

## **Integrating Social Practice Theory in Agent-Based Models A Review of Theories and Agents**

Mercuur, Rijk; Dignum, Virginia; Jonker, Catholijn M.

**DOI**

[10.1109/TCSS.2020.3007930](https://doi.org/10.1109/TCSS.2020.3007930)

**Publication date**

2020

**Document Version**

Accepted author manuscript

**Published in**

IEEE Transactions on Computational Social Systems

**Citation (APA)**

Mercuur, R., Dignum, V., & Jonker, C. M. (2020). Integrating Social Practice Theory in Agent-Based Models: A Review of Theories and Agents. *IEEE Transactions on Computational Social Systems*, 7(5), 1131-1145. Article 9151181. <https://doi.org/10.1109/TCSS.2020.3007930>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Integrating Social Practice Theory in Agent-Based Models: A Review of Theories and Agents

Rijk Mercur, Virginia Dignum, *Member, IEEE*, and Catholijn Jonker, *Member, IEEE*

## Abstract

Evidence-driven agent-based modelling plays a useful part in understanding social phenomena. By integrating social-cognitive theories in our agent models we bear evidence from social and psychological research on our models for human decision-making. Social practice theory (SPT) provides a socio-cognitive theory that emphasizes three empirically and theoretically grounded aspects of behaviour: habituality, sociality and interconnectivity. Previous work has emphasized the importance of SPT for agents, has made abstract models of SPT or used SPT to study energy systems. This paper provides a set of requirements for integrating SPT in agent models and an evaluation of 11 current agent models with respect to these requirements. We find that current agent models do not fully capture habituality, sociality or interconnectivity, nor is there a model that aims to integrate all three aspects. For example, current models do not support context-dependent habits, do not use a comprehensive set of collective concepts and do not support hierarchies of activities. Our evaluation allows researchers to pick one of the current agent models depending on their needs regarding habituality, sociality and interconnectivity. Furthermore, this paper shows the usefulness of an agent model that integrates SPT and provides requirements that help modellers to achieve this model.

## Index Terms

Agent-based modeling, Social intelligence, Cognition

# Integrating Social Practice Theory in Agent-Based Models: A Review of Theories and Agents

## I. INTRODUCTION

**E**vidence-driven agent-based modelling plays a useful part in understanding social phenomena [1]–[5]. This includes bearing evidence on human-decision making on our models for human decision-making: agents. To utilize the evidence of sociological and psychological research, [6] argues for integrating socio-cognitive theories in our agent models. One of these socio-cognitive theories, called social practice theory (SPT), fits agent-based modeling as both study the interaction of humans with their social environment and the direct and indirect effect of these interactions on society. By grounding agent models in socio-cognitive theories, and in particular on SPT, we stand on the shoulders of years of sociological and psychological research on modelling human decision-making.

SPT provides a theory that describes our ‘everyday doings and sayings’ [7] and emphasizes that these so called social practices (SPs) are habitual, social and interconnected [7]–[10]. Our day is full of SPs: working, dining, commuting, teaching, meeting, walking or sports. First, SPs emphasize that behaviour is habitual [8]. For example, when one is at the office, one habitually enacts the SP of working. Habituality helps us to understand why it might not be so easy to fall into the same working practice at home, how you intentionally go to the office to trigger the practice of working or how you can (with willpower) also develop a habit to work at home. Second, SPs emphasize that behaviour is social [7], [9]: a practice is not only individual but others have a similar practice. For example, when your colleague enters your office, he or she does not distract you but waits until the coffee break at 10.30 to discuss current matters. Sociality helps us to understand how your colleague concludes to wait until the coffee break based on her own practice: she believes reading is a kind of work, work promotes productivity, the coffee break starts at 10.30 and that you believe the same. Third, SPs emphasize that behaviour interconnected [10]. For example, your work-commute is connected to your sport-commute and you decide to take the car so you can do both [11]. Interconnectivity helps us to understand how you want both your work-commute and sport-commute to promote efficiency and therefore take the car or how your colleague understands that reading is connected to the practice of work and therefore promotes productivity. In short, SPs describe our everyday decisions and help us model the habitual, social and interconnected aspects of these decisions.

[12], [13] call for translating socio-cognitive theories to a domain-independent agent model to prevent researchers from reinventing the wheel. They identify that current agent-based models (ABMs) use similar socio-cognitive theories, but without a general framework researchers cannot reuse,

compare or recombine these. This hampers both the efficiency [13], [14] and the evolution of models [1]. By translating SPT to a domain-independent agent model, we would enhance the comparability and reusability of agent models that model habituality, sociality and interconnectivity. However, for this purpose, we first need to identify what is required of an agent model that integrates SPs and if there is a gap in the current literature on agent models given these requirements.

This paper provides a set of requirements for an agent model that integrates SPT and verifies whether current domain-independent agent models satisfy these requirements. Previous work has emphasized the importance of SPT for agents [15], has made abstract models of SPT [16] or used SPT to study energy systems [17]. So far a review that lays out specific requirements for integrating SPT with agent models and evaluates current agent models does not yet exist. Such a review is useful for ABM researchers who are interested in integrating SPT with ABM and want to know in more detail what aspects SPT comprises, and what requirements these imply for implementations. Furthermore, it allows ABM researchers to pick one of the current agent models depending on their needs regarding habituality, sociality and interconnectivity. For this purpose, we distilled requirements from the literature on SPT, agent theory and social psychology. Each aspect (habituality, sociality and interconnectivity) is studied from the two perspectives that ABM aims to integrate: the individual agent perspective and the collective system perspective. We evaluated 11 agent models against the requirements we elicited. The selection of agent models is based on the review by [18], to which we added two more recent agent models. We split up the models in three categories: reasoning models (PRS, BDI, eBDI), normative models (BOID, BRIDGE, EMIL-A, NOA, MAIA) and social-psychological models (Consumat, PECS, Agent-0). By distilling requirements and evaluating the three categories of models, we provide an overview of the aspects SPT comprises and the current state-of-affairs in integrating these aspects in ABM.

We find that habituality, sociality and interconnectivity are empirical and theoretically grounded aspects of behaviour but have not been fully captured by current agent models. First, behaviour is habitual and often not conscious, voluntary or intentional [19]–[21]. We find that habituality is reflected in reasoning models by modelling reactivity and in the Consumat model by the ability of agents to repeat past behaviour. However, current agent models do not support (1) explicit reasoning about habits, (2) context-dependent habits and (3) individual learning concerning habits. Second, behaviour is intrinsically social and not individual with a layer of sociality on top of it [14], [22]–[24]. We find that sociality is somewhat reflected in models that use norms or that use social

mechanisms (e.g., imitation), but current agent models do not use a comprehensive set of collective concepts, nor do they order social information around actions or relate individual and collective concepts (e.g., habits to norms) in order to guide interactions. For example, a normative agent model supports reasoning about the fact that most people work, but not that because an individual agent believes work promotes productivity it reasons that most agents believe work promotes productivity. Third, behaviour is interconnected: actions do not stand alone, but are similar and influence each other [10]. We find that interconnectivity is reflected in models that use plans, but current agent models do not model explicit relations between activities, between each activity and each other model concept (e.g., desires, needs, resources, locations) nor model hierarchies of activities. In summary, although current agent models capture some aspects of SPT, none fully captures any of the individual aspects, nor is there a model that aims to integrate all three empirical and theoretically grounded aspects.

This paper shows the usefulness of a computational agent model that integrates SPT and provides requirements that help modellers to achieve this model. We do not argue that the resulting model will provide a general theory of human behaviour, but that integrating SPT is useful for agent-based modellers because it enables new insights using mechanisms that are based on evidence. Although it would certainly be useful to give an exact scope of SPT and its relevance when compared to other theories, this is rather difficult: SPT is a high-level abstract theory, social practices theorists define a SP in different ways, SPT is ever-expanding and merges with other theories and the jury is still out on the so-claimed limitations of SPT [25]. As shown by [25], SPT has given insights in a wide and diverse range of domains: eating, Nordic walking, teaching, learning, washing machine use, cycling, mobility, day trading on the Nasdaq market, domestic energy use, household waste, sustainable design, sustainable consumption, temporalities of consumption, the work of ambulance paramedics or lawyers, anxiety, memory, communities of practice, and organizational learning and knowing. An agent model that integrates SPT will enable researchers to use ABS to gain insights regarding habituality, sociality and interconnectivity in a wide range of social systems.

The remainder of the paper is structured as follows. Section II distils requirements for modelling habitual, social and interconnected behaviour from the literature on SPT, agent theory and social psychology. Section III provides an overview of current agent models and review to what extent they satisfy our requirements. Section IV discusses the consequences of the limitations of current agent models, Turing-completeness and the need for integration of these aspects versus a reductionist scientific strategy. Section V concludes the paper.<sup>1</sup>

## II. DISTILLING REQUIREMENTS FROM LITERATURE

Habituality, sociality and interconnectivity express that SPs have similar properties in three dimensions: over time, over people and over different activities (see Figure 1). Habitually expresses that SP are similar over time. For example,

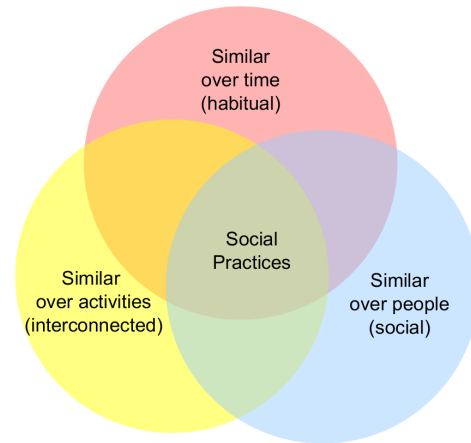


Fig. 1. A Venn diagram representing how habitual, social and interconnected actions come together in social practices.

commuting is habitual, because a person commutes by car everyday. Sociality express that behaviour is similar over people. For example, commuting is social, because most people associate commuting with a car. Interconnectivity expresses that behaviour is similar for different activities. For example, commuting is interconnected, because most people associate both shopping and commuting with a car. SPs thus truly capture our everyday doings and saying: behaviour that is expected to follow a predictable pattern as its similar with respect to time, people or other activities.

