How do commitments work? An agent based simulation using data from a recycling campaign in Santiago de Cuba

Hans-Joachim Mosler
Simulation, Integrated Assessment and Modeling
Swiss federal Institute of Aquatic Science and Technology
Dübendorf, Switzerland

Robert Tobias
Simulation, Integrated Assessment and Modeling
Swiss federal Institute of Aquatic Science and Technology
Dübendorf, Switzerland

Abstract - Diverse psychological findings were integrated in an agent based simulation model, on how situational cues affect behavior performance including the research areas of conviction-based behavior selection, norm-influences, prospective memory, habits, implementation intention and self-commitment. The model postulates, that a behavior selected for execution is the most preferred one between all remembered behavior alternatives. Whether a behavior is remembered or not depends mainly on its accessibility and habit strength. The stronger a person is committed to perform a behavior, the more certain situational cues remind and urge the person to perform it. To investigate the model time-series data was gathered on a daily base during a campaign in Santiago de Cuba promoting recycling behavior. The data confirmed the reminding and cognitive tension effect of prompts due to a commitment. Habits showed a reminding effect, but no effect on the preference. Implications of these results are discussed.

Keywords: Agent-based simulation, habits, commitment, implementation intention, time series data, environmental psychology.

1 Introduction

In environmental psychology many behavioral change measures are widely known to be effective [5]. Commitments like signing a 'contract' [2] [3] and prompts, for example in form of a sticker, are widely used and very efficient intervention techniques for changing individual behaviors. Although the effectiveness of the mentioned intervention techniques was investigated and proofed many times, their psychological mechanics are widely ignored. This might be due to the problem that these phenomena do not suite the traditional format of investigation done in psychology. The present paper aims at elaborating the way, commitments, and prompts together with habits affect behavior. For this, we will derive a formalized model, which integrates many relevant findings on effects of situational cues. The model is then applied to analyze time-series of behavior-intensities gathered in a campaign promoting recycling-behavior in Santiago de Cuba. We close with a short discussion of implications for theory, methodology and application.

2 A model of behavior selection

In the model presented, sets of compatible behaviors are selected in competition with each other. Each behavior and also each behavior-intensity is modeled as a behavior alternative with own attributes like habit, affective connotation, and corresponding convictions about their consequences. All effects on behaviors are calculated continuously, so the behavior-intensities can be defined arbitrarily, but the behaviors always have discrete intensities. For the investigations presented in this paper, only one behavior was specified and the behavior selection reduced to the determination of the behavior-intensity.

Based on the literature reviewed [1] [6] [7] [8] we assume that the preference for behaviors is determined by four influences: The comparison of convictions with norms; affective influences; habits; and cognitive tension-states. The strength of each influence depends on the ability and motivation to think about the behavior selection – modeled here as cognition intensity CI – and the cumulated pressure of all tension states – modeled as tension pressure TP. When ability and motivation to think about the behavior selection is high, i.e. when CI is high, the influence of convictions and norms dominate, while when CI is low, affective influences and habits dominate. The influence of convictions, norms, affective connotations and habits declines with raising tension pressure. This formalizes to

\[
\text{behaviorPreference} = TP \times IoTS + (1 - TP) \times (CI \times IoCN + (1 - CI) \times IoAH)
\]

with: TP = tension pressure
IoTS = influence of tension states
IC = cognition intensity
The influence of tension states on the behavior-preference depends on how strong the tension state is and how much a certain behavior has the potential to reduce this tension. To simplify the model for this paper only one tension is considered: The tension to fulfill a request as stated for example in a commitment. The influence of convictions and norms is calculated as sum of the differences between the values of the convictions and the values of the norms. The model presented counts only with two evaluative dimensions: A general return expressing costs and benefits of the behavior execution, and a behavior-specific dimension, that models influences of norms, that do not refer to behavior consequences. The return-conviction is calculated as sum from cost- and benefit-expectations. For this model we assumed, that the benefit-expectation follows a logistic production-function while cost-expectations follow a power-function. The norm, the return-conviction is compared to, is maximizing the return achieved by a certain behavior. The influence of convictions and norms can be calculated as follows:

\[
\text{IoCN} = \frac{\sum_{\text{allNorms}}((1 - \text{ABS}(\text{convictionValue} - \text{normValue})) \times \text{normSignificance})}{\sum_{\text{allNorms}}(\text{normSignificance})}
\]

