How People Respond to Policies: Integrating Behavioural and Decision-Making Theories into Systems Modelling for Sustainability Transformations

Christoph Schünemann

Leibniz Institute of Ecological Urban and Regional Development
Weberplatz 1, 01217 Dresden, Germany

Contact: c.schuenemann@ioer.de; ORCID: orcid.org/0000-0002-1214-8593

Abstract

Understanding how individuals, groups and societies respond to transformative policies is essential for designing effective sustainability policies and coordinated policy mixes. While psychology, behavioural economics and the social sciences provide a rich body of theories explaining human decision-making and behavioural change, these perspectives remain fragmented and are rarely translated into systems models. At the same time, the social component of socio-technical (STS), socio-ecological (SES) and socio-ecological-technological (SETS) systems is often represented only through highly aggregated concepts such as awareness, acceptance or behaviour, limiting the ability of models to realistically capture social responses to interventions. Behavioural representation in systems models can be grounded in different knowledge bases, including empirical evidence, stakeholder knowledge and established behavioural theories. This paper focuses specifically on theory-based representations and examines how behavioural theories can inform systems models at different degree of structural integration. It first reviews major decision-making and behavioural theories before distinguishing three forms of theory-based representation according to their degree of structural integration: (i) theory-derived behavioural mechanisms, (ii) theory-based behavioural modules embedded within broader SES, STS and SETS models, and (iii) systemic decision-making models in which behavioural theory contributes to the structural architecture of the model. It then discusses how these qualitative representations can, where appropriate, be translated into quantitative System Dynamics simulations for analysing temporal dynamics and policy scenarios. The paper argues that the challenge is not simply to integrate complete behavioural theories into systems models, but to identify and represent the behavioural mechanisms most relevant to the modelling purpose. The paper further discusses the potential of reusable theory- and evidence-grounded behavioural mechanisms, or 'behavioural molecules', as a future direction for making behavioural representations more systematic and transferable across sustainability modelling contexts.

Keywords:

behavioural public policy modelling, decision-making theories, behavioural theories, policy assessment, policy target group

1 Introduction

The design of effective policies and coordinated policy mixes to enable sustainability transformation increasingly hinges on a realistic representation of how individuals, groups and society respond to those interventions (Haddad et al., 2022; Raven & Walrave, 2020). While traditional policy analysis has often treated policy target groups as rational, fully informed decision-makers, a growing body of research in environmental psychology, behavioural economics, political science and the broader social sciences demonstrates that human behaviour is non-linear, bounded by cognition, social norms and institutional contexts (Kern et al., 2019; Lane & Rouwette, 2023; Rogge & Reichardt, 2016; Schrieks et al., 2021). Consequently, the emerging field of *behavioural public policy modelling* seeks to embed decision-making and behavioural-change processes within ex-ante policy assessments so that the intended effects of policies or coordinated policy mixes can be evaluated not only for their technical or ecological effectiveness but also for their likelihood of being adopted by target groups without causing unintended side effects (Kaufman et al., 2021; Strassheim, 2021).

Over recent decades, research from psychology (e.g. Ajzen, 1991; Bandura & Walters, 1977), behavioural economics (e.g. Kahneman & Tversky, 1979; Thaler, 2018), social sciences (sociology) (e.g. Macy & Willer, 2002; Rogers, 2003) and further disciplines has provided a diverse set of theories and frameworks to explain how individuals and groups make decisions and alter behaviour. While decision-making theories focus on how choices are made, behavioural theories explain why people act in specific ways, and behavioural change theories address how actions evolve over time. Integrating these perspectives is essential to assess whether policies achieve their intended impacts on target groups. Despite the richness of these individual approaches, the field remains fragmented, with limited guidance on how to combine insights from multiple traditions into holistic policy assessment frameworks (Schlüter et al., 2017).

Despite these advances, considerably less attention has been paid to how behavioural knowledge should be represented in systems models used to support policy assessment and design for sustainability transformations. Broadly speaking, two complementary modelling strategies can be distinguished. One approach places human decision-making itself at the centre of the analysis, resulting in systemic decision-making models that explicitly represent behavioural processes and their interactions through qualitative systems mapping. A second approach embeds human behaviour within broader socio-technical (STS; Geels, 2004), socio-ecological (SES; Berkes et al., 2001), and socio-ecological-technological systems (SETS; Ahlborg et al., 2019) systems models, where social processes interact with ecological and technical subsystems to explain the dynamics of complex sustainability systems.