To connect SPT to ABM we need to connect the collective view of SPT and the individual view of agents. [26] sees SPT as a way to abstract away from the individual. They see SPs as a collective entity that recruits or loses host (i.e. agents) over time. In contrast, [27] views SP as a way to connect agency and social structures. In the same line, [15] brings agency back on the table and connects collective concepts related to SPs (e.g., values, norms) with individual agent concepts (e.g., goals, beliefs). For ABM it is in particular important to connect the individual and collective view, because ABM uses agents as a primary concept and aims to connect the micro (individual) with the macro (collective).

To extract requirements for integrating SPT and agent models, the following subsections discuss in more detail how SPs relate to habits, sociality and interconnectivity. Each subsection connects the collective view of SPs to the individual view of agents. We end each subsection with a number of requirements for integrating SPs in ABM.

### A. Social Practices and Habits

SPs and habits both describe behaviour that is similar over time. SPT studies repetitive behaviour on a collective level and is interested in what aspects of a SP exactly repeat [8], [10]. Social psychologists study repetitive behaviour from the individual perspective as ‘habits’. They use the term ‘habit’ to refer to a phenomenon whereby behaviour persists because it has become an automatic response to a particular, regularly encountered, context [28].

We differentiate between two views on habits: habits as a behavioural dynamic notion and habits as a cognitive static notion. The switch from the behavioural view to the cognitive view entails that habits are not merely observable behaviour, but also mental phenomena. For example, one can refer to the habitual behaviour car-driving or the habitual cognitive connection between commuting and car-driving. The static view entails that habits are not only the repeated behaviour over time but also a mental configuration that persists in the mind (irrespective of the times when the behaviour is actually carried out). For example, one can express a habit dynamically as ‘to use the car everyday’ or statically as ‘at this moment there is a strong mental connection between the car and commuting’. An agent model that integrates SPT thus needs not only a representation of the dynamic decision, update and reasoning algorithms, but it also needs to provide the static configuration of objects, variables and relations. Thus, the model should provide researchers with the primary concepts and relations to model agents that make habitual decisions, updates and reason about habituality.

Models that aim to express habitual decisions and updates need to contrast these with intentional decisions and updates [19], [29], [30]. We follow Wood and Moors by recognizing that the automaticity of habits entails unintentional, uncontrollable, goal independent, autonomous, purely stimulus-driven, unconscious, efficient, and fast behaviour [20], [29]. The automaticity of habits gains meaning when contrasted with another decision mode: intentional decisions. Furthermore, habits and intentions interact and habitual decisions and updates are a product of this interaction [29]. For example, a car-driving habit emerges when agents intentionally drive cars over a long enough period of time. To model the automaticity of habits and interaction with non-habitual behaviour, the model should enable agents to differentiate between habits and intentions.

Habits are sensitive to contextual triggers. For example, the context ‘home’ can trigger the habit of taking the car (whereas the context ‘hotel’ might not). To be more precise: habitual decisions are triggered by particular context-elements (that together comprise the whole context) [29]. For example, the context ‘home’ consists of several context-elements, such as, ‘your house’, ‘breakfast time’, ‘kids at home’, that together trigger your car-driving habit. The strength of these ‘context element’-action relations is a continuous instead of a discrete parameter (e.g., a coffee machine is a slightly stronger habitual trigger than a colleague) [20]. Thus, to form the basis of habitual decisions, the model should enable habitual relations between an action and a context-element where the strength of that relationship is a continuous parameter.

The literature on habits differs in what they consider as context-elements. The common factor in these definitions is that context-elements are physical tangible resources or locations [29], [31], [32]. For example, nearby cigarettes can trigger a habit of smoking. Wider definitions of context-elements allow timepoints, other activities [31] and/or other people to trigger habits [29]. We choose the wider definition because it matches how habituality is described in SPT. [9] emphasized that habituality in SP does not only refer to physical resources but to mental associations as well. Thus

the model should capture that context-elements can comprise resources, activities, locations, timepoints or other people.

Habits depend on their actors. First, the strength of the habitual connection between context and actions is agent-specific. Bob’s habit to take the train is not the same as Alice’s habit to take the train. Second, how the strength changes over time differs per human. [32] empirically studied this strength gain in an experiment where subjects were asked to do the same action daily in the same context and report on automaticity. The subjects reported an increase in habit strength that followed a different asymptotic curve per subject and converged at a different maximum habit strength per subject. The model thus needs to enable agents to differ in the strength of the habitual connection, the maximum of this strength and the function over time to reach this maximum.

A habit is sensitive to the attention attributed to a decision [33]. The more attention attributed to the decision the lower the chance the action is done out of habit. The literature on the regulation of attention is extensive (e.g., see [34], [35] for an overview) and it goes too far to capture this concept in detail in this paper. Enough to say here is that the model should capture that agents can vary in their attention at different moments in time.

Intentional actions contrast with habitual behaviour as intentional actions are attempts to achieve some abstract aim [36]. Examples of concepts that capture this abstract aim are goals, desires, values, motives [36], [37]. The model should provide a concept that captures the abstract aim intentions are directed at.

The following requirements summarize this section:

- H.1** The model should capture the similarity of behaviour over time.
- H.2** The model should provide researchers with the primary concepts and relations to model agents that make habitual decisions, updates and reason about habituality.
- H.3** The model should provide researchers with the primary concepts and relations to model agents that differentiate between habits and intentions.
- H.4** The model should support habitual relations between an action and a context-element where the strength of that relationship is a continuous parameter.
- H.5** The model should capture that context-elements can comprise resources, activities, location, timepoints or other people.
- H.6** The model should capture that agents can differ in
  - a) the strength of a habitual connection
  - b) the maximum strength of a habitual connection
  - c) the time to reach this maximum
  - d) the amount of attention they attribute to a decision
- H.7** The model should provide a concept that captures the abstract aim intentions are directed at.

## B. Social Practices and Sociality

SPs and sociality are connected because both focus on the similarity of behaviour over people. SPT focuses on sociality primarily as a static group notion emphasizing that we have a similar view on the world that can be organized in terms of

our SPs. Social intelligence focuses on sociality as a dynamic individual notion emphasizing our ability to act wisely in interactions. This paper uses the literature on SPs and social intelligence to identify what is required to model that SPs are social.

There is a variety of definitions on what it means for an agent to be social or socially intelligent. For Thorndike it's the ability to act wisely in interactions [22]. This is close to the layman idea of sociality: an activity that is done in the presence of other people. For Goleman, it means that agents have social awareness and social influence [38]. For [15] it's the ability to form expectations about the behaviour of others and react to them. The commonality in these views is that there is some information to be had about other people and that this information is used to guide (social) decisions and (social) updates. The model should provide primary concepts and relations that capture this social information and enable agents to make socially intelligent decisions, update social information and reason about collective concepts.

In the agent literature, there has been an evolution about which primary concepts should be used to do this social decision-making, updating and reasoning. A first series of papers uses agents that only take into consideration the actions of others. For example, in the Consumat model [39] an agent takes into consideration what most other agents do. [40] emphasized that we need to extend such models to also consider the mental state of other agents. He claims the notion of social action cannot be a behavioural notion - just based on an external description, because what makes the action social is that it is based on certain mental states. A second series of papers focuses on such a representation of the mind of other agents based on individual notions such as beliefs, desires and intentions [41], [42]. Sociality is introduced as a secondary notion. For example, as the ability to form beliefs about other's goals [42] or as a mechanism to filter its intention to a socially desired set [41]. [14], [37], [43] argued that humans are at the core social beings and thus use social concepts as a primary concept. [44] agrees and sees these social concepts as the "new micro-foundations" for agent decision-making. We view social concepts here as referring to reasoning concepts that depend on being collective (and henceforth call them 'collective concepts'). For example, the notion of culture is hard to imagine without multiple agents. A third series of papers focuses on these collective concepts. For example, values [45], [46], norms [47], trust [48] and culture [49]. What characterizes this work is to use a collective notion in the individual reasoning of an agent. For [50], it's exactly this ability to reason about collective concepts (e.g., culture, or a political party) as individuals that makes us unique as humans. Furthermore, the use of collective notion in individual reasoning provides ABM with a way to connect the micro and macro. We require that an agent model of SPs uses a comprehensive set of collective concepts that support social decision-making, updating and reasoning.

In SPT, an SP is seen as social exactly because it is a concept that depends on being collective: they are the primary concepts that we use to order the world around us. We order our day by a series of practices (breakfast, commuting,

working, lunching, sports, showering, sleeping) and we have a similar view on these practices (e.g., to commute with the car one needs a car, believe car driving enables commuting and value going from A to B). [9] emphasized that SPT entails a paradigm shift: the social is captured in our collective SPs instead of in a collective mental world (i.e., mentalism) or collective texts (i.e., textualism). For [10] and [9] an SP is thus social in the sense that it stores social, collective and similar information and not necessarily in the sense that it's interactive. For example, one of [10]'s canonical examples of an SP is showering, an SP that is mostly done alone. Furthermore, SPs capture a particular view on the social world where the social world is ordered around our daily doings and sayings. We require that an agent model of SPs captures that SPs relate the collective social world with the individual world of interactions.

Whereas SPT makes a point of using SP to take a collective view on behaviour that does not concern individual interactions, we see SPs as a primary concept that individuals use to guide interactions. For an SP-theorist such as [9], SPs are thus something collective, "but not in the sense of a mere sum of the content of single minds, but in a time-space transcending non-subjective way". Although SPs thus emerge from individual enactments, [9] views it as a completely separate collective entity. [10] argued that what makes SPT valuable is, in particular, this shift from the individual view to the collective view. Individuals are merely carriers or hosts that are used by the SP to spread around. [10] sees SPT as a way to break with the view that behaviour is the result of individual decision-making. In contrast, for an agent theorist such as [15], SPs are an entity that (also) exist on the level of individual decision-making. As mentioned before, [15], [17] see SP as useful for individual interactions just because it exists on both levels: the individual and the collective level. For agent-based modelling, in particular, it's important to follow this second line of reasoning and include both the individual (micro) and collective (macro) view in our understanding of the social world. SPs are one way to model social information around a structure that exists both the individual and collective level: practices. We require that an agent model of SPs supports agents that order social information around their practices.