In the case of ought norms, the difference \(\text{convictionValue} - \text{normValue}\) is set to 0 if it is greater than 0. NormSignificance is an attribute each norm has that expresses, how important or in the terms of [4] how ‘salient’ a norm is for an individual in the moment of behavior selection. When neither strong cognitive tensions urge to the performance of a behavior nor a person is much motivated and able to think about what behavior to perform, the behavior selection is dominated by affective influences and habits. For the model presented, a simple weighted mean of the two influences is considered.

\[
\text{IoAH} = (1 - \text{habitWeight}) \times \text{affectiveInfluence} + \text{habitWeight} \times \text{habitInfluence}
\]

with: \(0 \leq \text{habitWeight} \leq 1\)

The weighting parameter habitWeight expresses the influence of habits on the behavior preference. The affective influence is modeled very similar to the influence of convictions. The affective influence is the result of the comparison of the actual affective state of an individual and the affective connotation of the behavior. The habit model is rather simple: With every behavior-execution in a specific situation the habit for this and similar behaviors raises and in every time-step it decays. The habit raise on similar behavior is modeled as an exponential function of the difference between the executed behavior-intensity and the behavior-intensity, for which the habit raise is calculated. Habits decay proportionally to their value over time. The habit raise and proportion of decay are set to the same value, so that after endless repetition of the behavior, habit reaches 1. The value of habit raise and decay is set in the calibration of the model.

What behavior is executed in a specific situation not only depends on the preference for different behaviors but also on what behavior alternatives are remembered. The model of memory effects is based on the concept of accessibility which expresses the ease with which something is remembered. To model the forgetting of a behavior alternative, the accessibility of this behavior is reduced, to model reminding, the accessibility is raised. Remembering is modeled with a comparison of this accessibility together with the habit of the behavior to a threshold. The raise of accessibility depends on what reminds the behavior. For this paper only the process of commitment to a behavior, the behavior execution and seeing the own prompt has a reminding effect, though the model would account also for reminding effects by the observation of others behaviors or signs of a public commitment. It is assumed, that when committing oneself, all behavior-intensities are remembered and so, the accessibility raised maximal. The accessibility-rise due to the behavior execution and seeing the own prompt was set by calibrating the model to the data. Forgetting is modeled by a proportional decay of the accessibility over time, where the speed of decay is set in the calibration of the model.

This model so far explains the behavior selection, but it still remains open, how prompts and commitment influence this process. This is explained in the next section.

### 3 Modeling prompt and commitment campaigns

A prompt and commitment campaign has effects in the physical world (i.e. signs like memory-aids or from a public self-commitment) and changes in the psychological structure of affected persons. The tension-states provoked by a prompt depend on the form and strength of the commitment of the person to perform behavior. In the presented study only one behavior is considered, so the behavior-specific tension-effect is set to 0 and the intensity-specific tension-effect is maximized (set to 1). These tension-effects are the stronger the stronger the commitment is. The following formulas are assumed.

\[
\text{tensionInfluence} = (\text{behaviorInflu} + \text{intensityInflu}) \times (\text{behaviorIntensity} ^ \text{comPower}) \times \text{comStrength}
\]

with: \(\text{tensionInfluence} = \text{influence of tension states due to a commitment on the preference for a specific behavior-intensity}\)

\[
\text{behaviorInflu} = \text{commitment effect on behavior preference independent of intensity}
\]

\[
\text{intensityInflu} = \text{commitment effect depending on behavior-intensity}
\]
comPower = parameter to define, how much more higher behavior-intensities are preferred to lower ones due to a commitment
comStrength = parameter to define, how much an individual is committed to perform a behavior
All simulations in this paper have been done with behaviorInflu = 0; intensityInflu = 1;
comPower = 2

The tension effects are set up in the moment of a commitment and then decay over time proportionally to their value, where the speed of decay again is calibrated. The tension pressure (the weight of the tension influence on the behavior preference) is set equal to the highest tension influence multiplied with a calibration parameter.

Memory effects of prompts and commitment are modeled simply by adding a constant to the accessibility of all behavior-intensities. These constants are products of parameters for each event type (commitment, behavior execution and observation of prompt as explained above) and the commitment strength. To model a commitment campaign, many agents, that all function as described above though with different individual variable values, are created. If a prompt is perceived by an individual, the accessibility of the committed behavior is raised and the tension-influence activated both proportional to the commitment strength. Then it is determined, which behavior-intensities have sufficiently high accessibility and habit to be remembered in the behavior-selection. For each of the remembered behavior-intensities the preference is calculated out of their tension influence, return-conviction, specific behavior norms, affective connotation and habit. The lowest intensity with the highest preference is then executed.