While these broader systems frameworks have become increasingly influential in sustainability research, the representation of the social component often remains considerably less developed than the ecological and technical components (Müller-Hansen et al., 2017). Social systems are frequently represented through highly aggregated concepts such as awareness, acceptance or behaviour, limiting the ability of systems models to realistically capture how actors respond to transformative policies and interventions (Schlüter et al., 2017). Conversely, systemic decision-making models provide much richer representations of behavioural processes but typically focus on specific actor groups rather than the wider socio-ecological-technical system. This raises a fundamental methodological question:

How detailed should the representation of human decision-making and behavioural change be for different types of systems models or how much social complexity should be represented in different types of systems models, and at what level of detail?

Behavioural representation in systems models can be grounded in different knowledge bases. In applied systems modelling, behavioural assumptions may be informed by empirical data, stakeholder knowledge, participatory modelling, local theories or context-specific evidence. These approaches are highly relevant be-

cause they can ground behavioural representations in observed behaviour and situated knowledge. However, they raise a different set of methodological questions from the translation of established behavioural theories into systems models.

This paper therefore *focuses specifically on theory-based behavioural representation*. We ask how established behavioural and decision-making theories can be translated into systems models, and how this translation varies in its degree of structural integration, from selected behavioural mechanisms or modules embedded in a broader SES/STS/SETS perspective to systemic decision-making models in which behavioural theory contributes to the architecture of the systems model itself.

Systems Thinking offers a useful framework for translating these behavioural insights into a systemic perspective. Rather than considering behavioural mechanisms in isolation, systems models allow them to be embedded in the wider social, technological and ecological context in which decisions and behaviours emerge. They explicitly represent feedback structures, interdependencies and interactions between different system components, making it possible to explore how behavioural mechanisms may contribute to wider system dynamics. This is particularly relevant for sustainability transformations, where policy interventions can generate indirect effects, adaptation and feedbacks that extend beyond the immediate behavioural response. Such representations can remain qualitative or, where appropriate, be further developed into quantitative simulation models.

This paper contributes to bridging this gap by providing a structured overview of existing theories and concepts to model decision-making and behavioural change in policy contexts. It first situates the origins of these theories, comparing perspectives from psychology, behavioural economics, and social sciences, and highlighting their complementary strengths and limitations in understanding individual and collective behaviour.

It then examines how established behavioural and decision-making theories can be translated into systems models at different degree of structural integration. Rather than treating theory-based representation as a binary choice, the paper distinguishes between (i) theory-derived behavioural mechanisms, (ii) theory-based behavioural modules embedded in broader STS/SES/SETS models, and (iii) theory-based systemic decision-making models in which behavioural theory provides the basis of the model structure. Finally, the paper discusses how qualitative modelling approaches can, where appropriate, be translated into quantitative System Dynamics simulation models to analyse temporal dynamics and evaluate alternative policy scenarios.

By systematically linking established decision-making and behavioural theories to the representation of behavioural mechanisms in complex systems models, this paper provides researchers and policymakers with a framework to better anticipate behavioural responses, identify feedback mechanisms and improve the design and assessment of transformative sustainability policies.

2 Existing theories and concepts to model decision-making or behavioural change processes for policy addressees

2.1 Behavioural policy modelling

Modelling human behaviour in public policy is an evolving field, shaped by ongoing research in environmental psychology, behavioural economics, political science, social science and other fields (Schrieks et al., 2021). Central questions include: whose behaviour should be modelled, and for what purpose should it be changed? To assess the ex-ante effectiveness of future transformative policy or policy mixes, it is essential not only to understand the complex dynamics they may trigger, but also to determine whether it will actually alter the decision-making or behaviour of the policy target groups in the intended way. Besides assessment by expert feedbacks, this can be addressed through behavioural modelling in policy assessment, which mainly uses decision rules to represent behavioural mechanisms at the individual or group level (Schrieks et al., 2021). The goal is to model decision-making and behavioural change processes which are inherently complex and non-linear. Approaches range from simple ad hoc assumptions and psychological theories to methods like

machine learning and artificial neural networks (Lane & Rouwette, 2023), which provide comprehensive syntheses for integrating human decision-making into complex systems models. A key challenge lies in balancing model complexity with the ability to parameterise behavioural rules empirically: rules must reflect real-world decision-making without becoming too intricate for empirical calibration via surveys, expert judgment, or data science.