By looking at SP from both an individual and collective viewpoint, we notice that these two views do not always match. In other words, there are multiple views possible on the same SP. For example, one can view car-driving as an action that promotes (the value of) pleasure or that demotes pleasure. Moreover, humans differ between what they believe an SP comprises and what they believe others believe an SP comprises. From the viewpoint of the individual, one can differ between a personal view and a collective view on an SP. For example, one believes that car-driving is usually seen as a means for transport but believes him or herself it's a fun activity. These personal and collective aspects of a view can differ: one can believe in a personal view on something without believing this view is collective or one can believe something is the collective view without believing in this view. We require that agents have beliefs about both their individual view on an SP as well as a collective view on an SP.

There is a difference between the collective view on an SP from the viewpoint of the modeller and the viewpoint of the agent. The modeller can see what beliefs are truly collective among all agents: the modeller can extract the collective SP from a model. For example, a modeller might extract that there are individually different beliefs about the relation between car-driving and pleasure, but that all individuals believe that to car-drive one needs a car. In contrast, the agents themselves have to guess what others believe and these guesses differ. For example, one agent believes that most others see car-driving as pleasurable while other agents believe there are individual differences. There is a reciprocal dynamic between the beliefs that are truly collective among all agents and the collective view of each individual agent.<sup>2</sup> A modeller is able to extract the true collective view from the fine-grained collective views of agents, but not vice-versa. Therefore we require that an agent model of SPs supports a collective view that can differ from agent to agent.

This view also makes clear that SPT only considers a subset of social intelligence. SPT takes a general collective view on our daily activities; SPs are a heuristic where humans generalize over a group of people. This contrasts with another strand of social intelligence called the theory of mind (ToM). Studies on ToM study human's ability to create a mental model of others' beliefs. [51]. In contrast with SPT, studies on ToM consider beliefs about specific others and chains of beliefs. For example, one can believe John believes car-driving is fun or believe that John believes that I believe that John believes car driving is fun. These aspects are out of the scope of SPT. SPs are a heuristic that considers only two agents: itself or the group. For example, when greeting someone in most cases it suffices to know that most people view greeting as polite and see shaking hands as a part of greeting. SP focuses on the social intelligence that works in most situations, in contrast, research on the theory of mind treats particular cases where more in-depth reasoning is needed. We thus require that an agent model of SPs supports both a personal and collective view on SP (but not necessarily beliefs about particular others or chains of beliefs).

The following requirements summarize this section:

- SI.1** The model should capture the similarity of behaviour over people.
- SI.2** The model should provide researchers with the primary concepts and relations to model agents that make socially intelligent decisions, updates and reason about collective concepts.
- SI.3** The model should use a comprehensive set of collective concepts that support social decision-making, updating and reasoning.
- SI.4** The model should enable agents to order social information around their practices.
- SI.5** The model should capture that social practices relate the collective social world with the individual world of interactions.
- SI.6** The model should capture that agents have a personal view on a social practice.
- SI.7** The model should capture that agents have a collective view on a social practice.

**SI.8** The model should enable agents to have a different personal view than their collective view.

**SI.9** The model should enable agents to each have a different collective view on a social practice.

### C. Social Practices and Interconnectivity

SPs and interconnectivity are connected because they both focus on the similarity of behaviour over different activities. Activities here refer to bodily movements. SPT focusses on interconnectivity on an abstract level of activity. For example, it discusses how the work-commute and school-commute are connected because they both use the car as a resource. The agent literature focuses on interconnectivity on a concrete level of activity. For example, it discusses how commuting comprises getting the car keys, driving the car and arriving at work. SPs and agent activities thus exist at different levels of abstraction, but both comprise bodily movements. As we will discuss in the next section, we view agent activities as part of SPs. This section uses the literature on agents and SPs to identify what is required to model that SPs are interconnected.

SPT argues that if SPs are connected in some aspects, then they become more connected in other aspects too [10]. For example, [10] mentions how in the early days of driving, cars easily broke down. To be able to drive a car one needed the competence to repair it. The SP of driving thus became connected with other SPs that related to the competence of repairing, for instance, plumbing or carpeting. The meaning of these SPs as something masculine influenced the meaning of car driving. Car driving and plumbing now share a masculine meaning. The model should provide researchers with the primary concepts and relations to model that if SPs are connected, they become more connected.

In the agent literature, activities are interconnected to enable agents to make decisions and inferences. [52] connects activities via goals in the agent GOAL language. For example, the goal of a successful day can be split up in you being in the car, you being at work and you being home. [52] views goals as states of the world and activities as ways to reach the state. In contrast, research on language protocols [53] uses activities as their primary concepts and specifies relations between them. A common factor in this work is that its necessary to define the relation between activities to enable agents to decide what to do next and reason about the properties of activities. Recent work in SPT reflects this view. For example, [11] found that interviewees decide to take the car, because they aim to connect leisure activities, healthcare activities and shopping activities. The model should provide the primary concepts and relations to enable agents to make interconnected decisions, updates and reason about the interconnectedness of activities,

In SPT, SPs are interconnected in terms of time, space or common elements [10]. First, SPs connect when they are enacted at the same time or in sequence. For example, the SP of breakfast and commuting are interconnected because they happen around the same time. Second, SPs connect when they are enacted in the same space. For example, the SP of working and getting coffee are connected because they are happening in the same place. Third, SPs are connected when they share an

element. For example, in the early mentioned example of [10] plumbing and car-driving are connected via the competence of repairing and the meaning of masculinity. The model should express that SPs are connected in terms of time, space and common elements.

From the agent perspective, [54]–[57] classified activities and studied in which possible ways activities relate. The central aim of this work is to recognize activities in the context of smart homes. For example, to recognize that a person is cooking, because he or she boils water and is looking for a cutting board. [57] makes a difference between actions and sequential activities.<sup>3</sup> Actions are atomic. Sequential activities are an ordered sequence of actions. For example, commuting is a sequence of taking the kids to school and going to work. Note that from this point on we will separate between activities and actions. Activities refer to any bodily movement (i.e., actions and sequential activities). Actions refer to the subset of activities that are atomic. The model should differentiate between different types of activities: atomic actions and sequential activities (an ordered sequence of actions).

[57] separate two types of relations between activities: an ontological and temporal relation.<sup>4</sup> The ontological part describes relations between actions such as subsumptions, equivalence or disjointness. For example, taking the train to school is a kind of commuting. A temporal relation encodes qualitative information regarding time. For example, the user performs two activities after another. This ontological and temporal information can be used by agents to decide what to do next or to make inferences. For example, humans infer that if taking the car to work is environmentally unfriendly then taking the car to school might be as well. The model should capture these temporal and ontological relations between activities to enable agents to make decisions and inferences.

The following requirements summarize this section:

- I.1** The model should capture the similarity of behaviour over different activities.
- I.2** The model should provide researchers with the primary concepts and relations to model agents that make interconnected decisions, updates and reason about the interconnectedness of activities.
- I.3** The model should express that social practices are connected in terms of time, space and common elements.
- I.4** The model should differentiate between different types of activities: atomic actions and sequential activities (an ordered sequence of actions).
- I.5** The model should capture both temporal and ontological relations between activities to enable agents to make decisions and inferences.

### III. EVALUATION OF CURRENT AGENT MODELS

We evaluate for 11 domain-independent agent-based models to what extent they satisfy the requirements for integrating SPT in agent models. These models are not designed to satisfy our requirements, they have their own purpose. However, the comparison makes clear that if one wants to integrate SPs in agent-based models current models do not suffice. To select related models, we use the overview by Balke and Gilbert

[18], but omit neuro-cognitive models because they study the neurology of the mind as viewed from the outside, while we are interested in a socio-cognitive view on the mind as we experience it from the inside. We add two relevant domain-independent agent-based models published after the review of [18]: MAIA [58] and Agent-0 [59].<sup>5</sup> MAIA is relevant as it aims to integrate sociality with ABM and adds a new concept of roles. Agent-0 is relevant as it aims to integrate three decision-making modules that include a context-dependent module and a social module. Using this method we come to a list of the following 11 domain-independent agent models: PRS [60], BDI [61], eBDI [62], BOID [41], BRIDGE [63], EMIL-A [64], NOA [65], MAIA [58], Consumat [39], [66], PECS [67] and Agent-0 [59].

We divide these agent-based models into three categories: reasoning models, normative models and social-psychological models. Reasoning models first emphasized autonomy, reactivity (e.g., PRS) and later added proactivity (e.g., BDI, eBDI). When it became clear that adding agent-communication language to such models does not suffice to successfully represent sociality in humans, researchers focused on adding norms to agent models [68], [69]. Normative agent models focus on different types of norms: social norms (e.g., EMIL-A), deontological norms (e.g., BOID) or both (e.g., MAIA); and different dynamics within norms: norm innovation (e.g., EMIL-A) or norm enforcement (e.g., BOID). Last, some researchers took their inspiration more directly from social-psychological literature and combined several socio-psychological mechanisms in one model (i.e., Consumat, PECS, Agent-0). Reasoning models, normative models, social-psychological models thus represent three categories in ABM that relate in a similar way to our requirements.

The remainder of this section compares the reasoning, normative, and social psychological models to our requirements on habituality, sociality and interconnectivity. If there are differences between the models within the category, then we follow the charitable principle; the comment in the table refers to the model(s) that relate(s) most closely to our requirement and the(se) model(s) are stated.