Due to the huge number of possible experiments, a systematic exploration of the model by variation of all the parameters can not be presented here. In this paper only the calibration of some of the parameters based on empirical data gathered during a campaign is depicted. This connection of simulation and empirical data is essential to assure the usability of the model. The investigation presented here focuses on the following parts of the model: a) effect of the tension-states provoked by a commitment on the behavior-selection, b) dynamics and interaction of the development of accessibility and habits, and c) dynamics and effect of convictions and affective connotations.

4 Methods of the empirical study on recycling behavior in Santiago de Cuba

The data was gathered in a campaign to reduce deposited solid waste in Santiago de Cuba in the beginning of 2005. The interventions were applied to a set of individuals, which were selected with the random-walk sampling method. Individuals filled out a short questionnaire on a daily base to track the dynamics of changes in behavior-intensity and some psychological variables. The interventions for these subjects was as follows: As prompt a simple sheet of white paper with the writing “Please, classify and separate!” and the fractions that shall be separated (i.e. glass, aluminum, paper, carton, and plastic) was handed out with the commentary, to hang it up where the solid-waste is gathered but without any further persuasion to do so. For the commitment we handed out a paper with the writing “Here we separate solid waste!” with the instruction to hang it outside the house, so that everybody on the street can see it. The behaviour-intensity was measured by the question “What have you done with your solid-waste today? I separated to bring to the recycling centre…” with 6 categories to answer: “nearly everything” (coded as 1); “most (about 3/4)” (coded as 0.75); “about half” (coded as 0.5); “a part (about 1/4)” (coded as 0.25); “nearly nothing” (coded as 0.1); and “nothing” (coded as 0). Seven of the households with complete data of the two measurement periods did two interventions. Although the model was examined with more data, in this paper we focus on these seven households. The simulation was run over 55 time-steps which can be interpreted as one day of the campaign each. In the steps or days 1 to 27 the first daily data-gathering took place and in the steps or days 33 to 55 the second. In the week of day 27 to 33 the second intervention and gathering of the panel-data was done and no daily data could be gathered. The first intervention – putting up a prompt – was set for step 0 and the second intervention – the public self-commitment – for step 30. The calibration of the model to the data was mainly done by hand.

5 Results: Calibrating the simulation model with empirical data

Before the simulation model is tested, the empirical data itself is presented and discussed. In figure 1 the dynamics of the mean behavior-intensities of the seven households, which performed both interventions and where complete data of the entire campaign is available are shown. In the first period of daily data gathering, when the households used prompts to raise recycling activity, behavior-intensity first falls rapidly and then raises more slowly again. Although the mean behavior-intensity at the beginning and end of this period are nearly the same, the changes in behavior-intensity are rather strong: The means of the intensity vary over two measurement units of the behavior, what corresponds to half the theoretically possible variance of the behavior. In the second phase, when the households committed themselves publicly to recycle solid waste, the dynamics of the mean behavior-intensities get more tranquil: The means vary in about one...
measurement unit of the behavior. Though, the level of the behavior-intensity is more than a measurement unit higher than the one of the first phase. In total, rather complex dynamics resulted for behavior-intensities even in this a priori defined group of households, where several types of dynamics might be mixed up. The simulation model will have to explain the fall and raise of the behavior-intensity in the first phase and the rather constant intensities at higher level in the second phase. To explain the dynamics of the behavior-intensity now the results of the calibration of the simulation are presented. In figure 1 the dynamics of behavior-intensities of the calibrated model are shown. It can be seen, that the simulated behavior-intensities show the same qualitative pattern. The only bigger deviations can be found at the beginning of each phase, where the simulated behavior-intensities are obviously higher than the empirical ones. Though the error count for the model is still tolerable with an average of 0.02, an absolute daily error of 0.08, and a maximal error of 0.25.

![Figure 1. Empirical and simulated behavior-intensity](image)

Surprisingly, no model could be fitted sufficiently well to the data, where the habits had any influence on the preference of behaviors. Though, habits are quite important in reminding one to perform the behavior! A surprising finding that will be discussed later.

With the model, the dynamics of the behavior-intensities can be explained as follows:

1. Hanging up a prompt seems to have the effect of a weaker commitment. The intervention reminds the persons to recycle (raise of accessibility) and provokes a cognitive tension to comply with the commitment to recycle. This leads to a raise in behavior-intensity from estimated 0.1 to about 0.5.

