can be selected in each point in time. If none of these cases applies, the agent chooses *do nothing*.

5 Simulating Social Pressure Towards Smoking within Protection Motivation Theory

The implementation of the concept is demonstrated with smoking as use case. Over the course of the simulation, with steps representing an abstract concept of 'situations', agents experience social pressure. For instance, people feel compelled by their peers to either smoke or refrain from doing so, depending on the group's attitude. The act of smoking itself is not modelled explicitly but is expressed via the attitude as proxy. The model was implemented using NetLogo 6.3.0. For the explanation of the model validation process, its in- and outputs are described along with the calibration process and exploration using sensitivity analysis.

5.1 Experiment Design and Calibration

The primary goal of this experiment is not a credible simulation of smoking (and its cessation), but rather to provide a use case in which this combination of social pressure and PMT is explored. The analyzes and experiments are performed with this goal in mind.

The implemented model includes a set of parameters which can be configured by users, listed and described in Table 1. These parameters mostly deal with the composition of agent archetypes and the rate at which its state changes. To examine the model behavior, three types of output variables were observed:

- *Smoking behavior:* The number of habitual smokers ($att \geq 0.75$), casual smokers ($0.25 < att < 0.75$) and convinced non-smokers ($0.25 \geq att$) in the population
- *Desires:* The average desire for conformity, security and hedonism in the population, which conveys the overall population trend
- *Actions:* The number of times the actions *do nothing*, *change environment* and *adapt attitude* chosen by the agents

A primary objective of this model's calibration is to examine whether there exists a parameter combination that results in a stable behavior without a severe stagnation. While habitual behaviors, such as smoking, have long delays between action and effect, other behaviors may show much higher degrees of change and more dynamic situations.

To establish a baseline parameter configuration at which real-world observations are imitated by the model, a calibration using BehaviorSearch, a tool integrated in NetLogo, was performed. In this study, a distribution of 29% average smokers in Germany was expected to be achieved, based on available empirical data (Starker et al., 2022). In prior trial experiments, the model showed that agents tend to build stable, homogenous groups after approximately 30 steps, each representing a situation in which social pressure might have been experienced. Thus, the goal of calibration is to minimize the deviation of the number of smokers from the target value at the end of the simulation, once equilibrium has been reached. Additionally, the composition of agent types should add up to 100%. Further, a system in which the initial number of smokers and non-smokers matches statistical data and starts in a stable equilibrium is not conducive to the creation of a realistic system, which is why static behavior (*do nothing*) is among the observations to be minimized. As a result, Equation 5.1 was used as minimization function for BehaviorSearch, with the weights chosen to adjust the priority of the different conditions according to their importance.

$$\begin{aligned} fitness = &\Big(\Big|\frac{|Smokers|}{|N|} - 0.3\Big| * 1000\Big) + \\ &(\sum do\_nothing * 0.1) + \\ &((|Conformists| + |Hedonists| + \\ &|Security|) - 1) * 10 \end{aligned} \tag{11}$$

The results of the calibration are included in the last column in table 1. To better understand how the parameters interact with each other, however, an in-depth sensitivity analysis is required. Still, a prior calibration as a first step towards model verification is necessary to ensure that a model can achieve valid results. In this case, while competing solutions exist, one possible configuration observes a dominant proportion of hedonistic agents, with security-oriented

Table 1: Parameters, ranges and the recommended default values based on calibration. Population = 100 was set as fixed parameter to restrict the search space.

Parameter	Description	Range	Value
n_people	Number of agents	int	100
Conformists	% of conformists	0 - 1.0	0.23
Hedonists	% of hedonists	0 - 1.0	0.64
Security	% of health-aware	0 - 1.0	0.18
pressure_mod.	Strength of perceived pressure	0 - 1.5	1.25
contact_term.	Allowed difference between attitude and perceived pressure before changing environment	0 - 1	0.4
min_stay	Min. number of actions in new env. before changing is possible	0 - 10	3
moving_distance	move forward set number of units	0 - 10	2
number_friends	Number of peers each agent has	0 - 10	2
change_rate	Rate of change of attitude	0 - 0.1	0.01

people and conformists in similar numbers. The pressure modifier is slightly above the intended default of 1, amplifying the intended effect. At 0.4, agents are rather tolerant to differences in opinions, and will only change their environment in the case of severe conflict of attitudes. With small spatial movements and a small social circle of two close friends as the default, agents are thus more likely to hold on to their environment than to change it, further confirmed by the low change rate which favours slow, gradual changes over abrupt shifts.