2.2 Decision-making and behavioural theories for policy modelling

The scientific landscape regarding decision-making and behavioural change draws on three distinct academic traditions, each with its own analytical focus, level of abstraction, and methodological orientation, as summarised in Table 1. Psychology and behavioural science operate primarily at the individual level, offering detailed, granular insight into cognitive and motivational processes through conceptual and qualitative frameworks such as the Theory of Planned Behaviour (Ajzen, 1991), the COM-B Model (Michie et al., 2011), and Habit Theory (Verplanken & Orbell, 2003). While rich in explanatory depth, these models are often limited in their generalisability across broader populations or societal contexts. Behavioural economics, by contrast, spans the individual to the societal level and is characterised by more formalised, quantitative models, based on Nudge Theory (Thaler, 2018), Prospect Theory (Kahneman & Tversky, 1979) and Bounded Rationality (Simon, 2000), that are frequently calibrated to empirical data (Schlüter et al., 2017). However, this precision comes at the cost of simplifying assumptions about human preferences and decision processes (Dold et al., 2025). Finally, sociology and social science situate behaviour within group and societal dynamics, using explanatory frameworks such as Diffusion of Innovations (Rogers, 2003), Social Practice Theory (Reckwitz, 2002), and Collective Action Theory (Heckelman, 2019) to account for social norms, networks, and collective patterns of behaviour. These approaches offer strong conceptual and explanatory power at scale, though they are less suited to generating precise, individual-level predictions. Together, these three traditions provide complementary lenses through which decision-making processes and behavioural change can be understood, from the cognitive processes of individuals to the social norms and values shaping societies.

Table 1: Academic backgrounds on modelling decision-making and behaviour

	Psychology / Behavioural science	**Behavioural economics**	**Sociology / Social science**
Aim	Psychological models provide detailed, granular insights into individual behaviour, capturing cognitive and motivational processes, but are often limited in scope and generalizability	Behavioural economics provides broad, quantitative insights into policy responses, but relies on simplifying assumptions about decision-making and preferences.	Sociological approaches focus on conceptualizing societal-level phenomena, such as norms, networks, and collective practices, offering explanatory power for group behaviour rather than precise individual-level predictions
Behavioural level	Individual, group to societal	Individual, group to societal	Group to societal
Exemplarily theories	- Theory of Planned Behaviour - COM-B Model - Habit Theory	- Nudge Theory - Prospect Theory - Bounded Rationality	- Diffusion of Innovations - Social Practice Theory - Collective Action Theory
Degree of formalisation	Primary explanatory, conceptual, qualitative or semi-quantitative models	Primarily quantitative, formalised models, often calibrated to data with simplifications	Primary explanatory, conceptual, qualitative or semi-quantitative models

Table 2 compares theories and concepts relevant to understanding decision-making, behaviour and behaviour change. Decision-making theories and concepts focus on describing how individuals or groups select among alternatives for a conscious decision (Kahneman & Tversky, 1979). They emphasise cognitive processes, heuristics and rationality, often in single, discrete decision situations. Behavioural theories or concepts explain why people behave in a certain way by considering factors such as attitudes, social norms, and perceived behavioural control (Ajzen, 1991). They are particularly useful for predicting short- to medium-

term behaviours. Behavioural change theories or concepts focus on the processes through which behaviour evolves over time (Kaufman et al., 2021; Prochaska & Velicer, 1997). They often describe stages, motivational factors, and environmental influences that facilitate long-term adoption of desired behaviours.

Table 2: Comparing decision-making, behavioural and behavioural change theories and concepts

	Decision-making theories & concepts	**Behavioural theories & concepts**	**Behavioural Change theories & concepts**
Aim of theory	Explain how individuals or groups choose between alternatives	Explain why people behave in a certain way	Explain how and why behaviour changes over time
Characteristics	- Focus on cognitive processes - Often single decision situations - Examines rationality, heuristics, biases	- Focus on behaviour determinants - Include attitudes, social norms, perceived control - Often applied in short- to medium-term prediction	- Focus on behavioural development over time - Include stages, motivation, environment, self-efficacy - Often used for intervention design
Example theories/models	- Rational Choice Theory - Bounded Rationality (Herbert Simon) - Prospect Theory (Kahneman & Tversky)	- Theory of Planned Behaviour (Ajzen) - Health Belief Model - COM-B Model	- Diffusion of Innovations (Rogers) - Social Cognitive Theory (Bandura) - Rubicon concept
Example: transformation towards climate-neutral buildings	Homeowners decide between different energy renovation measures by weighing the pros and cons	Homeowners adopt energy-saving behaviours based on attitudes toward sustainability, social influence and perceived ease	Homeowners gradually change behaviours (install insulation, adopt solar, use efficient appliances) through stages of awareness, preparation, action and maintenance