#### A. Habits

Table I shows that current agent models do not support (1) explicit reasoning about habits, (2) context-dependent habits and (3) individual learning concerning habits. In more detail:

*H1-2:* Current agent models do not support explicit reasoning about habits. Reasoning agents and normative agents do not make explicit habitual decision and updates or reason about habituality. Norms differ from habits in that they refer to similarity over people (e.g., most people usually drive a car), whereas habits refer to similarity over time for one individual (e.g., I usually drive a car). Consumat models habitual decisions by giving the agent a chance to repeat past behaviour [39]. However, there is no explicit variable capturing the properties of habits (e.g., the strength of the habit). Therefore, the Consumat agent makes habitual decisions but is not able to reason about habits. In summary, some agent models make habitual decisions (Consumat), but none reason about habits.



TABLE I  
HABITUALITY: VERIFICATION AND EXPLANATION OF DIFFERENT AGENT MODELS WITH RESPECT TO OUR REQUIREMENTS. THE ‘?’-COLUMN DENOTES WHETHER THE REQUIREMENT IS UNSATISFIED (X), SOMEWHAT SATISFIED (~) OR SATISFIED (✓) BY THE AGENT MODEL.

Nr.	Requirements	Reasoning Models (PRS, BDI, eBDI)		Normative Models (BOID, BRIDGE, EMIL-A, NOA, MAIA)		Social-Psychological Models (Consumat, PECS, Agent-0)	
		?	Explanation	?	Explanation	?	Explanation
H1-2	similarity over time captured in habits	X	no explicit habitual reasoning	X	norms do not refer to individual repetition	~	habitual decisions, but no reasoning about habits
H3	habits and intentions	~	no, but reactivity and deliberation	~	no, but reactivity and deliberation	✓	habits and intentions (Consumat)
H4	action-‘context-element’ relation	~	no difference between pre-condition and trigger	~	no difference between pre-condition and trigger	~	only in one module (Agent-0)
H5	a context-element comprises	X	activities, time and agents are not context-elements	X	activities, time and agents are not context-elements	X	activities, time and agents are not context-elements
H6	agents differ in their habits	X	reactivity is not adaptive nor agent-specific	X	reactivity is not adaptive nor agent-specific	~	only due to experience
H7	intentions are directed at an abstract aim	✓	intentions are a primary concept	✓	captured by intentions, utility, goals or normative goals	✓	intentions are captured by utility-maximization

*H3:* Only the Consumat agent has both a habitual and intentional decision mode. However, habitual decision-making in the Consumat is not context-dependent (H4-H5) or agent-specific (H6). Reasoning models and normative models aim for reactivity instead of habituality. Reactivity matches habituality in that it requires agent models to react to the environment (i.e., context) in a timely fashion [70]. Reactivity (in current implementations) differs from habituality in that it confounds pre-conditions and triggers (H4), does not consider agents and activities to be context-elements (H5) and is not adaptive or agent-specific (H6).<sup>6</sup> All agent models have an intentional mode of decision-making (sometimes called deliberate decision-making or rational decision-making) (H7). In summary, current agent models all have an intentional mode of decision-making, but only the Consumat also has a habitual decision mode (which has several limitations we will now expand on).

*H4:* Some agent models conceptualize context-action relations, but they confound pre-conditions and triggers. Reasoning models and normative models confound pre-conditions and habitual triggers [15]. Pre-conditions (or as [15] calls them: affordances) relate context-elements to when an action is possible, whereas habitual triggers relate context-elements to priorities over actions. Confounding pre-conditions and triggers leads to unintentional prioritization in reasoning and normative models [15].<sup>7</sup> In Agent-0 one of the three decision-making modules (the affective one) uses context to determine the appropriate fear reaction. However, this differs from habits where the conditioning happens directly between the context and action. Although the Consumat agent has a habitual decision-making mode, the habit is not sensitive to the context. For example, it is not relevant if the agent is at home or in the office to repeat past behaviour (instead repetition depends on the current satisfaction of the Consumat agent). In summary, reasoning models and normative confound pre-conditions and triggers and only one social-psychological model (Agent-0) supports context-action relations, but only in one module.

*H5:* Current agent models do not consider other activities, timepoints or agents to be context-elements. Although we found no formal definitions as specifications are example-

based, instances of context-element in the models refer only to resources or locations.

*H6:* Current agent models do not model adaptive and agent-specific reactions to the context. Reasoning and normative models hard-code the reaction to the environment (at all times and for all agents) in the form of plans. Agents thus do not learn an agent-specific reaction to the environment over time. In the Consumat model, the experience of an agent influences its propensity to repeat past behaviour. However, there are no agent-specific parameters that specify an agent’s personal tendency to go into a habit or develop stronger habits over time. In summary, in some agent models (Consumat) habits depend on experience, but in none of the agent models, the agents have a personal tendency to go into habits or develop stronger habits over time.

*H7:* Current agent models all have intentions that are direct at an abstract aim. In reasoning and most normative models intention is a primary concept. In normative models, intentions are captured by utility (MAIA), goals (EMIL-A) or normative goals (EMIL-A). In social-psychological models, intentions are captured by utility-maximization. Current agent models thus capture intentions either as a primary concept or reframe it as goals or utility-maximization.

## B. Sociality

Table II shows that current models do not use a comprehensive set of collective concepts, order information around actions and relate individual and collective concepts in order to guide interactions. In more detail:

*S1:* Only normative models have an explicit concept that denotes a similarity over people. Norms denote the similarity of actions over people (e.g., most people drive a car). However, normative models do not conceptualize the similarity of people concerning other mental constructs in the model. For example, no concept expresses that most people have a certain goal or that most people relate a certain action to a certain goal. In reasoning models, agents enact the same actions, goals or plans, but there is no explicit concept capturing this similarity. Models for the theory of mind and mental models enhance the architecture of BDI agents (see e.g., [42], [71]). Although

TABLE II

SOCIALITY: VERIFICATION AND EXPLANATION OF DIFFERENT AGENT MODELS WITH RESPECT TO OUR REQUIREMENTS. THE ‘?’-COLUMN DENOTES WHETHER THE REQUIREMENT IS UNSATISFIED (✗), SOMEWHAT SATISFIED (∼) OR SATISFIED (✓) BY THE AGENT MODEL.

Requirements		Reasoning Models (PRS, BDI, eBDI)		Normative Models (BOID, BRIDGE, EMIL-A, NOA, MAIA)		Social-Psychological Models (Consumat, PECS, Agent-0)	
Nr.	Shorthand	?	Explanation	?	Explanation	?	Explanation
S1	conceptualizes similarity over people	✗	only implicit in similar actions, goals or plans	✓	norms explicitly capture similarity wrt actions	✗	only implicit in similar actions or motivations
S2	social decisions, updates and reasoning	∼	no explicit sociality	∼	explicit decision, updates and reasoning about norms	∼	imitation of other agents, but no social reasoning
S3	use collective concepts	✗	only emotions (eBDI)	∼	only roles (MAIA) and norms	✗	only emotions (Agent-0) and implicit norms
S4	order sociality around practices	✗	mental models are not ordered around actions	✗	mental models are not ordered around actions	✗	mental models are not ordered around actions
S5a	relate collective and individual concepts	✗	limited individual/collective concepts and relation	∼	only relates personal actions and collective actions	✗	limited individual/collective concepts and relation
S5b	use collective concepts to guide interactions	✗	no, because of limitations in S5a	∼	does not adequately relate norms to habits in interactions	✗	no, because of limitations in S5a
S6-9	different personal and collective view	✗	no collective view	✓	agents each track a local norm wrt actions	✓	agents each track a local norm wrt actions

these approaches model some aspects of the collective social world they are not systematically build-up from collective concepts such as values, culture, norms or social practices. This comment is in line with [37]. Likewise, in social psychological models, agents enact similar actions or have similar motivational concepts (e.g., needs in Consumat or fears in Agent-0), but they do not have explicit concepts capturing similarities and dissimilarities and these agents do not reason about such similarities. In summary, reasoning and social-psychological models have no explicit concept that denotes the similarity of agents; normative models capture the similarity of people with respect to actions as a characteristic of the concept norm.

S2: There are aspects of social intelligence that fall outside the scope of this paper, therefore we withhold from a complete evaluation on the ability of models to do social decision, updates and reasoning. Recall that this paper looks at the intersection of social practices and social intelligence (section II-B) and aspects like ToM or social impact are not required of a social practice model. Having said that, different agent models emphasize sociality to a different extent. This results in differences in how explicit these models incorporate social decisions, updates and reasoning: the reasoning models do not emphasize explicit sociality, the normative models emphasize sociality wrt norms and the social-psychological models emphasize social mechanisms more than explicit collective concepts and reasoning. The details of these differences are covered in the following sub-sub sections that treat S3-S9. In summary, we state the differences concerning S2 that are relevant for this paper in the next sub-sub sections but withhold from a complete evaluation of the models concerning S2.

S3: Current agent models do not use a comprehensive set of collective concepts. For example, there is no concept of values, culture or identity. Reasoning models only use individual concepts: beliefs, desires and intentions and eBDI adds one collective concept to this list: emotions. Normative models focus on the concept of norms and MAIA uses the concept of role to denote the subset of agents that are similar in their

goals or norms. Social-psychological models use emotions (Agent-0) and social mechanisms (imitation), but no explicit collective concepts. In summary, the current agent models use some collective concepts (norms, emotions, roles), but omit others (values, culture and identity).

S4: Current agent models do not order sociality around practices. Reasoning models, normative models and social-psychological models do not conceptualize practices but do have a concept of action. Because practices are a series of actions, we instead inquire: do current agent models order their information around these actions? More precisely, do current agent models conceptualize the relation between each action and each other concept within the model and use these relations to understand each other? We find that they do not. In reasoning and normative models, actions, desires and intentions are linked via plans. This does not specify for every action if the action promotes a desire or not. In social-psychological models, agents conceptualize the relation between actions and mental concepts (e.g., needs) and physical concepts (i.e., in the fear-module of Agent-0), but not other actions. Because actions (and their relation with other concepts within the model) do not take center stage, they are not used to form mental models of other agents (i.e., order social information). In summary, as agents do not use actions and practices as a central component of their model these cannot be used to order social information.