When testing viable parameter combinations, it was observed that the system typically reaches a stable state after approximately 30 simulation steps, with agents forming homogenous groups similar to various segregation models (Goles et al., 2011).

5.2 Simulation Results and Discussion

Once the feasibility of realistic model behavior was confirmed using calibration, a sensitivity analysis was performed based on the example provided by Jaxa-Rozen and Kwakkel (Jaxa-Rozen and Kwakkel, 2018). To show how the different design decisions in the model impact the overall behavior, we analyze the first- and second-order sensitivity indices of the parameters in regard to the observation variables. A major finding was the initial distribution of character types ($Hedonists$, $Conformists$ and $Self-Protectors$) dominating the model behavior, which aligns with empirical data that confirms most will retain their initial attitude, e.g. those who are smokers and hedonists at the same time will continue to smoke, while those who are non-smokers and health-conscious are unlikely to start smoking (Høie et al., 2010).

To analyze the effects of other variables on the different observations, the calibrated values for the population distribution were used and the input variable was removed from the sensitivity analysis. The change rate (c_r), describing the amplitude of changes agents make and how strongly they perceive the pressure, is the major determinant for model behavior, accounting for approximately 70% of the variance in most observations. The pressure modifier, adjusting the strength of pressure agents experience, offered insight into the role that peer pressure plays in the application of PMT in this model. Figure 4 summarizes the findings regarding the impact of the pressure modifier, and thus the social pressure altogether, with respect to the different observation variables.

Social pressure contributes towards the distribution of behavioral types in the population. The low influence on the number of habitual smokers, in combination with the number of non-smokers and casual smokers being more sensitive to social pressure, indicates that people with moderate to low attitudes are more strongly affected by external influences. In that regard, the population shows a strong leaning towards persuading agents to give up smoking. Strong interactions between the pressure modifier and other parameters, such as the number of friends and contact termination rates, indicate that social pressure is a major contributor towards reduction of smoking, contrary to the initial expectation that social pressure might be a driving factor in previously non-smoking agents starting to smoke. This gradual reduction of smoking attitude does indeed align with the global trend of a decreasing number of smokers (WHO, 2021).

Moreover, social pressure is strongly related to agents deciding to be inactive — its absence or presence is the main explanator for the frequency of inaction. While the desire for conformity is barely affected by pressure, the average desire for hedonism and security are dominated by the pressure modifier — without social pressure, agents will retain their initial beliefs. Assuming that agents will, at the beginning of the simulation, already perform the behavior that fulfills their desires, they will have no incentive to change their attitudes or perform any adaptive or maladaptive behaviors due to the lack of external influences. In this regard, this study comes to the same conclusions as (Kurchyna et al., 2022), which noted stagnation in a system with purely cognitive processes that do not receive any external inputs in the form of social influence.

Such findings indicate that the combination of external/social and internal/cognitive components is well-suited for autonomous systems of continuous simulation without clearly defined end states. In this example, agents pursue the continuous goal of a balanced needs satisfaction in a mixed social environment that may hinder or encourage their quest for homeostasis. Due to the mechanistic design of PMT and the two-layered approach to contact networks, the

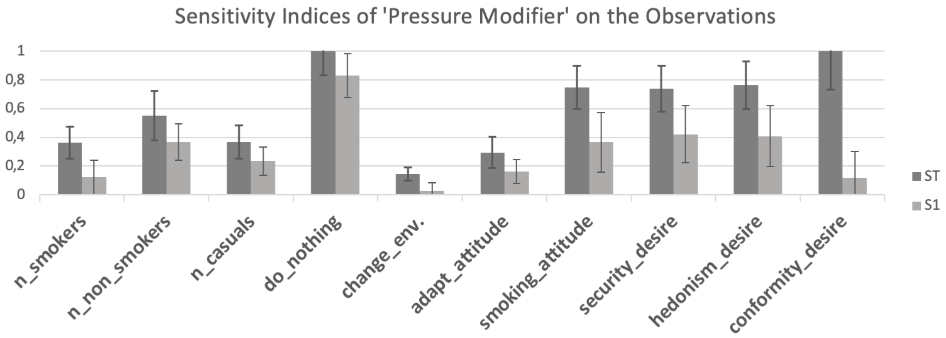

Figure 4: First (S1) and total-order (ST) indices for pressure sensitivity towards different observations.

ideas of this model can also be transferred to other application areas in which agents act under uncertainty and conflicting goals and options. A good example for such scenarios is human-AI-interaction and assistance systems, which need to be able to anticipate and consider inconsistent goals and seemingly irrational user behavior.

