A second-order adaptive social-cognitive agent model for prisoner recidivism

Melman, Dorien; Ploeger, Janne B.; Treur, Jan

published in
Highlights in Practical Applications of Agents, Multi-Agent Systems, and Trust-worthiness. The PAAMS Collection
2020

DOI (link to publisher)
10.1007/978-3-030-51999-5_13

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

A Second-Order Adaptive Social-Cognitive Agent Model for Prisoner Recidivism

Dorien Melman, Janne B. Ploeger, and Jan Treur(✉)

Social AI Group, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
dorienmelman@hotmail.com, janneploeger@gmail.com,
j.treur@vu.nl

Abstract. In this study, a second-order adaptive social-cognitive agent model is introduced to examine prisoner recidivism. For comparison between different kinds of prisons and prison policies, recidivism rates from Norway and the USA are used. Two scenarios were used to model the effects of environmental, prison-related, and personal influences on recidivism rates. The presented adaptive social-cognitive agent model is based on a second-order reified network model. The model allows to computationally explore the effects on prisoner recidivism and the learning process for a prisoner's social-cognitive state of mind as a main determinant of recidivism risk.

Keywords: Prisoner recidivism · Hebbian learning · Adaptive agent model

1 Introduction

One of the main goals of imprisonment is to rehabilitate prisoners to a life without crime [1]. However, many studies have shown that a large proportion of prisoners recidivate within a short time [2, 3]. In fact, literature suggests that some prisons have criminogenic effect [4]. In other words, offenders become more, rather than less, criminally oriented due to their experience in prison. Recidivism rates demonstrate - among other things - a prison system's capacity to rehabilitate its offenders [5].

A country that is well-known for its low recidivism rate is Norway [6]. This is remarkable, as Norway is also seen as exceptional in that they have moderate punitive policies. Twenty years ago, Norway exchanged a punitive 'lock-up' approach for prisons that focus on rehabilitation and maintain exceptionally humane conditions [7, 8]. As a consequence, the prison environment is relatively similar to the outside world, except for the restriction of freedom. For example, prisoners are not locked up between bars, kitchens are fully equipped with sharp objects, and prison guards are more concerned with the prisoners. Although it may be against a more revenge-oriented sense of justice to treat criminals well, the recidivism rate has shown an impressive decrease since this change in the prison system and is with 20% now one of the lowest of the world [9]. In contrast to Norway, the recidivism rates are very high for the USA: 60% of the released prisoners is reconvicted within two years [9]. These high recidivism rates suggest that many American offenders are not moved by imprisonment to stay out of trouble [4]. A comparison of the prisons in the USA and the Norwegian prison system reveals large differences. The approach of the USA is much more

F. De La Prieta et al. (Eds.): PAAMS 2020 Workshops, CCIS 1233, pp. 154–167, 2020.
https://doi.org/10.1007/978-3-030-51999-5_13

punitive. Guards have a more hierarchical role and are less concerned with the educational aspect for prisoners and more focused on the punishment against them. More custodial sanctions are being imposed on prisoners, which risk disrupting conventional relationships and pushing offenders into more antisocial contexts [4].

The major differences in prison systems and recidivism rates have created a demand for a better understanding of the relationship between prison environment, societal norms and recidivism. There has been much research into possible influences on prisoner recidivism. For example, one prominent strand of research has focused on how prison experiences influence offending [10–12]. For instance, Mitchel et al. [13] have investigated the impact of personal aspects such as age, race, and gender. Their analysis indicates that imprisonment has a more criminogenic effect among males than females.

Other studies have focused on examining prison policies and practices that improve reentry outcomes [14, 15]. It has been suggested, for example, that visitation reduces recidivism. As [16] has emphasized, the loss of contact with the outside world - especially with regard to family members - might have negative consequences for reintegration back into society. This thought is supported by Mears et al. [17], who found that the extent of visitation has a significant effect in reducing recidivism of all types, which suggests that visitation improves the reintegration of ex-prisoners back into society. Another example of a concept that is related to recidivism is post-release employment. The transition from prison to employment can be very difficult for released prisoners. However, investigation of the relationship between post-release employment and recidivism shows that employment reduces the risk of recidivism. Hence, prisoners benefit from getting a suitable job.

The aim of the study reported here was to model from a Social Cognition perspective the influences on prisoner recidivism using an adaptive agent model based on an adaptive mental network, i.e., a network of mental states that describes the agent's social-cognitive functioning [18, 19, 23, 24]. In Sect. 2 the designed adaptive social-cognitive agent model is introduced. Section 3 illustrates the agent model by example simulations and indicates how parameter tuning was applied. Validation of the model by comparison to empirical data is discussed in Sect. 4. Section 5 provides verification of the model by mathematical analysis.

2 The Adaptive Agent Model

A Network-Oriented Modeling approach was used to design the adaptive agent model with a focus on Social Cognition [18, 19, 23, 24]. This approach can be considered generic and was suitable to create a second-order adaptive agent model for the social-cognitive processes described in Sect. 1. The modeling approach used can be considered as a branch in the causal modeling area which has a long tradition in AI; e.g., see [20–22]. It distinguishes itself by a dynamic perspective on causal relations, according to which causal relations exert causal effects over time, and in addition these causal relations themselves can change over time as well. The basic type of network model used is called *temporal-causal network model*. It provides a useful concept to translate (supported by a dedicated modeling environment [23]) qualitative processes as known from empirical literature into dynamic, numerical computational models that can be used for simulation. It takes states and their causal effects to other states into account as nodes and connections in a causal network.