For the purpose of modelling the policy impacts on the target group response, the selection of the most suitable theory or concept depends on the intended purpose of the policy intervention. If the policy or policy mix aims to change the long-term behaviour of a target group, then behavioural change theories and concepts are appropriate. An example might be homeowners that gradually change their behaviours (install insulation, adopt renewable energies, use efficient appliances) through stages of awareness, preparation, action and maintenance. If the policy is focussed more short-term to change the decision-making process of its target group, then decision-making theories and concepts are suitable to use. An example might be new subsidies for energy renovation measures of buildings that aim to enhance homeowners' decisions towards performing more energy renovation measures by influencing the weighing the pros and cons. It is important to note, that it is reasonable to select not only one decision-making theory or behavioural change theory to model human response because the different theories have their specific purpose in describing decision-making and behavioural change while neglecting other elements a different theory might consider more explicitly. Accordingly, for modelling the policy impacts on the target group response, it might also be expedient to include theories from both, behavioural change and decision-making.

2.3 From behavioural theories to behavioural mechanisms in systems models

While the theories introduced above originate from different scientific disciplines and differ in their analytical focus, explanatory purpose and level of abstraction, they all seek to explain how individuals and groups make decisions, behave and respond to external interventions. Rather than representing competing explanations, they provide complementary perspectives on human behaviour and decision-making. For systems modelling, however, complete behavioural theories are rarely transferred directly into models. Instead, they provide conceptual foundations from which relevant behavioural mechanisms can be identified and represented according to the purpose of the model.

This distinction is particularly important because decision-making theories, behavioural theories and behavioural change theories address different aspects of human behaviour. Decision-making theories primarily

explain how choices are made, behavioural theories explain why people behave in particular ways, while behavioural change theories describe how behaviour evolves over time in response to interventions or changing behaviour evolves over time in response to interventions or changing conditions. Rather than selecting a single theoretical perspective, systems models can combine complementary behavioural mechanisms originating from different theories to provide a more comprehensive representation of the social component of complex systems.

The central challenge is therefore not deciding which behavioural theory should be incorporated into a model, but rather *which behavioural mechanisms are required to adequately represent human decision-making and behaviour for the specific modelling purpose*. Importantly, theory-based behavioural representation is not a binary choice between 'using a theory' and 'not using a theory'. Established theories can inform systems models at different degree of structural integration. A theory may contribute a single behavioural mechanism or a small number of causal relationships, several theory-derived mechanisms may be assembled into a coherent behavioural module embedded in a broader SES/STS/SETS model, or a behavioural theory may provide the structural backbone of a model centred on the decision-making process itself. Table 3 conceptualises these different degrees of structural integration by distinguishing how behavioural theory shapes the representation of behaviour and how this representation relates to the wider systems model.

Table 3: Different degrees of structural integration of behavioural and decision-making theories in systems models

	Theory-derived mechanisms	**Theory-based behavioural modules**	**Theory-based systemic decision-making models**
How behaviour is represented	A behavioural concept or mechanism is included in the systems model.	Multiple theory-derived variables and mechanisms are combined to represent a more complex behavioural process.	The structure of the systems model itself is derived from a behavioural or decision-making theory.
Exemplary representation in systems models	Social norms → perceived social norm → retrofit adoption	Attitudes + social norms + perceived behavioural control + self-efficacy → intention/decision → retrofit behaviour	Motivation → decision/intention → action → evaluation, based on the Rubicon model, complemented by mechanisms from other behavioural theories
Role of theory	Identifies a relevant behavioural mechanism.	Provides the conceptual basis for a coherent behavioural substructure.	Provides the structural architecture of the decision-making/behavioural process.
Role in the wider system	Supporting element within a broader SES/STS/SETS model.	Distinct behavioural module interacting with other system components.	Behavioural/decision-making process is the central object of the systems model.

To make this distinction more concrete, Table 4 illustrates how established behavioural theories can be translated into specific behavioural mechanisms and subsequently into possible system-model representations. The examples show that a behavioural theory does not need to be transferred as a complete theoretical framework into a systems model. Instead, specific mechanisms or constructs can be selected and operationalised according to the modelling purpose.

Importantly, behavioural mechanisms should not be interpreted as predefined model variables. Rather, they represent generic conceptual elements that may be operationalised differently depending on the modelling approach, the level of abstraction and the research question. For example, perceived social norms may be represented as state variables, causal relationships, reinforcing feedback loops or explicit social network structures, whereas bounded rationality may be operationalised through limited information availability, heuristic decision rules or perception delays.