S5: Through comparison with current models, we found that requirement SI.5 needs to be evaluated on two aspects. Requirement SI.5 states that the model should capture that social practices relate the collective social world with the individual world of interaction. This encompasses two aspects. First, sociality requires that the agent model connects collective concepts to individual concepts (SI.5a). Second, sociality requires that agents use collective concepts to guide interactions (SI.5b). We now continue to evaluate current agent models on both these aspects.

S5a: Current models do not relate individual concepts to collective concepts, because they do not comprise both the individual and collective concepts or because the ontology

does not relate them adequately. First, all of the evaluated models lack collective concepts. In particular, values, culture or identity. Second, all models lack individual concepts. In particular, habits (except, as discussed, the Consumat model). Habits form the individual counterpart to norms. Where habits state what an individual mostly does, given a certain situation, norms state what the collective mostly does, given a certain situation. None of the models (including normative models) relate habits explicitly to norms. Third, for some models, the individual and collective concept is included but not adequately connected. We give two examples (in these examples the collective and individual concept have the same name):

- all models do have a concept of the physical world, which is both an individual and collective concept, but do not reason about the fact that the physical world is collective. The models do not reason about the fact that others, given the current physical context, will do the same action as they do, or have the same desire (BDI), same emotion (agent-0) or same need (Consumat).<sup>8</sup>
- none of the models reasons about the fact that the motivational constructs they use are collective. The models do not reason about the fact that most other agents have the same desires (reasoning and normative models), emotions or needs as themselves.

In summary, current models do not relate individual concepts to collective concepts, because they do not comprise certain individual concepts (i.e., contextualized habits) and collective concepts (i.e., values, culture or identity) or because the ontology does not relate them adequately (i.e., physical context, desires, emotions, needs).

*S5b*: Current agent models are limited in guiding interactions because they do not relate individual and collective concepts. The different cases of limited individual and collective connection (*S5a*) have a direct consequence for guiding interactions. First, if an agent model does not comprise a collective concept (e.g., culture), the agent cannot form expectations of other agents that are based on that concept (e.g., the other agent does not shake hands because that's not part of its culture). Second, if a model does not comprise an individual concept (e.g., habits), the agent cannot reason about the collectiveness of this individual concept (e.g., most other people will also have the habit to drive to work, so I can ask someone to carpool with). Third, if a model does not make an adequate connection between an individual and collective concept (e.g., the individual and collective view on motivation), the agent cannot use its mental model to form expectations about others (e.g., most other people will also have a desire to be healthy).<sup>9</sup> In summary, in current agent models agents are limited in guiding interactions, because they cannot form expectations regarding collective concepts they do not conceptualize or use their mental model as a proxy for others.

*S6-9*: Normative and social-psychological models agents have a different personal and collective view on actions. In these models, agent each have a different view on the best action based on their intentions (i.e., intentions or utility-maximization). Besides, they have a different view on what the collective views as the best action. In BOID and BRIDGE,

this collective view is expressed in the different obligations that hold for different agents. In MAIA [58], this collective view is expressed in roles, where a different role for an agent means a different conceptualization of the norm. In Consumat and Agent-0, this collective view expressed in the form of a local norm: agents each imitate the agents around them and thus have a different view on what the collective best action holds. Because none of the models connects other individual and collective concepts, the agents do not have a different personal and collective view on these concepts.<sup>10</sup> In summary, in normative and social-psychological models each agent has a different personal and collective view on actions, but not on other concepts.

### C. Interconnectivity

Table III shows that current models do not make explicit relations between activities, between each activities and each other model concept (e.g., desires, context) nor model hierarchies of activities. Reasoning and some normative models (BOID, BRIDGE) do have plans that implicitly connect activities and context-elements and motivations. Social-psychological models connect activities to the same motivations. However, without explicit hierarchies and connections between activities the models do not directly support inferences about the similarity of activities on a personal level (e.g., two activities need the same resource, need to be performed in sequence, promote the same value) or on a social level (e.g., other people will need this resource). In more detail:

*I1*: Current models do not model an explicit similarity relation between each activity. Reasoning and normative models use plans to relate activities to other activities. The relations between activities and the similarities in activities are implicit in these plans. For example, two activities lead to the same goal, because the two activities lead to two different subgoals that eventually lead to the same goal (see [72] for BDI-based planning). Social psychological models model relate actions with motivational constructs in the model. For example, a certain action satisfies a certain need (Consumat), fear or rational intention (Agent-0). In summary, reasoning and normative models specify a relation between activities via plans, all models specify relations between activities and some other elements, but there are no explicit relations between activities that denote the similarity.

*I2*: Current models focus on deciding what's next, some models (extensions of BDI) plan ahead, but none reason about the interconnectivity of activities. Reasoning and normative models use plans to make decisions and reason about the interconnection of activities. BDI-based models in particular reason about what is the next action. [73], [74] extend such models by planning ahead: they approach the interconnection of actions as a coordination problem where agents search optimal sequences of activities that satisfy a set of time constraints. This is useful for ABS that zoom in on a small time-scale (where precise coordination of actions with other agents has a big influence on the overall system). However, ABS aims to models human limitations in sequencing actions: action sequences in humans are the sub-optimal product habits

TABLE III

INTERCONNECTIVITY: VERIFICATION AND EXPLANATION OF DIFFERENT AGENT MODELS WITH RESPECT TO OUR REQUIREMENTS. THE ‘?’-COLUMN DENOTES WHETHER THE REQUIREMENT IS UNSATISFIED (✗), SOMEWHAT SATISFIED (∼) OR SATISFIED (✓) BY THE AGENT MODEL.

Requirements		Reasoning Models (PRS, BDI, eBDI)		Normative Models (BOID, BRIDGE, EMIL-A, NOA, MAIA)		Social-Psychological Models (Consumat, PECS, Agent-0)	
Nr.	Shorthand	?	Explanation	?	Explanation	?	Explanation
11	similar over activities	✗	only implicit similarity in plans	✗	only implicit similarity in plans (BOID, BRIDGE) and norms	✗	only implicit similarity in needs (Consumat) or emotions and rationality (Agent-0)
12	interconnected decisions, updates and reasoning	∼	use plans to decide what next and plan ahead	∼	use plans (BOID, BRIDGE) and norms to decide what is next, but no reasoning	✗	no planning or reasoning
13	SPs are connected in time, space and common elements	∼	only implicit connection through plans	∼	only implicit connection through plans and norms	∼	connected to needs/emotions, but limited spatial and no temporal connections
14	different types of activities	✗	no separation between different levels of abstractness in activities	✗	no separation between different levels of abstractness in activities	✗	no separation between different levels of abstractness in activities
15	temporal and ontological connections	∼	implicit connection through plans	∼	implicit connection through plans	✗	no temporal and ontological relations

and, coordination between humans is based on the limited beliefs agents have about others. Reasoning and normative models do not emphasize making inferences using the interconnection of activities to further resource management or social expectations. For example, agents do not reason that ‘car commuting to school’ relates to ‘car commuting’ and therefore cannot infer that the resource ‘car’ relates to the abstract activity of ‘commuting’. As a consequence, the agent does not immediately know that it will need a car to commute nor is the agent able to form expectations about others needing a car to commute (also see Section IV about the interplay of sociality and interconnectivity). Social-psychological models focus on deciding what is next, but do not plan ahead nor reason about the interconnectivity of models. In summary, all models focus on deciding what’s next, some extensions of BDI models focus on optimal plans, but none emphasize reasoning about the interconnectivity of activities.

*13:* Current reasoning and normative model link temporal, spatial and elemental through plans. Social-psychological models connect activities directly to motivational constructs. Agent-0 makes a spatial connection between activities in the fear-module by making the fear an agent experiences context-dependent. All other explicit temporal and/or spatial relationships have to be specified by the designer of the system. However, there are no further spatial and no temporal connections in the model. In summary, in reasoning and normative models the connection of activities (temporal, spatial, to other elements) is implicit in plans; in social-psychological models, actions are connected directly to motivational constructs, but there are no temporal and limited spatial (only Agent-0 and only in one module) connections.

*14:* Current agent models do not have different types of activities. BDI-based models (eBDI, BOID, BRIDGE) are centered around states of the world: goals/desires are states of the world and agents have beliefs about these states being currently true or not true. For example, an agent reasons that it wants to go from state  $on(block-a, floor)$  to  $on(block-a, block-b)$  and therefore first moves block-

b on the table and then block-a on block-b, but does not have explicit knowledge about the sequence of activities it should take. Social-psychological models emphasize the mental models of agents modelled around motivational construct (e.g., an agent has a disposition towards this need or this fear), but do not focus on hierarchies of activities. In summary, current models do not have different types of activities, because they take states of the world or agents – and not activities – as the primary concepts of their reasoning.

*15:* Current models have implicit (reasoning and normative), limited or no (social-psychological) temporal and ontological relations between activities. As mentioned by now, reasoning and normative models make connections between activities via plans, but these are not explicit. For example, there are no statements about ‘car commuting’ being a kind-of ‘commuting’ or ‘bringing your kids to school’ being a part-of ‘commuting’. Social-psychological models make no temporal or ontological relation between activities.

#### IV. DISCUSSION

The agent models we reviewed are all Turing-complete: they are expressive enough to simulate the computational aspects of any general-purpose computer or computer language. However, we did not evaluate the models on their ability to make certain calculations, but on the extent to which they emphasize concepts and relations from SPT. This entails that the model expresses the semantics of the concept: it models the meaning of the concept by specifying the relation with other objects and imposing certain restrictions on the allowed deductions. Thus we make a difference between ‘enabling’ in a wider sense (i.e., allowing certain computations) and ‘enabling’ in a narrow sense (i.e., supporting modellers to express certain semantics). A good example to illustrate this difference lies in how habits are modelled. Reasoning models enable — in the wider sense — the modeller to simulate habitual behaviour in an agent: a series of very-specific plans ensures that the agent keeps repeating the behaviour in a certain context. However, they do not enable — in the narrow sense — the

modeller to specify and interpret habitual behaviour. Likewise, the Consumat model enables — in the wider sense — the modeller to simulate content-dependent habits: a series of very specific needs and actions ensures the agents only repeat their behaviour in a certain context. However, it does not specify the semantics of context-dependent habits enabling, for example, the direct comparison between (the strength of) two habitual context-action relations. In both cases, the models would become difficult to manage and interpret if one wants to analyze habits. In summary, we are not interested in enabling modellers to make certain computations, but in enabling modellers to express the semantics of habits, sociality and interconnectedness and integrating these aspects via primary concepts and relations in the model.