The causal effects are represented by the connections between the states, that are labeled with weights that determine the causal strength of the effect. These *connection weights* from state X to state Y are denoted by $\boldsymbol{\omega}_{X,Y}$. A state Y's *speed factor* $\boldsymbol{\eta}_Y$ expresses how fast a state changes upon causal impact. The causal impacts from multiple incoming connections for a state Y are aggregated by a *combination function* $\mathbf{c}_Y(..)$. Some of the states have a single incoming connection, of themselves or of one other state. In that case, the simplest combination function, the identity function **id(.)** can be used, which just multiplies the source state value by the connection weight. For multiple incoming connections, more complex functions can be used like the scaled sum **ssum(..)** or advanced logistic function **alogistic(..)**. All three factors together, the connection weights, speed factors and combination functions specify the *network structure characteristics* defining the conceptual representation of the temporal-causal network model.

Adaptiveness of a model is obtained when some of these characteristics, for example, the connection weights $\boldsymbol{\omega}_{X,Y}$, are dynamic, and represented by additional *reification states* $\mathbf{W}_{X,Y}$. Similarly, the speed factor $\boldsymbol{\eta}_Y$ of state Y can be made adaptive by representing it by a reification state $\mathbf{H}_Y$. Adding such states leads to a two-level network with the original base states at the base level and the added reification states at the (first-order) reification level. This construction can be repeated, leading to a second-order reification level. In that way network characteristics of the first-order reification level can become adaptive as well. For example, to allow an adaptive speed of adaptation for $\boldsymbol{\omega}_{X,Y}$, the speed factor $\boldsymbol{\eta}\mathbf{w}_{X,Y}$ of reification state $\mathbf{W}_{X,Y}$ can be made adaptive in this way and represented by a second-order reification state $\mathbf{Hw}_{X,Y}$. This results in a second-order adaptive network model [18, 23]. A conceptual representation of this type of network structure is shown in Fig. 1.

The numerical representation derived from a conceptual representation is shown in Table 1. This numerical representation is automatically derived from the conceptual representation by the dedicated modeling environment that has been developed and was used for the simulation experiments; see [23] or [18], Ch. 9.

Table 1. Numerical representation derived from a conceptual representation of a temporal-causal network model [21, 22].

Concept	Representation	Explanation
State values over time t	$Y(t)$	At each time point t each state Y in the model has a real number value in [0, 1]
Single causal impact	$\mathbf{impact}_{X,Y}(t) = \boldsymbol{\omega}_{X,Y} X(t)$	At t state X with connection to state Y has an impact on Y, using connection weight $\boldsymbol{\omega}_{X,Y}$
Aggregating multiple impacts	$\mathbf{aggimpact}_Y(t) = \mathbf{c}_Y(\mathbf{impact}_{X1,Y}(t), \ldots, \mathbf{impact}_{Xk,Y}(t)) = \mathbf{c}_Y(\boldsymbol{\omega}_{X1,Y} X_1(t), \ldots, \boldsymbol{\omega}_{Xk,Y} X_k(t))$	The aggregated causal impact of multiple states X_i on Y at t, is determined using combination function $\mathbf{c}_Y(\ldots)$
Timing of the causal effect	$Y(t + \Delta t) = Y(t) + \boldsymbol{\eta}_Y[\mathbf{aggimpact}_Y(t) - Y(t)]\Delta t = Y(t) + \boldsymbol{\eta}_Y[\mathbf{c}_Y(\boldsymbol{\omega}_{X1,Y} X_1(t), \ldots, \boldsymbol{\omega}_{Xk,Y} X_k(t)) - Y(t)]\Delta t$	The causal impact on Y is exerted over time gradually, using speed factor $\boldsymbol{\eta}_Y$; here the X_i are all states with connections to state Y

Thus, the following difference and differential equations are obtained:

$$
\begin{aligned}
Y(t+\Delta t) &= Y(t) + \boldsymbol{\eta}_{\gamma}[\mathbf{c}_{\gamma}\left(\boldsymbol{\omega}_{X_1,Y}X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y}X_k(t)\right) - Y(t)]\Delta t \\
\frac{\mathbf{d}Y(t)}{\mathbf{d}t} &= \boldsymbol{\eta}_{\gamma}[\mathbf{c}_{\gamma}(\boldsymbol{\omega}_{X_1,Y}X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y}X_k(t)) - Y(t)]
\end{aligned}
\tag{1}
$$

In this study, three combination functions were used. For states with single incoming connections, the identity function **id(.)** was used. This function does not have any additional parameters and is shown in Eq. 2:

$$\mathbf{id}(V) = V \tag{2}$$

However, two other combination functions were used to aggregate multiple incoming connections, namely the scaled sum functions $\mathbf{ssum}_{\boldsymbol{\lambda}}\mathbf{(.)}$ and the advanced logistic sum function **alogistic(..)** defined in Eq. 3 and Eq. 4 respectively:

$$\mathbf{ssum}_{\boldsymbol{\lambda}}(V_1, \ldots, V_k) = \frac{V_1 + \ldots + V_k}{\lambda} \tag{3}$$

$$\mathbf{alogistic}_{\boldsymbol{\sigma},\boldsymbol{\tau}}(V_1, \ldots, V_k) = [\frac{1}{1+e^{-\boldsymbol{\sigma}(V_1+\ldots V_k-\tau)}} + \frac{1}{1+e^{\boldsymbol{\sigma\tau}}}](1+e^{-\boldsymbol{\sigma\tau}}) \tag{4}$$

The subscripts indicate the dependence of the parameters. Often, when connection weights are non-adaptive, for the scaled sum, as a normalisation the value of parameter $\boldsymbol{\lambda}$ is the sum of all incoming connection weights $\boldsymbol{\omega}_{X,Y}$. This ensures that the outcome is always between [0, 1]. However, in this study $\boldsymbol{\lambda} = 1$ is used, which makes it a (non-scaled) sum combination function. In case of the **alogistic** function, $\boldsymbol{\sigma}$ is the steepness factor and $\boldsymbol{\tau}$ the threshold.