The selection of an appropriate modelling approach therefore depends not only on the research question but also on the intended level of behavioural representation. Qualitative systems models like Causal Loop Diagrams provide a flexible means of representing behavioural mechanisms, feedback structures and causal interdependencies without requiring precise empirical parameterisation. This structural perspective can be

sufficient to understand how policies influence decision-making processes, identify behavioural barriers and reveal potential policy leverage points.

Table 4: From behavioural theories to behavioural mechanisms in systems models for three examples

Behavioural theory	Behavioural mechanism	Possible representation in systems models
Bounded Rationality	Individuals make decisions under limited information, limited cognitive capacity and uncertainty. Rather than optimising, they rely on heuristics, simplify complex situations and satisfice. Decision quality depends on perceived rather than objective information.	Information availability; perceived complexity; cognitive load; uncertainty; search effort; experience; heuristics; aspiration level; perceived costs and benefits; information delays
Theory of Planned Behaviour (TPB)	Behaviour is determined by behavioural intention, which emerges from the interaction of attitudes towards the behaviour, perceived social norms and perceived behavioural control. Intention is translated into action depending on perceived ability and external constraints.	Attitudes; behavioural intention; perceived behavioural control; subjective norms; perceived behavioural feasibility; expected outcomes; implementation probability; decision threshold
Social Norm Theory	Behaviour is influenced by perceived expectations and observed behaviour of others. Individuals adapt their behaviour to conform to descriptive norms (what others do) and injunctive norms (what others approve of), creating social feedback processes.	Descriptive norms; injunctive norms; peer influence; social conformity; visibility of others' behaviour; social approval; network effects; diffusion processes; reinforcing social feedback loops

When research questions require the analysis of temporal dynamics, policy scenarios or emergent system behaviour, these qualitative representations can be translated into quantitative simulation models such as System Dynamics or agent-based models. However, quantitative representation of behavioural mechanisms is considerably more demanding because it requires empirical parameterisation of social and cognitive variables. Collecting reliable behavioural data across heterogeneous populations is resource-intensive, while differences in values, norms, preferences and decision-making styles limit the transferability of empirical findings across contexts (Schlüter et al., 2017). Furthermore, bounded rationality (Lipman, 1995), subjective perception of information and social desirability bias (Larson, 2019) introduce fundamental uncertainties that cannot be eliminated through additional data collection alone. Consequently, behavioural simulation models remain inherently exploratory, and different operationalisations of the same behavioural theory may produce substantially different model outcomes (Kwakkel & Pruyt, 2013). These limitations do not reduce the value of quantitative behavioural modelling but emphasise the importance of selecting an appropriate level of behavioural representation for the modelling purpose. Qualitative models prioritise structural understanding of behavioural mechanisms and feedback structures, whereas quantitative models enable the exploration of temporal dynamics and policy scenarios at the expense of substantially higher parameterisation effort and uncertainty. Consequently, the appropriate level of behavioural detail should be determined by the purpose of the model rather than by the availability of behavioural theories.

3 Translating decision-making and behaviour theories into systems models

To assess how policies or coordinated policy mixes influence target-group behaviour in sustainability transformations, systems modelling provides a versatile framework to integrate decision-making and behavioural change theories. The following sections focus specifically on theory-based behavioural representation and illustrate how the same general knowledge base, established behavioural and decision-making theories, can be translated into systems models at different degree of structural integration. First, theory-based behavioural mechanisms and modules can be embedded within broader socio-technical (STS), socio-ecological (SES) and socio-ecological-technological (SETS) models. Second, behavioural theories can provide the structural basis of systemic decision-making models in which the decision-making or behavioural-change process itself is the central object of analysis. The translation of these qualitative models into quantitative System

Dynamics or agent-based simulations to explore emergent, time-resolved dynamics is discussed in the last section.

3.1 Theory-based behavioural representation in SES, STS and SETS models

A first way to integrate behavioural theory is to include selected behavioural mechanisms or coherent behavioural modules within systems models with a strong technical or ecological component (see first two columns in Table 3). Such systems models commonly belong to the SES/STS/SETS approaches. Despite their different disciplinary origins, these frameworks consider the social component and its interactions with ecological and technical subsystems.

In this modelling strategy, the overall architecture of the model remains centred on the broader SES/STS/SETS perspective. Behavioural theory therefore does not necessarily determine the structure of the entire model. Instead, selected mechanisms can be introduced where they are necessary to explain how actors respond to policies, technologies, environmental changes or social conditions.