6 Conclusion

The study of the decision-making processes leading to people's behavior and large-scale emergent phenomena is an important research topic in ABSS. To examine how psychological theories can be used to model decision-making processes of agents in continuous simulations, an implementation of PMT and social pressure for the use case of smoking was presented. While PMT provides the cognitive process leading to decisions within agents, the concepts of cognitive dissonance and social pressure fill the model with life. This proof-of-concept implementation shows that without social pressure, cognitive concepts risk a vacuum in which agents lose incentive to act, consistent with previous findings in similar types of studies; cf. e.g., (Kurchyna et al., 2022).

In the broader context of simulations in which agents interact under uncertainty and with potentially conflicting goals, concepts such as PMT may be well suited to include a variety of factors into decision-making without having the large scope of alternative approaches such as the theory of planned behavior (Ajzen, 1991), which may be more suited for use cases which require more complex behaviors and deliberate planning. The insights from the model will enable AI assistance systems to support individuals in their decision-making processes and motivate them to pursue more desirable actions when faced with conflicting objectives and perspectives on each action category. Additionally, this use case demonstrated two-layered networks of contacts exerting potentially contradictory influences. The insights provided by this model and its evaluation are promising in regard to various dimensions of future works.

Such future work therefore involves validating the model using empirical data, as well as testing its transferability to other health-related application areas. Furthermore, the authors plan on improving the model with the aim of adjustments such as more individualized behavior. For example, in the current model, adjustments in attitude and social network are determined by the change rate and the contact termination threshhold. However, these variables are set to a global static value. According to psychological theory, a person's emotional autonomy defines an individual bound for each, which determines when perceived social pressure will contribute to behavior adjustment (Savarimuthu et al., 2008). Additionally, the similarity of the person to its surrounding network may play a role in the adaptation of behavior. If the person does not identify enough with the group, the social pressure exerted will have no influence on the person's behavior (Terry and Hogg, 1996). In the same vein, a more elaborate model of social pressure, which includes detailed perception of other's opinions, would provide benefits to the realistic representation of the agent's decision making process. The importance of other characteristics (such as demographic variables) should be examined as additional variables, since, e.g., a person's age may contribute to willingness to change and thus adopt protective behaviors (Badham and Gilbert, 2015).

In general, the authors intend to extend and adapt the agent architecture based on the BDI approach (Rao and Georgeff, 1995) to achieve a more differentiated model of individual desires as well as complex decision-making, including planning steps towards a behavior and the formation of habits (see, e.g., (Kurchyna et al., 2022)). Finally, a future step extends the model towards the possibility of testing different intervention strategies, for example, different communication strategies (c.f. (Haer et al., 2016)), in order to lead agents to beneficial behavior such as quitting smoking.

ACKNOWLEDGEMENTS

This submission is a result of the work in the context of the SEMSAI Project, funded by the German Federal Ministry of Education and Research under the grant number 031L0295A.

REFERENCES

Abdulkareem, S. A., Augustijn, E.-W., Mustafa, Y. T., and Filatova, T. (2018). Intelligent judgements over health risks in a spatial agent-based model. *International journal of health geographics*, 17(1):1–19.

Ajzen, I. (1991). The theory of planned behavior. In *Organizational behavior and human decision processes*.

Badham, J. and Gilbert, N. (2014). Personal protective behaviour during an epidemic. In *Social Simulation Conference-SSC 2014*.

Badham, J. and Gilbert, N. (2015). Tell me design: Protective behaviour during an epidemic. CRESS.

Berndt, J. O., Rodermund, S. C., and Timm, I. J. (2018). Social contagion of fertility: an agent-based simulation study. In *2018 Winter Simulation Conference*. IEEE.

Casetta, B., Videla, A. J., Bardach, A., Morello, P., Soto, N., Lee, K., Camacho, P. A., Hermoza Moquillaza, R. V., and Ciapponi, A. (2017). Association between cigarette smoking prevalence and income level: a systematic review and meta-analysis. *Nicotine & Tobacco Research*, 19(12):1401–1407.