This far, all combination functions and connections weights are considered for states in the base level of the agent model. The parameter values and connection weights are mostly static, and have a constant value that differs per state. However, since learning in the sense of improving certain social-cognitive skills plays an important role in prison, some of the parameters and connection weights are made adaptive [19]. To this end a (second-order) multilevel reified network model is used for the agent. There are two states in the base level which have adaptive incoming connection weights, which is discussed in more detail later below. These connection weights are modeled at the first reification level as states with their own dedicated combination function: $\mathbf{hebb}_{\boldsymbol{\mu}}\mathbf{(..)}$. The idea of this function is that connection weights between base level states are learned. If both states show activity simultaneously, the connection between those states gets stronger and its weight value is therefore learned. This learning process includes some forgetting as well, the extent of which is indicated by the persistence factor $\boldsymbol{\mu}$. A high persistence factor means low level of forgetness, while a low value for the persistence factor shows high level of forgetness which makes learning more difficult [25]. Thus, a fourth combination function used for the adaptive connection weights in the first reification level is the following Hebbian learning function $\mathbf{hebb}_{\boldsymbol{\mu}}\mathbf{(..)}$:

$$\mathbf{hebb}_{\boldsymbol{\mu}}(V_1, V_2, W) = V_1 V_2(1 - W) + \boldsymbol{\mu} W \tag{5}$$

The parameter $\boldsymbol{\mu}$ in this function is the persistence factor, V_1 and V_2 indicate the activation levels of the two connected states, and W is the connection weight that changes over time.

The use of the Hebbian learning combination function results within the modeling environment [23] in the following difference and differential equation for a connection weight reification state $\mathbf{W}_{X_1,X_2}$ for $\boldsymbol{\omega}_{X_1,X_2}$ (according to Eq. 1 applied to state $\mathbf{W}_{X_1,X_2}$):

$$\begin{aligned}
\mathbf{W}_{X_1,X_2}(t+\Delta t) &= \mathbf{W}_{X_1,X_2}(t) + \boldsymbol{\eta}_{\mathbf{W}_{X_1,X_2}}[X_1(t)X_2(t)\left(1-\mathbf{W}_{X_1,X_2}(t)\right) + \boldsymbol{\mu}\mathbf{W}_{X_1,X_2}(t) - \mathbf{W}_{X_1,X_2}(t)]\Delta t \\
\frac{\mathrm{d}\mathbf{W}_{X_1,X_2}(t)}{\mathrm{d}t} &= \boldsymbol{\eta}_{\mathbf{W}_{X_1,X_2}}[X_1(t)X_2(t)\left(1-\mathbf{W}_{X_1,X_2}(t)\right) + \boldsymbol{\mu}\mathbf{W}_{X_1,X_2}(t) - \mathbf{W}_{X_1,X_2}(t)]
\end{aligned} \tag{6}$$

In this study, the states that make use of adaptive connection weights are prisoners' social-cognitive mental states. The social-cognitive wellbeing of a person changes in prison over time, hence the related connections should be learned. In this way, the adaptation of these mental states happens gradually.

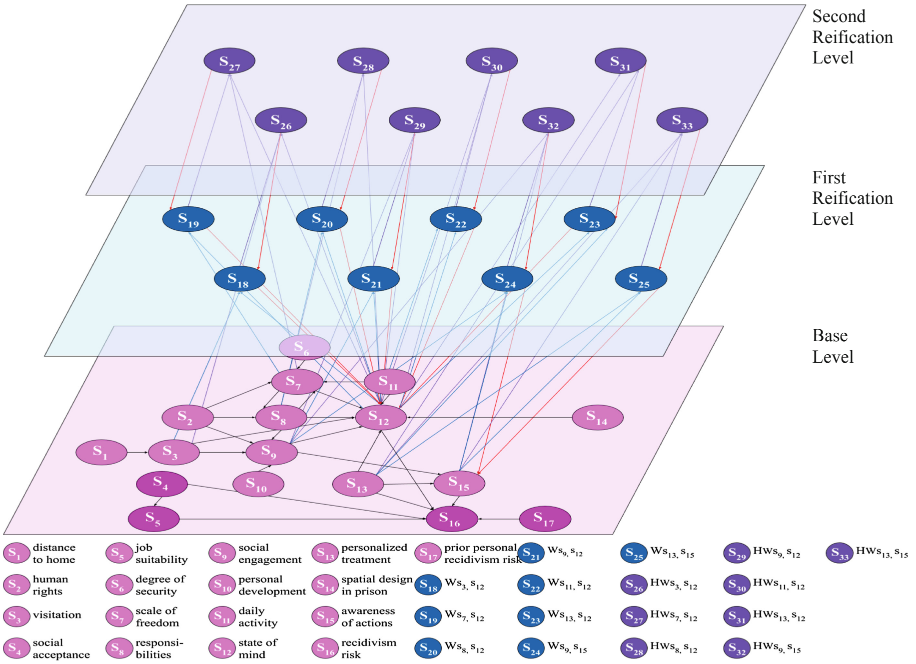


Fig. 1. Second-order adaptive agent based on a three-level reified network model (Color figure online)