2. The raise in accessibility is very short-lived and the behavior is forgotten more and more often in the first week. Though not visible in the empirical data, still the behavior-intensity of about 0.5 is the preferred one. People still want to recycle more, but they forget it too often.

3. In the second and third week a habit is developing for the new behavior-intensity and similar behaviors. This habit helps remembering the behavior, so that higher intensities of recycling become accessible again. This leads to a stronger raise in habits of higher behavior-intensities, which then are remembered, too. So slowly the behavior-intensity rises due to a shaping-effect of habits.
4. In the forth week the behavior-intensity becomes stable as the preferred behavior-intensity becomes habitual. Higher behavior-intensities can be remembered, but as they have a lower preference, the behavior-intensity does not rise further.

5. In the fifth week a public self-commitment was applied. This public commitment has a stronger impact than just hanging up a prompt. So the behavior-intensity rises abruptly due to the same effects as mentioned under 1.

6. In the second phase no quick decay of the behavior-intensity due to forgetting is visible. The reason for this is the stronger reminding effect of the prompts (and other situational cues) due to the stronger commitment in this phase. Although not visible in the data, the simulation shows, that equally to the first phase the accessibility falls quickly, but never that deep, that the preferred behavior is not remembered anymore. So no forgetting-effect becomes visible.

7. Although the behavior-intensity in phase two is nearly stable, a slow decay can be observed in the simulation data. This is due to the decay of the tension-state provoked by the commitment.

It was shown, that the model presented can be fitted well and with plausible parameter values to the empirical data. Further the model could explain in detail, why the behavior-intensities develop as observed. These results now shall be discussed on a more general level.

6 Discussion

In the following we interpret the results in an integrative way, discuss some shortcomings of the model, and point out directions of future research.

6.1 Interpretation of the results

An important result of the parameter-settings determined by the calibration-process is the time-frame of psychological processes. While accessibilities decay over days, habits develop over weeks. This result is not surprising, but it has never been shown with dynamic field data, that these common-sense assumptions hold in reality. Probably the most surprising result is the way, habits affect behavior selection. According to the calibrated parameter settings, habits do not have any influence on behavior preferences. Though, they have a rather strong influence on remembering behaviors. This result might be due to the kind of behavior investigated. We assume that extremely simple behaviors, which consist of not more than a simple movement, might be strongly influenced by habits, as their ‘initialization’, which would help remembering them, already means their completion. On the other hand, it seems, that most ‘bad habits’ are not really habits. Many behaviors, that are explained by ‘bad habits’ like for example eating to much sweets, are not automatic reactions to situational cues but much more characterized by a discrepancy of affective and convictional influences. We do not eat so many sweets, because we perform the behavior automatically, but because we like it so much to eat them, although we know, we should not. As most studies to habits do not consider affective influences but only contrast habitual with instrumental influences, this difference could not be found. Finally it was surprising, that in the presented case-study habits did not act as inertia impeding the behavior change but in contrary, they fostered behavior change by helping remembering the newly preferred behavior.

6.2 Shortcomings of the model

Obviously the model integrates a large amount of theories and findings and still remains easily manageable, because the derivation of conclusions is supported by computers. The prize for this, are some strong simplifications and many untested hypotheses used for formalizing the model. Though, exactly because of the formalization this simplifications and hypotheses becomes explicit and it would be possible to change or test assumptions easily. It must be noted that the formulas express only, what is stated verbally and often implicitly in the literature. Naturally, there are many possibilities to make these implicit assumptions explicit and the ones presented, are not necessarily the best. But they are usually the simplest, though linear relations often proved to be not adequate. Further, due to the parameterization, many relations could be varied in the calibration process to test their influence on the results.

6.3 Further research

The used assumptions pinpoint lacks of the actual investigation and suggest directions for further research – not only in the field, but also in the laboratory. But even with all the simplifications and unproved assumptions, the model scored well in replicating and explaining empirical time-series. Naturally, even though much more data was investigated with the model, one campaign alone can not prove the practicality of a model. Not only, that many model parts could not even be examined with the data gathered in Santiago de Cuba, even for the parts examined, much more data and investigation is necessary to find out the limits of the model or to develop better formulas and algorithms. Nevertheless, the presented investigation shows, that the model is not completely off the world.

7 References


[2] Burn, S. M. & Oskamp, S., 1986, “Increasing community recycling with persuasive communication and