For example, Theory of Planned Behaviour may contribute a behavioural substructure linking attitudes, subjective norms and perceived behavioural control to behavioural intention, while Social Norm Theory may contribute reinforcing social feedbacks. Bounded Rationality may instead be represented through limited information, search effort, perception delays or heuristic decision rules. These mechanisms can then interact with technical or ecological processes already represented in the broader systems model.

This approach allows behavioural and decision-making theories to increase the explanatory depth of the social component without making human decision-making the sole object of the model. The level of behavioural detail can therefore range from a small number of theory-derived causal links to a more coherent behavioural module containing several interacting mechanisms. A key methodological challenge is that theories developed at the individual level must be translated to the level of aggregated actors, groups or populations represented in SES/STS/SETS models. This requires explicit assumptions about aggregation, heterogeneity and the relationship between individual decisions and collective system behaviour. The resulting representation should therefore make transparent which theoretical mechanisms are retained, which are aggregated, and which are omitted.

The main advantage of this approach is modularity: behavioural mechanisms can be added where they are relevant without requiring the entire systems model to be reformulated around a behavioural theory. At the same time, this modularity can become a limitation if the selected behavioural mechanisms are disconnected from one another or if theoretically important interactions are omitted. Theory-based behavioural modules therefore require a clear rationale for mechanism selection and structural integration.

3.2 Behaviour-centred systems models: Theory-based systemic decision-making/behaviour models

At the highest degree of structural integration, behavioural theory does not merely contribute selected mechanisms or a distinct behavioural module to an otherwise established systems model. Instead, the structure of the systems model itself is derived from a behavioural or decision-making theory, with the decision-making or behavioural-change process constituting the central object of modelling. Such systemic decision-making models represent the different stages, mechanisms and feedbacks through which policy target groups form decisions and translate them into behaviour, while allowing additional behavioural mechanisms from complementary theories to be incorporated into this structure.

Systemic decision-making models based on qualitative systems modelling or mapping methods can therefore provide a way to translate established behavioural theories into a dynamic representation of decision-making processes and their interdependencies. Rather than representing behavioural constructs as isolated variables, they make explicit how these constructs interact and how feedbacks within the decision-making process may shape behavioural responses to policy interventions (Schünemann et al., 2025).

In the following, it will be explained for a specific example how this can be managed. Schünemann et al. (2025) dedicated their work to the question of why, despite existing policy frameworks, energy retrofit dynamics in most European countries remain far below the levels required to meet climate mitigation targets. To address this challenge, they developed a systemic decision-making model that maps the decision-making process of single-family homeowners regarding energy retrofit integrating psychological, sociological and economic factors within a complex systems perspective, and thereby accounting for the systemic interdependencies that individual-level barrier analyses have thus far left poorly understood. This is achieved by developing a qualitative Causal Loop Diagram (CLD) depicted in Figure 1 whose core structure is based on the psychological Rubicon model, a framework that describes the transition between different phases in goal-directed human behaviour (Heckhausen & Gollwitzer, 1987). It divides the process into motivation, planning, action, and evaluation phases, emphasising a 'Rubicon', or point of no return, when a decision is made, shifting the focus from deliberation to goal pursuit. The focus of the model explanation centres on the pre-decisional motivation phase (Rubicon phase 1), as it describes the crucial process by which uninterested homeowners decide to dive into the topic energy retrofit. In this pre-decisional phase, the fundamental condition for action is established, based on the homeowner's cost–benefit analysis and a range of psychological factors, while the remaining three phases reflect the circumstances of homeowners in realising the retrofit and their satisfaction with the outcome, which in turn evoke important feedback loops.

Especially within the pre-decisional motivation phase, the model incorporates behavioural mechanisms derived from ten established decision-making and behavioural theories from psychology, behavioural economics and the social sciences (see Table 5). Rather than reproducing these theories individually, the systemic decision-making model synthesises their complementary behavioural mechanisms within an integrated systems perspective, thereby providing a more holistic representation of how policies influence individual decision-making and behaviour.