The second reification level includes states that enable adaptivity of the speed factor of the adaptive connection weight states for some of the mental states at the base level: the adaptive learning rate. However, a person's ability to learn depends on someone's state of mind in the first place. For example, people that are depressed or stressed seem to show a low learning capacity to get out of that situation. This is due to the high cortisol level and low brain-derived neurotrophic factor (BDNF) level in the body

[26–28]. Together, these levels reduce or block the long-lasting changes in synaptic strength, that are associated with (adaptive) cognitive processes. This block of learning capacity is associated with what has been called negative metaplasticity, which therefore decreases cognitive functioning. Hence, bad cognitive functioning makes it difficult for a person to get out of such a stressed or depressed situation [26–28]. To model this, the speed of the learning process was made adaptive as well: the speed factors of the states in the first reification level are represented by states in the second reification level. The speed factor increases when its input increases. Since the input is learned, the speed factor will be low in the beginning, and raise over time, hand in hand with the learning. Hence, the learning process and the speed of learning are positively correlated in a cyclic manner. The complete model, including both levels of reification states is shown in Fig. 1. The total number of states within the model is 33.

The base level includes 17 states. A distinction is made between states that appear explicitly outside of prison (dark pink) and states that persists both in- and outside prison (light pink). The states outside of prison include S_4 social acceptance to prisoners of society, S_5 job suitability outside of prison, S_{16} recidivism risk, and S_{17} prior recidivism risk of a person.

The last one covers the personal background of an individual. The other states are all states that are maintained in- and outside prison. Two states, S_{12} (state of mind) and S_{15} (awareness of actions) can be seen as the mental states of the prisoner that both are crucial for social-cognitive functioning. Hence, to model the important social-cognitive learning process, these are the states of which most incoming connections are adaptive. A full specification of the adaptive agent model by role matrices [23] can be found at URL https://www.researchgate.net/publication/338547387.

3 Simulation of Example Scenarios

In this section, it is discussed how the influence of prison properties and prisoner states was simulated through the adaptive agent model described in Sect. 2. Two scenarios are taken into account to investigate these influences on a prisoner's social-cognitive mental states and the corresponding recidivism risk. In the first example scenario, both the Norwegian and American imprisonment are simulated for one individual. In the second scenario, the same Norwegian prison is used to simulate different individuals.

3.1 First Scenario: Norway Versus USA

As mentioned above, the first scenario simulates the difference of imprisonment between Norway and the USA for the same individual. This means that for both simulations, the initial values of the personal characteristics' states are equal. These states include S_3 *(family contact)*, S_5 *(job suitability)*, S_9 *(social engagement)*, S_{12} *(state of mind)*, S_{15} *(awareness of actions)*, and S_{17} *(prior personal recidivism risk)*. On the other hand, the prison and country characteristic states are different. These states include S_1 *(proximity to home)*, S_2 *(human rights)*, S_4 *(social acceptance)*, S_6 *(degree of security)*, S_{10} *(personal development)*, S_{11} *(daily activity)*, S_{13} *(personalized treatment)*, *and* S_{14} *(spatial design)*. The remaining states, which are S_7 *(scale of freedom)*, S_8

(responsibilities), and S_{16} *(recidivism risk)* are equal for both simulations, since they involve general aspects before prison. For example, the individual is not restricted in freedom before prison, independent of personal characteristics or the country.

Figure 2 shows the difference between imprisonment for the same individual between a Norwegian prison (a) and an American prison (b). All 33 states from the conceptual model of Fig. 1 are included. As discussed in Sect. 2, the most important states for the person's social-cognitive functioning and development thereof are the mental states S_{12} *(state of mind)* in bold green and S_{15} *(awareness of actions)* in bold blue, and in relation to them S_{16} *(recidivism risk)* in bold orange. The other lines are the remaining states. In both simulations, eight lines, starting at 0.05 show a linear progression. These lines represent the states in the second reification level, which include the adaptive speed factor of the first reification level states. The linear increase indicates that the speed factor is actually adapted hand in hand with the learning process.