Castelfranchi, C., Conte, R., Paolucci, M., et al. (1998). Normative reputation and the costs of compliance. *Journal of Artificial Societies and Social Simulation*, 1(3):3.

Davidsson, P. (2002). Agent based social simulation: A computer science view. *Journal of artificial societies and social simulation*, 5(1).

Dechesne, F., Di Tosto, G., Dignum, V., and Dignum, F. (2013). No smoking here: values, norms and culture in multi-agent systems. In *AI and Law*.

Festinger, L. (1957). A theory of cognitive dissonance. Stanford University Press.

Floyd, D. L., Prentice-Dunn, S., and Rogers, R. W. (2000). A meta-analysis of research on protection motivation theory. In *Journal of applied social psychology*.

Ganley, B. J. and Rosario, D. I. (2013). The smoking attitudes, knowledge, intent, and behaviors of adolescents and young adults: Implications for nursing practice. *Journal of Nursing Education and Practice*, 3(1):40.

Goles, D. N., Goles, E., and Rica, S. (2011). Dynamics and complexity of the schelling segregation model. In *Phys. Rev. E. 83*. American Physical Society.

Haer, T., Botzen, W., and Aerts, J. (2016). The effectiveness of flood risk communication strategies and the influence of social networks—insights from an agent-based model. In *Environmental Science & Policy*.

Hashimoto, T. and Egashira, S. (2001). Formation of social norms in communicating agents with cognitive frameworks. *Journal of Systems Science and Complexity*.

Hedayati, S., Damghanian, H., Farhadinejad, M., and Rastgar, A. A. (2023). Meta-analysis on application of protection motivation theory in preventive behaviors against covid-19. *International Journal of Disaster Risk Reduction*, page 103758.

Heidari, S., Jensen, M., and Dignum, F. (2020). Simulations with values. In *Advances in Social Simulation: Looking in the Mirror*. Springer International Publishing.

Huang, C. Y. and Wen, T. H. (2014). A novel private attitude and public opinion dynamics model for simulating pluralistic ignorance and minority influence. In *Journal of Artificial Societies and Social Simulation*.

Høie, M., J., M. I., and Rise, J. (2010). An extended version of the theory of planned behavour: Prediction of intentions to quit smoking using past behaviour as moderator. In *Addiction Research & Theory*.

Jaxa-Rozen, M. and Kwakkel, J. H. (2018). Pynetlogo: Linking netlogo with python. In *Journal of Artificial Societies and Social Simulation*.

Jovanovic, M., Mitrov, G., Zdravevski, E., Lameski, P., Colantonio, S., Kampel, M., Tellioglu, H., and Florez-Revuelta, F. (2022). Ambient assisted living: scoping review of artificial intelligence models, domains, technology, and concerns. *Journal of Medical Internet Research*, 24(11):e36553.

Kothe, E. J., Ling, M., North, M., Klas, A., Mullan, B. A., and Novoradovskaya, L. (2019). Protection motivation theory and pro-environmental behaviour: A systematic mapping review. *Australian Journal of Psychology*, 71(4):411–432.

Krömker, D., Eierdanz, F., and Stolberg, A. (2008). Who is susceptible and why? an agent-based approach to assessing vulnerability to drought. *Regional Environmental Change*, 8:173–185.

Kurchyna, V., Rodermund, S., Berndt, J. O., Spaderna, H., and Timm, I. J. (2022). Health and habit: An agent-based approach. In *KI 2022: Advances in Artificial Intelligence: 45th German Conference on AI Proceedings*. Springer-Verlag Berlin, Heidelberg.

Li, L., Wang, S., and Lin, Y. (2022). The casual effect of relational mobility on integration of social networks: An agent-based modeling approach. In *Current Psychology*.

López y Lopéz, F., Luck, M., and d'Inverno, M. (2002). Constraining autonomy through norms. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 2*.

Macy, M. W. and Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. In *Annial review of sociology*.

Maddux, J. and Rogers, R. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. In *Experimental Social Psychology*.

Martin-Lapoirie, D., D'onofrio, A., McColl, K., and Raude, J. (2023). Testing a simple and frugal model of health protective behaviour in epidemic times. *Epidemics*, 42:100658.

McEligot, K., Brouse, P., and Crooks, A. (2019). Sea bright, new jersey reconstructed: Agent-based protection theory model responses to hurricane sandy. In *2019 Winter Simulation Conference (WSC)*, pages 251–262. IEEE.