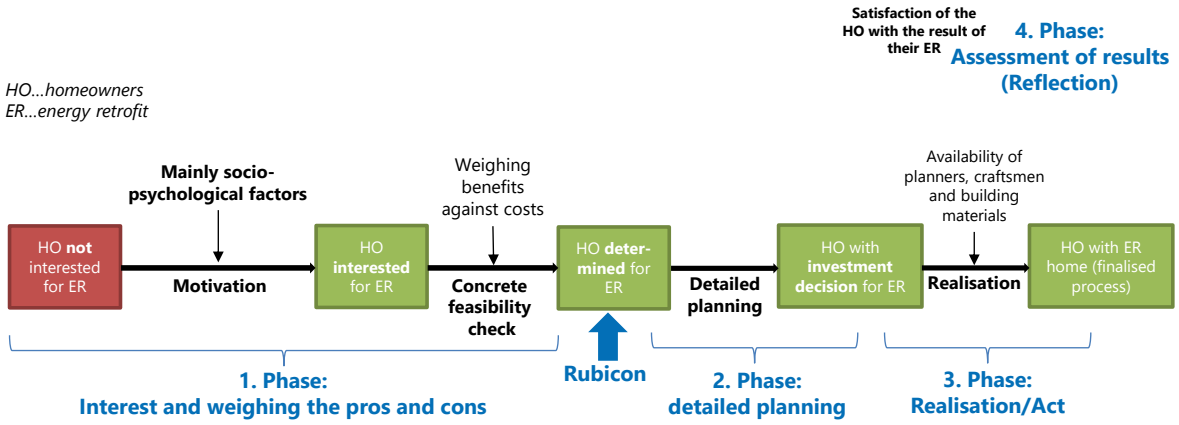
Combining this individual decision-making perspective with a systemic view, by including perceived social norms formed by media and the social environment, the model demonstrates the added value of this integrative approach. Well established theories with a specific focus on human decision-making and behavioural change could be integrated into one systemic decision-making model aiming for a more holistic description where policies are affecting the decision-making process and which feedback loops are influenced. As central outcomes, the exemplary systemic decision-making model reveals that especially media and word-of-mouth effects in the social environment have a remarkable impact on the perceived social norm about energy retrofit among homeowners and thus their decision. Because human decision-making is subject to bounded rationality and systematic cognitive biases, policies premised on purely rational actor assumptions will consistently underestimate the behavioural inertia to be overcome. Accordingly, the model implies that considerably stronger and more behaviourally informed policy interventions are required to achieve the pace of energy retrofit necessary for climate mitigation targets.

The example represents the high end of the structural-integration continuum described above. Behavioural theory does not merely provide individual variables that are inserted into an otherwise independent systems model; rather, the Rubicon model provides the basic architecture of the decision process, while mechanisms from additional theories enrich this architecture. The model therefore illustrates how a systems model can be centred on the behavioural process while still representing interactions with broader social and contextual dynamics.

The underlying theoretical structure may provide a transferable starting point for other contexts of sustainability transformations, but the specific mechanisms, variables and causal relationships would need to be adapted to the respective decision context and system boundary. The main benefit of such behaviour-centred systems models is that they can integrate mechanisms from multiple established decision-making and behavioural theories into a coherent representation of decision-making or behavioural change. However, for policy assessment, especially the pre-decisional phase should be analysed in more detail. Here, quantitative

systems modelling approaches like System Dynamics or agent-based modelling might be suitable to capture time-resolved dynamics based on the emergent system behaviour.

a) Rubicon model to describe the decision-making process of homeowners in energy retrofit



b) Systemic decision-making model (Causal loop diagram) based on the Rubicon structure