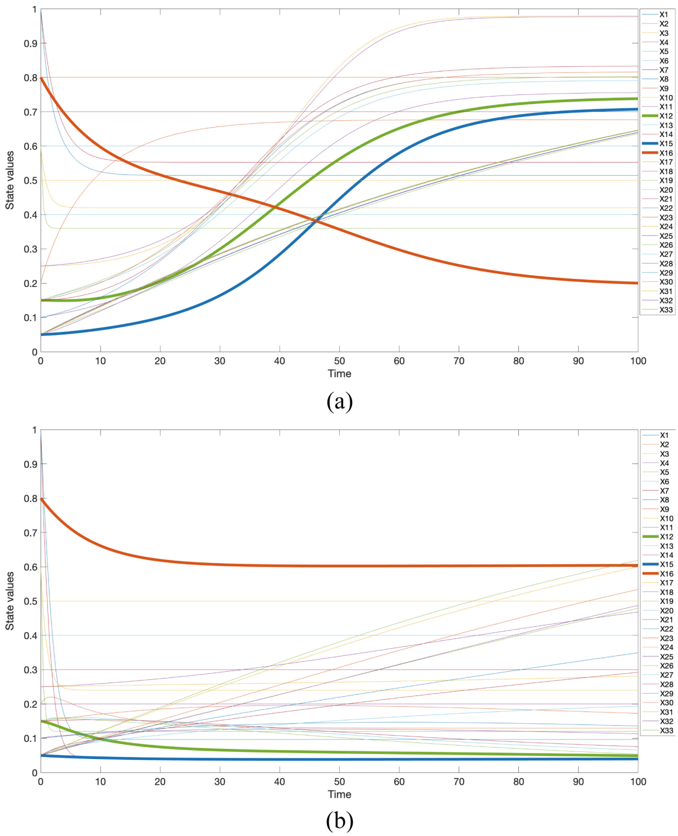


Fig. 2. Simulation of first example scenario: one individual in a Norwegian prison (a) and the same individual in an American prison (b). The y-axis represents the state values, whereas the x-axis represents the time. The time step Δt is set on 0.5. The bold lines represent the most important states: the most relevant mental states in green and blue, and recidivism risk in orange. (Color figure online)

The states in the base level and first reification level show a different course for both Norwegian prison and American prison. First of all, the first-order reification level states that represent the adaptive connection weights, show an seemingly exponential increase until they reach an equilibrium for the Norwegian prison. These states show a different progress for the American imprisonment. Since the imprisonment properties for the USA are in general much lower, the incoming values for the adaptive connection weights are much lower as well. Therefore, these states show a minimal increasing course, after which they decrease to a equilibrium value slightly below the initial state value. Although the speed of learning is slightly learned, the learning itself shows almost no progression.

For the Norwegian imprisonment, the corresponding base level states with adaptive incoming connections (bold green and blue) show an increase that starts more gradually and increase faster when the receiving input gets larger. That is, when the adaptive connection weights have gained a higher value. The slow increase in the beginning of these states and the exponential progress shows that the learning blockade decreases over time. Hence, a Norwegian prisoner learns to be in a better (social-cognitive) state of mind, and during this process, it gets more and more easy to increase this positive state of mind. The final recidivism risk of the Norwegian prisoner shows a value of 0.193.

On the other hand, for the American imprisonment, the corresponding base level states with adaptive incoming connections (bold green and blue) show a slight decrease in the beginning and end up very quick in an equilibrium. The equilibrium value is slightly lower than the initial value, which means that the prisoner even starts feeling worse in prison instead of getting better. The corresponding equilibrium value of the recidivism risk of this prisoner is 0.601.

In order to achieve these outcomes, which almost perfectly match the recidivism rates found by Yukhnenko et al. [9], some model parameters have been tuned automatically. Through Simulated Annealing – a well-known optimization technique that is effective in finding good parameter values for models with a large numbers of parameters [19] – all connection weights to *recidivism risk* (S_{16}) were tuned by using the empirical data in [9]. These weights include $\boldsymbol{\omega}_{S4,S16}$, $\boldsymbol{\omega}_{S5,S16}$, $\boldsymbol{\omega}_{S12,S16}$, $\boldsymbol{\omega}_{S13,S16}$, $\boldsymbol{\omega}_{S15,S16}$, $\boldsymbol{\omega}_{S17,S16}$. The parameter values that resulted in the lowest error after 5000 iterations are included in the final model, of which the results are shown in the subsequent figures.

3.2 Second Scenario: Different Initial Emotional States

The second scenario simulates the difference of imprisonment between two different individuals in a Norwegian prison. This means that for both simulations, the initial values of the prison and country characteristics are equal, whereas the personal characteristics are different. Within this main scenario, two scenarios can be distinguished.

The first scenario is applied to examine the influence of the (social-cognitive) *state of mind* (S_{12}) of the individual on the course of *recidivism risk* (S_{16}). In this scenario, the only difference with the scenario for Norway in Sect. 3.1 is the initial value of the state of mind. The value of 0.2 has been changed into 0.4, which represents a more positive person. The result, depicted in Fig. 3, shows an increase in learning from the

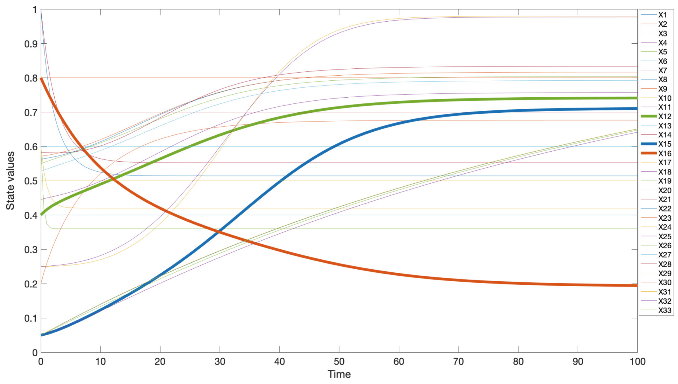
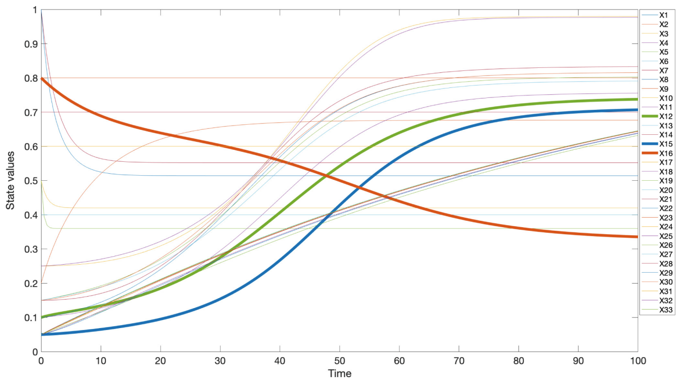


Fig. 3. Simulation of a Norwegian prison for an individual with a relatively positive state of mind.