This paper shows that current agent models do not support (1) explicit reasoning about habits, (2) context-dependent habits and (3) individual learning concerning habits. As shown in Section II-A, both in SPT and social psychology habits are recognized as a key component of behaviour. This paper answers recent calls for thorough integration of habits in agent models [12], [75]. By not supporting explicit reasoning about habits, agents are not able to correctly combine habits with other decision-making concepts. For example, an agent is not able to reproduce the ability of humans to put itself intentionally in a context (e.g., the desk) to trigger a habit (e.g., to work) [29]. Without modelling how habits depend on context the activities of agents will repeat an activity in any context. For example, an agent is not able to reproduce the behaviour of a person that habitually drinks coffee at work, but tea at home. Without modelling individual learning concerning habits an agent is not able to acquire new personal habits or lose old ones. For example, agents are not able to reproduce differences in humans where one person gets easily stuck in the habit of car-driving, while another person switches between driving a car and using a train. Concluding, new agent models need to be developed to integrate social-psychological research on habits in agent models.

This paper shows that current models do not use a comprehensive set of collective concepts, order information around actions and relate individual and collective concepts in order to guide interactions. As shown in Section II-B, both in SPT and agent theory sociality is recognized as a key component of behaviour. This paper answers recent calls for thorough integration of sociality in agent models [24], [37], [76], [77]. Without integrating a comprehensive set of collective concepts agents cannot fully reproduce human ability to reason about a collective world. For example, without concepts such as values, culture and identity an agent cannot understand why another agent refuses to shake hands or travel by car. Although practices (i.e., actions) are not the only concept around which social information can be ordered, ordering information around practices has at least two advantages: a practice is social, that is, it exists on both the individual and collective level (see Section II-B) and ordering social information around practices corresponds to empirical work in neurology on social reasoning [78]. Without connecting individual and collective concepts, agents cannot extend their reasoning about their *own* preferences to form expectations about *others'* preferences.

For example, an agent cannot reason that because the agent itself has a habit to drive a car, chances are high that others share this habit and therefore ask a colleague to carpool. Concluding, new agent models need to be developed to integrate research on SPT and social agents to model sociality in agent models.

This paper shows that current models do not make explicit relations between activities, between each activity and each other model concept (e.g., desires, context) nor model hierarchies of activities. As shown in Section II-C, in SPT, interconnectivity is a key component of behaviour and, in agent theory, interconnectivity is gaining evidence as a useful way of modelling decisions. Without explicit hierarchies and connections between activities, current models do not directly support inferences about the similarity of activities on a personal level (e.g., two activities need the same resource, need to be performed in sequence, promote the same value) or on a social level (e.g., other people will need this resource). Concluding, new agent models need to be developed to integrate SPT and agent research on interconnectivity in agent models.

To integrate SPT in ABM, we need to integrate habituality, sociality and interconnectivity in one agent model. This paper follows a reductionistic approach by splitting up SPT into aspects and splitting up these aspects into requirements. This approach has been highly successful in the physical sciences and makes it possible to understand complex systems by understanding the properties of presumably more basic components. However, it gives the false impression that investigating the organizational features of things is less informative than investigating component properties [79].<sup>11</sup> For example, it is the integration of habits and interconnectivity that enables a model to express the routine of first taking the kids to school and then going to work. And, it is the integration of sociality and interconnectivity that enables agents to infer that others also connect the work-commute with the school-commute. Thus, although some of the agent models perform relatively well when evaluated against a single aspect or a single requirement, to truly integrate SPT in agent models we need to integrate habituality, sociality in interconnectivity in one model.

## V. CONCLUSION

This paper provided a set of requirements for integrating SPT in agent models. We identified three empirically and theoretically relevant aspects of SPT for modelling agent decision-making: habituality, sociality and interconnectivity (Figure 1). Section II discussed these aspects using literature on SPT, agent theory and social psychology and provided a list of requirements for an agent model that aims to integrate SPT.

This paper provided an evaluation of 11 current agent models against the requirements we elicited. We found that current agent models do not fully capture habituality, sociality or interconnectivity nor is there a model that aims to integrate all three aspects. First, current agent models do not support (1) explicit reasoning about habits, (2) context-dependent habits and (3) individual learning concerning habits (see Table I).

Second, current models do not use a comprehensive set of collective concepts, order information around actions and relate individual and collective concepts in order to guide interactions (see Table II). Third, current models do not make explicit relations between activities, between each activity and each other model concept (e.g., desires, context) nor model hierarchies of activities (see Table III). In addition to detailing these specific differences, we discussed that to utilize SPT in ABM, we need to integrate habituality, sociality and interconnectivity in one agent model. In short, although all agent models capture some aspects of SPT, none fully captures any of the individual aspects, nor is there a model that aims to integrate all three empirical and theoretically grounded aspects.

This paper shows the usefulness of a computational agent model that integrates SPT and provides requirements that help modellers to achieve this model. As we discussed, all the agent models we review are all Turing-complete, but they do not incorporate aspects from SPT as primary concepts and relations. Therefore, modellers are not supported in modelling habits, sociality and interconnectivity. We discussed several examples of human behaviour that current domain-independent agent models do not support. First, without an adequate model of habits, an agent is not able to reproduce the ability of humans to put itself intentionally in a context (e.g., the desk) to trigger a habit (e.g., to work). Second, without an adequate model of sociality, an agent cannot reason that because an agent itself has a habit to drive a car, chances are high that others share this habit and therefore ask a colleague to carpool. Third, without an adequate model of interconnectivity, an agent cannot reason that because the school-commute is a kind of commuting, and it commutes by car, the agent will also need to use the car for a school-commute. Last, without integrating habituality, sociality and interconnectivity in one model, agent models do not support agents that combine habits and interconnectivity in a routine of first taking the kids to school and then going to work. Furthermore, the model does not support agents that, in addition, use sociality to expect others have a similar commuting routine and therefore decide to carpool together. Integrating SPT in agent models will help us understand the world in terms of three key aspects of behaviour: lazily habitual, lovingly social and actively interlinked.

#### ACKNOWLEDGMENT

The authors would like to thank Mathijs de Weerd and Gal Kaminka for shedding light on related work from multi-agent planning, and Frank Dignum for the many insightful conversations on social practices.

#### REFERENCES

- [1] B. Edmonds, "Bootstrapping Knowledge About Social Phenomena," *Journal of Artificial Societies and Social Simulation*, vol. 13, no. 2010, pp. 1–16, 2010.
- [2] —, "Different modelling purposes," *Understanding Complex Systems*, no. 9783319669472, pp. 39–58, 2017.
- [3] R. Boero and F. Squazzoni, "Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science," *Journal of Artificial Societies and Social Simulation*, vol. 8, no. 4, pp. 1–31, 2005.
- [4] F. Squazzoni, W. Jager, and B. Edmonds, "Social Simulation in the Social Sciences: A Brief Overview," *Social Science Computer Review*, vol. 32, no. 3, pp. 279–294, 12 2013. [Online]. Available: <http://ssc.sagepub.com/cgi/doi/10.1177/0894439313512975>
- [5] S. Moss and B. Edmonds, "Towards good social science," *JASSS*, vol. 8, no. 4, pp. 1–16, 2005.
- [6] W. Jager, "Enhancing the Realism of Simulation ( EROS ): On Implementing and Developing Psychological Theory in Social Simulation," *Journal of Artificial Societies and Social Simulation*, vol. 20, no. 3, p. 14, 2017.
- [7] T. R. Schatzki, *Social Practices. A Wittgensteinian approach to human activity and the social*. Cambridge University Press, 1996.
- [8] P. Bourdieu, *Outline of a theory of practice*. Cambridge University Press, 1977, vol. 16. [Online]. Available: <papers://cfc50b6a-2d9e-4feb-87e5-d6012043bd5a/Paper/p2118>
- [9] A. Reckwitz, "Toward a theory of social practices: A development in culturalist theorizing," *European Journal of Social Theory*, vol. 5, no. 2, pp. 243–263, 2002.
- [10] E. Shove, M. Pantzar, and M. Watson, *The Dynamics of Social Practice: Everyday Life and How it Changes*. London: SAGE Publications, 2012.
- [11] N. Cass and J. Faulconbridge, "Commuting practices: New insights into modal shift from theories of social practice," *Transport Policy*, vol. 45, pp. 1–14, 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.tranpol.2015.08.002>
- [12] F. Dignum, V. Dignum, R. Prada, and C. M. Jonker, "A Conceptual Architecture for Social Deliberation in Multi-Agent Organizations," *Multiagent and Grid Systems*, vol. 11, no. 3, pp. 147–166, 2015.
- [13] G. Kaminka, "Curing robot autism: A challenge," *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, pp. 801–804, 2013.
- [14] F. Dignum, R. Prada, and G. Hofstede, "From Autistic to Social Agents," in *Proceedings of the 2014 international conference on Autonomous Agents and Multi-Agent Systems*, 2014, pp. 1161–1164. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2617431>
- [15] V. Dignum and F. Dignum, "Contextualized Planning Using Social Practices," in *Coordination, Organizations, Institutions, and Norms in Agent Systems*, A. Ghose, N. Oren, P. Telang, and J. Thangarajah, Eds. Cham: Springer International Publishing, 2015, pp. 36–52.
- [16] G. Holtz, "Generating Social Practices," *Journal of Artificial Societies and Social Simulation*, vol. 17, no. 1, p. 17, 2014.
- [17] K. Narasimhan, T. Roberts, M. Xenitidou, and N. Gilbert, "Using ABM to Clarify and Refine Social Practice Theory," in *Advances in Social Simulation 2015*, W. Jager, Ed., vol. 528. Springer International Publishing AG 2017, 2017. [Online]. Available: <http://link.springer.com/10.1007/978-3-319-47253-9>
- [18] T. Balke, T. Roberts, M. Xenitidou, and N. Gilbert, "Modelling Energy-Consuming Social Practices as Agents," *Advances in Computational Social Science and Social Simulation*, 2014.
- [19] H. C. Triandis, "Values, attitudes, and interpersonal behavior," *Nebraska Symposium on Motivation. Nebraska Symposium on Motivation*, vol. 27, pp. 195–259, 1 1980. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7242748>
- [20] A. Moors and J. De Houwer, "Automaticity: a theoretical and conceptual analysis," *Psychological bulletin*, vol. 132, no. 2, pp. 297–326, 2006.
- [21] C. Hoffmann, C. Abraham, M. P. White, S. Ball, and S. M. Skippon, "What cognitive mechanisms predict travel mode choice? A systematic review with meta-analysis," *Transport Reviews*, vol. 0, no. 0, pp. 1–22, 2017. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/01441647.2017.1285819>
- [22] E. Thorndike, "Intelligence and its uses," *Harper's Magazine*, vol. 140, pp. 227–235, 1920.
- [23] C. Castelfranchi, "Cognitive architecture and contents for social structures and interactions," in *Cognition and multi-agent interaction: from cognitive modeling to social simulation*, 2006, vol. 37, no. 26, pp. 349–363. [Online]. Available: <http://stacks.iop.org/0305-4470/37/i=26/a=B01?key=crossref.f6cb4117463d9c8d19a33f7ad17928b5>
- [24] F. Y. Wang, P. Ye, and J. Li, "Social Intelligence: The Way We Interact, the Way We Go," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1139–1146, 2019.
- [25] T. Schatzki, "On Practice Theory, or What's Practices Got to Do (Got to Do) with It?" in *Education in an Era of Schooling*. Springer Singapore, 2018, pp. 151–165. [Online]. Available: <http://link.springer.com/10.1007/978-981-13-2053-8>
- [26] E. Shove, M. Watson, and N. Spurling, "Conceptualizing connections: Energy demand, infrastructures and social practices," *European Journal of Social Theory*, vol. 18, no. 3, pp. 274–287, 2015. [Online]. Available: <http://journals.sagepub.com/doi/10.1177/1368431015579964>