Mou, J., Cohen, J. F., Bhattacherjee, A., and Kim, J. (2022). A test of protection motivation theory in the information security literature: A meta-analytic structural equation modeling approach. *Journal of the Association for Information Systems*, 23(1):196–236.

Nardin, L. G., Balke-Visser, T., Ajmeri, N., Kalia, A. K., Sichman, J. S., and Singh, M. P. (2016). Classifying sanctions and designing a conceptual sanctioning process model for socio-technical systems. *The Knowledge Engineering Review*, 31(2):142–166.

Perreau de Pinninck, A., Sierra, C., and Schorlemmer, M. (2010). A multiagent network for peer norm enforcement. In *AAMAS 2010*.

Qian, K., Zhang, Z., Yamamoto, Y., and Schuller, B. W. (2021). Artificial intelligence internet of things for the elderly: From assisted living to health-care monitoring. *IEEE Signal Processing Magazine*, 38(4):78–88.

Radtke, R. (2022, accessed on 05 March 2023). Anteil rauchbedingter Todesfälle in ausgewählten Ländern im Vergleich der Jahre 1990 und 2019. In *https://de.statista.com/statistik/daten/studie/1172056/\umfrage/anteil-rauchbedingter-todesfaelle-nach-laendern/*. Statista.

Rao, A. S. and Georgeff, M. P. (1995). Bdi agents: from theory to practice. In *International Conference on Multi-Agent Systems*.

Rogers, R. (1983). Cognitive and psychological processes in fear appeals and attitude change: A revised theory of protection motivation. In *Social psychophysiology: A sourcebook*.

Savarimuthu, B., Cranefield, S., Purvis, M., and Purvis, M. (2008). Role model based mechanism for norm emergence in artificial agent societies. In *Coordination, Organizations, Institutions, and Norms in Agent Systems*. Springer Berlin Heidelberg.

Savarimuthu, B. T. R., Purvis, M., Purvis, M., and Cranefield, S. (2009). Social norm emergence in virtual agent societies. In *Declarative Agent Languages and Technologies VI: 6th International Workshop, DALT 2008, Estoril, Portugal, May 12, 2008, Revised Selected and Invited Papers 6*, pages 18–28. Springer.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology*. Academic Press.

Sheeran, P. and Webb, T. L. (2016). The intention–behavior gap. *Social and personality psychology compass*, 10(9):503–518.

Smith, E. and Conrey, F. (2007). Agent-based modeling: A new approach for theory building in social psychology. In *Personality and social psychology review*.

Squazzoni, F., Jager, W., and Edmonds, B. (2014). Social simulation in the social sciences: A brief overview. In *Social Science Computer Review*.

Starker, A., Kuhnert, R., and Hoebel, J., a. R. A. (2022). Rauchverhalten und Passivrauchbelastung Erwachsener–Ergebnisse . In *GEDA 2019/2020-EHIS.*

Tapp, L., Kurchyna, V., Nogatz, F., Berndt, J., and Timm, I. J. (2022). School's out? simulating schooling strategies during covid-19. In *Autonomous Agents and Multiagent Systems. Best and Visionary Papers: AAMAS 2022 Workshops*. Springer.

Terry, D. and Hogg, M. (1996). Group norms and the attitude-behavior relationship: A role for group identification. In *Personality and Social Psychology Bulletin*.

Tesser, A. (1980). When individual dispositions and social pressure conflict: A catastrophe. In *Human Relations*.

Wagner, N. and Agrawal, V. (2014). An agent-based simulation system for concert venue crowd evacuation modeling in the presence of a fire disaster. In *Expert Systems with Applications*.

Wens, M., Veldkamp, T. I., Mwangi, M., Johnson, J. M., Lasage, R., Haer, T., and Aerts, J. C. (2020). Simulating small-scale agricultural adaptation decisions in response to drought risk: an empirical agent-based model for semi-arid kenya. *Frontiers in water*, 2:15.

WHO (2021). Who global report on trends in prevalence of tobacco use 2000-2025. In *Geneva: World Health Organization*. Licence: CC BY-NC-SA 3.0 IGO.

Xu, X., Leung, D., Li, B., Wang, P., and Zhao, Y. (2015). Smoking-related knowledge, attitude, social pressure, and environmental constraints among new undergraduates in chongqing, china. In *International journal of environmental research and public health*.

APPENDIX

The model's NetLogo code can be viewed and executed under http://modelingcommons.org/browse/one_model/7256.