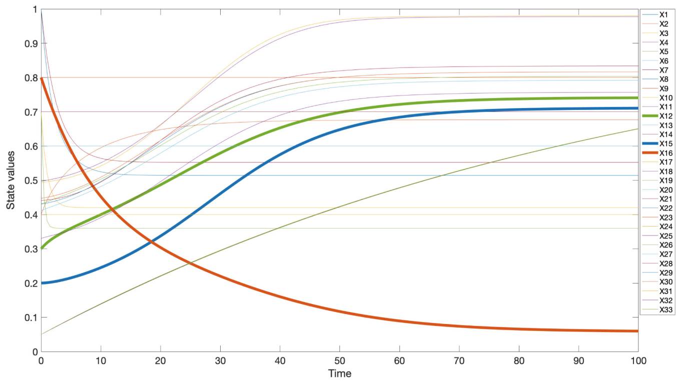
beginning of the simulation. Due to the more positive initial state of mind, the learning blockade is less which enables an earlier increase in state of mind. This leads to a faster increase of the state of mind and therefore a faster decrease of recidivism risk. The final value of recidivism risk is equal to the value in the simulation of Norway in Sect. 3.1.

The second scenario within the main scenario, focuses on the impact of individual differences in a broader sense. The effects of a Norwegian prison are modelled for two different individuals: one overall more positive person and one person that is in a relatively more tough situation (which we will call 'negative'). The differences between the two individuals can be found in the initial values of states that involve personal characteristics, which include *family contact* (S_3), *job suitability* (S_5), *social engagement* (S_9), *state of mind* (S_{12}), *awareness of actions* (S_{15}), and *prior personal recidivism risk* (S_{17}). The relatively positive person is, in comparison with the negative individual, assigned with higher values for all mentioned personal states (since these characteristics are assumed to lead to less criminal behavior), except for the prior personal recidivism risk. The initial value for this latter state is higher for the more negative individual as a higher value represents a person with a more difficult background. The results are shown in Fig. 4.

It can be observed that the recidivism risk decreases less for the simulation of negative individual (Fig. 4a) when compared to the positive individual (Fig. 4b). This results in a recidivism risk of 0.34 and 0.06 at $t = 100$ for the negative and positive individual respectively. Hence, for both the negative and positive individual the recidivism risk has been decreased from the beginning in the Norwegian prison, but it is more difficult for the negative individual to lower this rate when compared to the positive individual. This difference can be explained by the lower initial values of the 'positive' personal characteristics (i.e., characteristics that are assumed to decrease criminality) and higher value of prior personal risk for the negative individual. Due to these initial values, it takes longer to decrease the learning blockade.



(a)



(b)

Fig. 4. Simulation of second example scenario: a relatively negative person in a Norwegian prison (a) and a relatively positive person in the same Norwegian prison (b).

4 Validation

The review of criminal recidivism rates worldwide by [9] provided recent recidivism data. Therefore, these recidivism rates were used to validate the results of the proposed adaptive agent model. Since there exist different definitions of recidivism (e.g. reconviction and reimprisonment) and the definition affects the rate, it is important to use rates that result from the same definition for a proper comparison. In this study, the two-year reconviction rates of Norway and USA (federal) are used. According to the review of Yukhnenko et al. [9], 20% of the Norwegian ex-prisoners is reconvicted within two years, while the reconviction rate for this period is 60% for the USA.

A comparison of the simulation results and the empirical data shows that the outcomes of the proposed model are consistent with these data from literature. This is depicted in Fig. 5, where the empirical values are included. The recidivism rates of the individual used in the simulation scenario are 19.3% and 60.1% for Norway and the

USA respectively (represented with orange in Fig. 5). Hence, the Root Mean Square Error (RMSE) is 0.007 for the simulation of Norway and 0.001 for the USA.

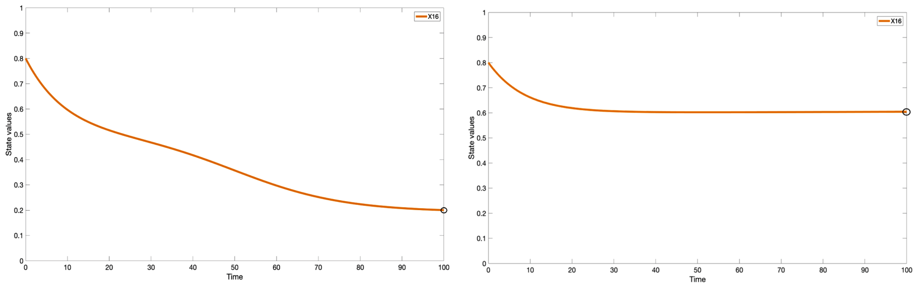


Fig. 5. The simulation results for recidivism of the first example scenario with the empirical data represented in black. (Color figure online)

5 Model Verification by Mathematical Analysis

In order to verify the model, a mathematical analysis of stationary points was performed. This analysis is based on the first scenario for Norway only. The only stationary points verified are the points which reach an equilibrium at the end of the simulation.