- [27] A. Giddens, "The constitution of society: Outline of the theory of structuration," *Cognitive Therapy and Research*, vol. 12, no. 4, p. 448, 1984.
- [28] T. Kurz and B. Gardner, "Habitual behaviors or patterns of practice? Explaining and changing repetitive climate-relevant actions," *Wiley Interdisciplinary Reviews: Climate Change*, vol. 6, no. 1, pp. 113–128, 11 2015. [Online]. Available: <http://doi.wiley.com/10.1002/wcc.327> <http://onlinelibrary.wiley.com/doi/10.1002/wcc.327/full>
- [29] W. Wood and D. T. Neal, "A new look at habits and the habit-goal interface." *Psychological review*, vol. 114, no. 4, pp. 843–863, 2007.
- [30] B. Gardner, G. J. De Bruijn, and P. Lally, "A systematic review and meta-analysis of applications of the self-report habit index to nutrition and physical activity behaviours," *Annals of Behavioral Medicine*, vol. 42, no. 2, pp. 174–187, 2011.
- [31] B. Verplanken, "Habits and implementation intentions," in *The ABC of behavioural change*. Oxford, UK: Elsevier, 2005, pp. 99–109. [Online]. Available: <http://opus.bath.ac.uk/9438/>
- [32] P. Lally, C. H. M. Van Jaarsveld, H. W. W. Potts, and J. Wardle, "How are habits formed: Modelling habit formation in the real world," *European Journal of Social Psychology*, vol. 40, no. 6, pp. 998–1009, 2010.
- [33] D. A. Norman and T. Shallice, "Attention to action," in *Consciousness and self-regulation*. Boston: Springer, 1986, pp. 1–18.
- [34] K. Rømhøj-Romluc, "Habit and Attention," in *The Phenomenology of Embodied Subjectivity. Contributions to Phenomenology*. Springer, Cham, 2013, vol. 31, no. 2, pp. 159–167. [Online]. Available: <http://link.springer.com/10.1007/s10743-014-9162-0>
- [35] B. A. Anderson, "The attention habit: How reward learning shapes attentional selection," *Annals of the New York Academy of Sciences*, vol. 1369, no. 1, pp. 24–39, 2016.
- [36] J. Knobe and B. Malle, "The Folk Concept of Intentionality," *Journal of Experimental Social Psychology*, vol. 33, no. 33, pp. 101–121, 1997.
- [37] G. J. Hofstede, "GRASP agents: social first, intelligent later," *AI and Society*, vol. 0, no. 0, pp. 1–9, 2017. [Online]. Available: <http://dx.doi.org/10.1007/s00146-017-0783-7>
- [38] D. Goleman, *Social Intelligence*. Random House, 2011.
- [39] W. Jager, *Modelling consumer behaviour*. The Netherlands: Universal Press, 2000.
- [40] C. Castelfranchi, "Modeling social action for AI agents," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2, no. c, pp. 1567–1576, 1997.
- [41] J. Broersen, M. Dastani, J. Hulstijn, Z. Huang, and L. van der Torre, "The BOID architecture," in *Proceedings of the fifth international conference on Autonomous agents*, 2001, pp. 9–16.
- [42] P. Felli, T. Miller, C. Muis, A. R. Pearce, and L. Sonenberg, "Artificial social reasoning: Computational mechanisms for reasoning about others," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8755, pp. 146–155, 2014.
- [43] G. J. Hofstede and C. Liu, "To Stay or Not to Stay? Artificial Sociality in GRASP World," in *Advances in Social Simulation*. Springer, 2020, pp. 217–231.
- [44] C. Castelfranchi, "For a science of layered mechanisms: beyond laws, statistics, and correlations," *Theoretical and Philosophical Psychology*, vol. 5, no. June, p. 536, 2014.
- [45] S. Cranefield, M. Winikoff, V. Dignum, and F. Dignum, "No pizza for you: Value-based plan selection in BDI agents," *IJCAI International Joint Conference on Artificial Intelligence*, pp. 178–184, 2017.
- [46] R. Mercur, V. Dignum, C. Jonker, and others, "The value of values and norms in social simulation," *Journal of Artificial Societies and Social Simulation*, vol. 22, no. 1, pp. 1–9, 2019.
- [47] F. Dignum, "Autonomous agents with norms," *Artificial Intelligence and Law*, vol. 7, no. 1, pp. 69–79, 1999.
- [48] C. M. Jonker and J. Treur, "Formal Analysis of Models for the Dynamics of Trust based on Experiences," in *European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, vol. 1647, no. June. Berlin: Springer, 1999, pp. 221–231. [Online]. Available: <http://link.springer.com/10.1007/3-540-48437-X>
- [49] G. J. Hofstede, C. Jonker, and T. Verwaart, "An agent model for the influence of culture on bargaining," in *HUCOM 2008 - 1st International Working Conference on Human Factors and Computational Models in Negotiation*, 2008, pp. 39–46.
- [50] N. Gilbert, "When does social simulation need cognitive models?" *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, pp. 428–432, 2005.
- [51] G. Woodruff, "Premack and Woodruff : Chimpanzee theory of mind," *Behavioral and Brain Sciences*, no. 1978, pp. 515–526, 1978.
- [52] K. V. Hindriks, "Programming Rational Agents in GOAL," in *Multi-Agent Programming*, 2009, ch. 4.
- [53] S. Kumar, M. J. Huber, P. R. Cohen, and D. R. McGee, "Toward a formalism for conversation protocols using joint intention theory," *Computational Intelligence*, vol. 18, no. 2, pp. 174–228, 2002.
- [54] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, "Accurate activity recognition in a home setting," *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08*, p. 1, 2008. [Online]. Available: <http://portal.acm.org/citation.cfm?doid=1409635.1409637>
- [55] H. Storf, M. Becker, and M. Riedl, "Rule-based activity recognition framework: Challenges, technique and learning," *Proceedings of the 3d International ICST Conference on Pervasive Computing Technologies for Healthcare*, 2009. [Online]. Available: <http://eudl.eu/doi/10.4108/ICST.PERVASIVEHEALTH2009.6108>
- [56] L. Chen, C. D. Nugent, and H. Wang, "A Knowledge-Driven Approach to Activity Recognition in Smart Homes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 6, pp. 961–974, 2012.
- [57] G. Okeyo, L. Chen, and H. Wang, "Combining ontological and temporal formalisms for composite activity modelling and recognition in smart homes," *Future Generation Computer Systems*, vol. 39, pp. 29–43, 2014.
- [58] A. Ghorbani, P. Bots, V. Dignum, and G. Dijkema, "MAIA: A framework for developing agent-based social simulations," *Jasss*, vol. 16, no. 2, pp. 1–15, 2013.
- [59] J. M. Epstein, *Agent\_Zero: Toward neurocognitive foundations for generative social science*. Princeton University Press, 2014, vol. 25.
- [60] M. P. Georgeff and A. L. Lansky, "Reactive reasoning and planning," *Proceedings of the Sixth National on Artificial Intelligence (AAAI-87)*, pp. 677–682, 1987.
- [61] A. Rao and M. Georgeff, "BDI Agents: From Theory to Practice," *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, 1995.
- [62] D. Pereira, E. Oliveira, and N. Moreira, "Formal modelling of emotions in BDI agents," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5056 LNAI, pp. 62–81, 2008.
- [63] V. Dignum and F. Dignum, "Emergence and enforcement of social behavior," *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, no. July, pp. 2942–2948, 2009. [Online]. Available: <http://www.mssanz.org.au/modsim09/H4/dignum.pdf>
- [64] G. Andrighetto and R. Conte, "Emergence in the Loop: simulating the two way dynamics of norm innovation," *Dagstuhl Seminar Proceedings. Schloss Dagstuhl-Leibniz-Zentrum für Informatik*, pp. 1–30, 2007. [Online]. Available: <http://drops.dagstuhl.de/opus/volltexte/2007/907>
- [65] M. J. Kollingbaum and T. J. Norman, "Norm adoption and consistency in the NoA agent architecture," *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, vol. 3067, no. September, pp. 169–186, 2004.
- [66] W. Jager and M. Janssen, "An updated conceptual framework for integrated modeling of human decision making: The Consumat II," *Eccs 2012*, p. 10, 2012.
- [67] C. Urban, "PECS: A Reference Model for the Simulation of Multi-Agent Systems," *Tools and Techniques for Social Science Simulation*, pp. 83–114, 2000.
- [68] K. Carley and A. Newell, "The nature of the social agent," pp. 221–262, 1994.
- [69] F. Dignum, D. Morley, E. A. Sonenberg, and L. Cavedon, "Towards socially sophisticated BDI agents," *Proceedings - 4th International Conference on MultiAgent Systems, ICMAS 2000*, pp. 111–118, 2000.
- [70] M. Wooldridge and R. Jennings, "Intelligent Agents: Theory and Practice," *The knowledge engineering review*, vol. 10, no. 2, pp. 115–152, 1995.
- [71] C. M. Jonker, M. B. Van Riemsdijk, and B. Vermeulen, "Shared mental models," in *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*, no. January 2010. Springer, 2010, pp. 132–151.
- [72] S. Sardina, L. De Silva, and L. Padgham, "Hierarchical planning in BDI agent programming languages: A formal approach," *Proceedings of the International Conference on Autonomous Agents*, vol. 2006, pp. 1001–1008, 2006.
- [73] L. R. Planken, M. M. de Weerdt, and C. Witteveen, "Optimal temporal decoupling in multiagent systems," *Belgian/Netherlands Artificial Intelligence Conference*, no. January, 2010.
- [74] L. Hunsberger, "Algorithms for a Temporal Decoupling Problem in Multi-Agent Planning," *Aaai-02*, pp. 468–475, 2002.

- [75] A. Solaki, F. Berto, and S. Smets, "The Logic of Fast and Slow Thinking," *Erkenntnis*, no. 0123456789, 2019. [Online]. Available: <https://doi.org/10.1007/s10670-019-00128-z>
- [76] V. Dignum and F. Dignum, "Agents are Dead. Long live Agents!" in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 2020, pp. 1701–1705.
- [77] E. Norling, "Don't Lose Sight of the Forest: Why the Big Picture of Social Intelligence is Essential," in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, no. May, 2016, pp. 984–987.
- [78] T. Metzinger and V. Gallese, "The emergence of a shared action ontology: Building blocks for a theory," *Consciousness and Cognition*, vol. 12, no. 4, pp. 549–571, 2003.
- [79] T. Deacon, *Incomplete nature: How mind emerged from matter*. WW Norton & Company, 2011.
- [80] D. V. Pynadath and S. C. Marsella, "PsychSim: Modeling theory of mind with decision-theoretic agents," *IJCAI International Joint Conference on Artificial Intelligence*, pp. 1181–1186, 2005.
- [81] F. Dignum, "Interactions as Social Practices: towards a formalization," *ArXiv e-prints*, pp. 1–36, 2018. [Online]. Available: <http://arxiv.org/abs/1809.08751>



Graduate School of Natural Sciences.

**Rijk Mercurius** Rijk Mercurius is a PhD Researcher working at Delft University of Technology. He investigates to what extent a computer model of daily routines improves simulation studies. He uses these simulations to gain insights into social systems: environmental commuting, environmental dining, the spread of rumours and hospitals. Rijk graduated cum laude at Utrecht Universities master's program in Artificial Intelligence in 2015. His master thesis compared interventions that promote vegetarian dining and was awarded the Best Thesis Award of the



**Virginia Dignum** Virginia Dignum is Full Professor at Umeå University, Chair of the Social and Ethical Artificial Intelligence group. She is also associated with the Faculty of Technology Policy and Management at the Delft University of Technology in the Netherlands. Her research focuses on value-sensitive design of intelligent systems and multi-agent organisations, in particular on the ethical and societal impact of AI. She is a Fellow of the European Artificial Intelligence Association (EURAI), a member of the European Commission High Level Expert Group on Artificial Intelligence, and of the Executive Committee of the IEEE Initiative on Ethics of Autonomous Systems. She is member of the scientific boards of the Delft Design for Values Institute, the AI4People - European Global Forum on AI, the Responsible Robotics Foundation, the SIDNfonds, and ALLAI-NL the Dutch AI Alliance. In 2006, she was awarded the prestigious Veni grant by the NWO (Dutch Organization for Scientific Research) for her work on agent-based organizational frameworks. She a well-known speaker on the social and ethical impacts of Artificial Intelligence, and is member of the reviewing boards for all major journals and conferences in AI. She has also chaired many international conferences and workshops, including ECAI2016, the European Conference on AI. She has published more than 180 peer-reviewed papers and edited several books, currently yielding a h-index of 30. From 2011 to 2017, she was vice-president of the Benelux AI Association (BNVKI).



**Catholijn Jonker** Prof. dr. Catholijn Jonker is head of the Interactive Intelligence group of the faculty of Electrical Engineering, Mathematics, and Computer Science, TU Delft. Jonker is also full professor of Explainable Artificial Intelligence at the Leiden Institute of Advanced Computer Science of Leiden University. She is a board member and a Fellow of EurAI, member of the Academia Europaea, member of the Royal Holland Society of Sciences and Humanities. In the past she was chair of the Dutch Network of Female Full Professors and of De Jonge Akademie of the Royal Netherlands Academy of Arts and Sciences. Prestigious grants are the NWO VICI (1.5 M€, 2007) personal grant negotiation support systems, and NWO Gravitation consortium grants on "Hybrid Intelligence" (19 M€ subsidy, 41.9 M€, 2019) of which she is vice-coordinator, and "Ethics of Socially Disruptive Technologies" (18 M€ subsidy, 23.6 M€, 2019) of which she is co-applicant.



Manuscript received ...; revised ...

NOTES

<sup>1</sup>A previous version of (part) of this paper is available as a pre-print [Mer-  
cuur, Rijk, Virginia Dignum, and Catholijn M. Jonker. "Modelling Agents En-  
dowed with Social Practices: Static Aspects." arXiv preprint arXiv:1811.10981  
(2018)]. This improved version differs significantly in that it focuses on current  
literature: it includes a review of current agent models and excludes a proposed  
model. In addition, the paper has been thoroughly rewritten to increase clarity,  
motivation and relevance.

<sup>2</sup>And another reciprocal relation between the individual view of the agent  
and the collective view.

<sup>3</sup> [57] identify two other types: simple activities and multi-task activities.  
These types of activities enable a precise temporal activity model where,  
for example, activities overlap. However, modelling these type of activities  
requires a complex quantitative temporal specification that is not needed for  
the longer temporal scale at which ABS studies systems.

<sup>4</sup>Note that other authors use the term ontology to refer to any kind  
of relation between two objects. Temporal relations are thus a subset of  
ontological relations. However, [57] uses the term ontological to refer to  
inferences one can easily make in description logic, whereas he uses the  
term temporal to refer to relations he can make in Allan's temporal logic.

<sup>5</sup>We do not focus on agent-programming languages or agent communica-  
tion protocols, but on conceptual or formal models of agent decision-making.

<sup>6</sup>As [18] mentioned, reactivity is modelled on the assumption behaviour  
is optimal. Habituality differs in that habits are a heuristic: they are a fast  
automatic response that works in most cases but can be contra-intentional.

<sup>7</sup>In addition to confounding pre-conditions and habitual triggers, reasoning  
and normative models do not have a many-to-many relation between actions  
and context-elements, which is necessary to express how strongly a context-  
element triggers an action.

<sup>8</sup>The only exception being the MAIA model where agents reason about the  
collectively of physicality through contextualized norms, that is, other agents  
mostly do X given physical context Y

<sup>9</sup>This limitation connects to [12], [14], [37], [43] who state that sociality  
should be ingrained in the core of their reasoning. Instead of adding extra  
concepts to model sociality (e.g., by extending BDI models), the same  
concepts should be used to form individual reasoning as well in forming  
expectations about others. As such, sociality is not added as a layer on top  
of individual reasoning but is used to shape the reasoning of the agent.

<sup>10</sup>Reasoning models that aim at representing others' mind focus on social  
expectations about *specific* others (e.g., [42], [80]), whereas SPT emphasizes  
a heuristic humans use where they focus on *the* others. Humans make  
assumptions about how most others view an activity [81]. We require that  
models differ between themselves and 'the group'.

<sup>11</sup>Neuroscience is an example where a nearly complete theory of synaptic  
function and only a slightly less complete understanding of neurons has led  
to a less dramatic understanding of human behaviour or social systems than  
one envisioned [79].

LIST OF FIGURES

I A Venn diagram representing how habitual, so-  
cial and interconnected actions come together in  
social practices. . . . . 2

LIST OF TABLES

I Habituality: Verification and Explanation of Dif-  
ferent Agent Models with Respect to our Re-  
quirements. The '?'-column denotes weather the  
requirement is unsatisfied (✗), somewhat satisfied  
(~) or satisfied (✓) by the agent model. . . . . 7

II Sociality: Verification and Explanation of Dif-  
ferent Agent Models with Respect to our Re-  
quirements. The '?'-column denotes weather the  
requirement is unsatisfied (✗), somewhat satisfied  
(~) or satisfied (✓) by the agent model. . . . . 8

III Interconnectivity: Verification and Explanation of  
Different Agent Models with Respect to our Re-  
quirements. The '?'-column denotes weather the  
requirement is unsatisfied (✗), somewhat satisfied  
(~) or satisfied (✓) by the agent model. . . . . 10