A stationary point of state Y at time t is defined as a point where d$Y(t)$/dt = 0. The complete model is said to be in an equilibrium at time t when all states, including connections weights, are in a stationary point [18]. As discussed in Sect. 2, the differential Eq. (1) obtained is

$$\frac{\mathbf{d}Y(t)}{\mathbf{d}t} = \boldsymbol{\eta}_{\gamma}[\mathbf{c}_{\gamma}(\boldsymbol{\omega}_{X_1,Y}X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y}X_k(t)) - Y(t)]$$

with the X_i all states from which state Y gets incoming connections;
here $\mathbf{c}_{\gamma}(\boldsymbol{\omega}_{X_1,Y}X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y}X_k(t))$ = $\mathbf{aggimpact}_Y(t)$

Therefore, in a temporal-causal model, there only exist a stationary point if $\boldsymbol{\eta}_Y = 0$ or $\mathbf{aggimpact}_Y(t) = Y(t)$. Since all speed factors are non-zero, $\mathbf{aggimpact}_Y(t) = Y(t)$ most hold for stationary points at each of the levels. The base level includes 17 states. A distinction is made between states that appear explicitly outside of prison (dark pink) and states that persists both in- and outside prison (light pink). The states outside of prison include S_4 social acceptance to prisoners of society, S_5 job suitability outside of prison, S_{16} recidivism risk, and S_{17} prior recidivism risk of a person. The last one covers the personal background of an individual. The other states are all states that are maintained in- and outside prison. Two states, S_{12} (state of mind) and S_{15} (awareness of actions) can be seen as the most relevant social-cognitive mental states of the prisoner. Hence, these are the states of which most incoming connections are adaptive.

From the simulated model of Scenario 1, the data were extracted for a number of states. A longer time span was used than depicted in Fig. 2(a), to ensure the model has reached an equilibrium if possible. The final values of *Y(t)* and the corresponding connection weight values of that time *t* are used the verify if the model reaches an equilibrium; here $\mathbf{aggimpact}_Y(t)$ is calculated by the corresponding combination function per state. If the derived $\mathbf{aggimpact}_Y$(t) is similar to the extracted value of Y(t), the definition of a stationary point is fulfilled. The results are shown in Table 2, which provides insight in the mathematical verification of the model.

Table 2. An overview of the stationary point state values of the non-constant, non-adaptive states, and states with non-adaptive incoming connection weights at time point t = 300 for which the network-model has reached an equilibrium.

States	S_3	S_5	S_7	S_8	S_9	S_{16}
Time point	300	300	300	300	300	300
$S_i(t)$	0.42	0.36	0.552195	0.514057	0.676508	0.193002
$\mathbf{aggimpact}_{S_i}$(t)	0.42	0.36	0.55219426	0.51405765	0.67650787	0.19300038
Deviation	0	0	−0.0000007	0.0000007	−0.0000002	−0.000002

Hence, all states values in Table 2 fulfill the definition of a stationary point. Additionally, the adaptive connection weights were analyzed as well. Recall the following combination function for the hebbian learning principle: $\mathbf{hebb}_{\boldsymbol{\mu}}(V_1, V_2, W) = V_1 V_2(1 - W) + \boldsymbol{\mu} W$. Also this equation can be used to verify the model by substitution of the values at time point *t* gathered from the simulation. The results are shown in Table 3. Since all deviations are smaller than 0.001, the reification states for the adaptive connection weights in the model show expected behavior as well [18].

Table 3. An overview of the stationary point state values of adaptive connection weight states at time point *t* = 300 for which the network-model has reached an equilibrium.

States	$W_{S3, S12}$	$W_{S7, S12}$	$W_{S8, S12}$	$W_{S9, S12}$	$W_{S11, S12}$	$W_{S13, S12}$	$W_{S9, S15}$	$W_{S13, S15}$
Time point	300	300	300	300	300	300	300	300
State value	0.756923	0.803692	0.792155	0.833768	0.8164642	0.8164642	0.9799634	0.977096
hebb(..)	0.75692320	0.803691945	0.792154892	0.833768436	0.816462101	0.816462101	0.979635529	0.977098257
Deviation	−0.0000002	0.00000006	0.0000001	−0.0000004	0.000002	0.000002	0.0003	−0.000002

6 Discussion

This paper describes an adaptive social-cognitive agent model for the dynamic and adaptive interplay between prison properties, societal aspects, personal characteristics and mental states relating to social-cognitive functioning, and prisoner recidivism. To enable learning of the connection weights to the relevant social-cognitive mental states of the prisoner and to create an adaptive learning speed, that makes the speed of learning dependent of someone's wellbeing [26], two types of levels were included in the model on top of the base level: one for the adaptive connection weights, and on top of that, one for adaptive learning rates. This resulted in a second-order adaptive social-cognitive agent model. The possible influences on prisoner recidivism were modeled and applied in two scenarios. As far as the authors know, these processes were not modeled computationally before.

The first scenario was used to simulate the effect of different prison systems on one individual. Two countries and their prison systems were taken into account: one concerned Norway, and the second one the USA. A comparison of the two simulations showed that the Norwegian prison system resulted in a lower recidivism risk compared to the USA for the same individual. A validation of the model with empirical data reveals that the results correspond with the criminal recidivism rates in [9].

In the second scenario, different individuals were simulated to examine the effect of a prison on different states of mind. These simulations showed that the recidivism risk of an individual with a more positive state of mind, coming from a better environment and good personal circumstances, decreases faster and to a lower level compared to an individual in a more tough situation.

Validation by parameter tuning and verification by mathematical analysis were performed using a first scenario. The parameter tuning focuses on the values that were available as empirical data: recidivism rates. More detailed data on the exact learning process of a prisoner could be useful to improve the model. However, this kind of data is very hard to quantify. The model verification by mathematical analysis showed that the model behaves as expected.

Although the introduced adaptive agent model is a good attempt in modeling prisoner recidivism by taking into account prison, societal and personal aspects, attention should be given to some limitations of this study. First of all, the generalizability of the model is restricted to one definition of recidivism: recidivism as reconviction after two years has been used in this study. However, more definitions of recidivism exist (e.g., reimprisonment) and these come with different recidivism rates [6]. Besides, the time span of recidivism after prison is important as well, since most measurements are done between one and five years after prison. Secondly, only the learning effect of the connection weights to mental states and the speed of learning were taken into account in the current study; that could be extended to more states. Furthermore, this study focused on Norway and the USA only. In order to make the model more robust, it would be a good improvement to investigate more countries for future work and to validate the model on more recidivism data.

References

1. Skardhamar, T., Telle, K.: Post-release employment and recidivism in Norway. J. Quant. Criminol. **28**(4), 629–649 (2012)
2. Baumer, E.: Levels and predictors of recidivism: the Malta experience. Criminology **35**, 601–628 (1997)
3. Langan, P., Levin, D.: Recidivism of prisoners released in 1994. US Department of Justice, Washington, DC (2002)
4. Cullen, F., Jonson, C., Nagin, D.: Prisons do not reduce recidivism: the high cost of ignoring science. Prison J. **91**(3), 48–65 (2011)

5. US Department of Justice: Why recidivism is a core criminal justice concern (2008)
6. Andersen, S., Skardhamar, T.: Pick a number: mapping recidivism measures and their consequences. Crime Delinquency **5**, 613–635 (2017)
7. Ugelvik, T., Dullum, J.: Nordic Prison Practice and Policy - Exceptional Or Not? Nordic Prison Policy and Practice. Routledge, London (2011)
8. Sterbenz, C.: Why Norway's prison system is so successful. Bus. Insider **11** (2014)
9. Yukhnenko, D., Sridhar, S., Fazel, S.: A systematic review of criminal recidivism rates worldwide: 3-year update. Wellcome Open Res. **4**, 28 (2019)
10. Berg, M., Huebner, B.: Reentry and the ties that bind: an examination of social ties, employment, and recidivism. Justice Q. **28**(2), 382–410 (2010)
11. Nagin, D., Cullen, F., Jonson, C.: Imprisonment and reoffending. Crime Justice **38**, 115–200 (2009)
12. Visher, C., Travis, J.: Transitions from prison to community: understanding individual pathways. Ann. Rev. Sociol. **29**, 89–113 (2003)
13. Mitchell, O., Cochran, J., Mears, D., Bales, W.: Examining prison effects on recidivism: a regression discontinuity approach. Justice Q. **4**, 571–596 (2017)
14. Cullen, F., Gendreau, P.: Assessing Correctional Rehabilitation: Policy, Practice, and Prospects. Policies, Processes, and Decisions of the Criminal Justice System, pp. 109–175 (2000)
15. Gideon, L., Sung, H.-E.: Rethinking Corrections: Rehabilitation, Reentry, and Reintegration. Sage, Thousand Oaks (2010)
16. Adams, K.: Adjusting to prison life. Crime Justice **16**, 275–359 (1992)
17. Mears, D., Cochran, J., Siennick, S., Bales, W.: Prison visitation and recidivism. Justice Q. **29**(6), 888–918 (2012)
18. Treur, J.: Network-Oriented Modeling for Adaptive Networks: Designing Higher-Order Adaptive Biological, Mental and Social Network Models. Studies in Systems, Decision and Control, vol. 251. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-31445-3
19. Treur, J.: Network-Oriented Modeling: Addressing Complexity of Cognitive, Affective and Social Interactions. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45213-5
20. Kuipers, B., Kassirer, J.: How to discover a knowledge representation for causal reasoning by studying an expert physician. In: Proceedings of IJCAI 1983 (1983)
21. Kuipers, B.: Commonsense reasoning about causality: deriving behavior from structure. Artif. Intell. **24**, 169–203 (1984)
22. Pearl, J.: Causality. Cambridge University Press, Cambridge (2009)
23. Treur, J.: A modeling environment for reified temporal-causal networks: modeling plasticity and metaplasticity in cognitive agent models. In: Baldoni, M., Dastani, M., Liao, B., Sakurai, Y., Zalila Wenkstern, R. (eds.) PRIMA 2019. LNCS (LNAI), vol. 11873, pp. 487–495. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33792-6_33
24. Treur, J.: Modeling higher-order adaptive evolutionary processes by multilevel adaptive agent models. In: Baldoni, M., Dastani, M., Liao, B., Sakurai, Y., Zalila Wenkstern, R. (eds.) PRIMA 2019. LNCS (LNAI), vol. 11873, pp. 505–513. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33792-6_35
25. Hebb, D.: The Organisation of Behavior. Wiley, New York (1949)
26. Garcia, R.: Stress, metaplasticity, and antidepressants. Curr. Mol. Med. **2**(7), 629–638 (2002)
27. Kirschbaum, C., Wolf, O., May, M., Wippich, W., Hellhammer, D.: Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults. Life Sci. **58**(17), 1475–1483 (1996)
28. Lupien, S., et al.: Stress-induced declarative memory impairment in healthy elderly subjects: relationship to cortisol reactivity. J. Clin. Endocrinol. Metab. **82**(7), 2070–2075 (1997)