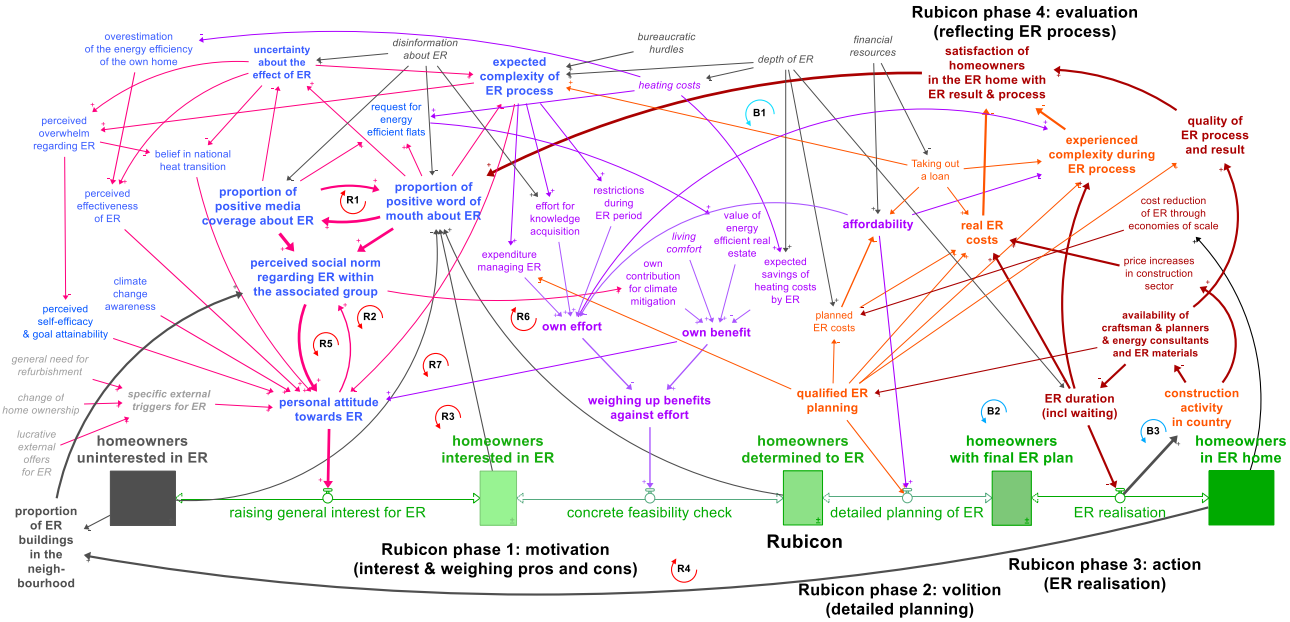


Figure 1: Exemplary systemic decision-making model describing homeowners' (HO) energy retrofit (ER) decision-making process based on the Rubicon model (taken from Schünemann et al. (2025)).

Table 5: Decision-making and behavioural theories represented in the systemic decision-making model (CLD) describing homeowners' (HO) energy retrofit (ER) decision-making process illustrated in Figure 1.

Theory	Core mechanism contributed	Primary disciplinary origin
Bounded Rationality	Decisions under limited information and cognitive constraints	Economics, Psychology
Theory of Planned Behaviour	Behaviour driven by attitudes, social norms and perceived behavioural control	Social Psychology
Social Cognitive Theory	Self-efficacy and observational learning shape behaviour	Psychology
Social Norm Theory	Behaviour influenced by perceived social expectations and peer behaviour	Sociology, Social Psychology
COM-B Model	Behaviour requires capability, opportunity and motivation	Behavioural Science, Psychology
Protection Motivation Theory	Threat and coping appraisal drive protective behaviour	Psychology
Value-Belief-Norm Theory (VBN)	Values and personal norms motivate pro-environmental behaviour	Psychology
Prospect Theory	Decisions depend on perceived gains, losses and risk framing	Behavioural Economics
Social Identity Theory	Group identity and social belonging influence behaviour	Social Psychology
Institutional Theory	Organisational behaviour shaped by formal and informal institutions	Sociology, Political Science

3.3 From qualitative models to quantitative simulation

The previously discussed CLD is an example of a qualitative decision-making representation in systems modelling in the context of sustainability transformations. While these qualitative models already provide valuable insights into behavioural mechanisms, feedback structures, and policy leverage points, quantification becomes particularly relevant when the research objective is to analyse time-dependent system behaviour, compare policy scenarios dynamically, or investigate emergent system behaviour. System Dynamics represents one possible quantitative extension of qualitative systems models when research questions require the explicit analysis of temporal dynamics, nonlinear feedbacks or policy scenarios.

For developing quantitative simulation models based on theory-derived behavioural structures, the key challenge arises because the social and cognitive mechanisms represented qualitatively must be operationalised and parameterised in mathematical form. In the conceptual CLD, variables such as social norms, awareness, perceived benefits, or self-efficacy represent important drivers of decision-making and are connected through feedback loops with economic and environmental variables. However, these variables are inherently qualitative and represent cognitive or social processes rather than directly measurable system states. As a result, they cannot be transferred directly into simulation models without additional assumptions.

A key challenge in this quantification is that parameters governing behavioural relationships are often difficult to estimate. Behavioural variables may represent subjective perceptions, learning processes, social norms or institutional factors that are difficult to quantify, especially when individual differences in decision-making and behaviours are considered. Consequently, theory-based behavioural simulation models may require empirical parameterisation even though their structural logic is derived from theory. Surveys, behavioural experiments, choice experiments, longitudinal studies, expert judgement or other socio-empirical methods may contribute to this parameterisation. However, even when parameterisation is carried out with considerable care, quantified behavioural variables entail uncertainty because future human decision-making and behavioural change under different conditions cannot be predicted with certainty. Accordingly, such

System Dynamics simulations remain exploratory. The theoretical grounding of the model can improve the transparency and justification of behavioural assumptions, but it does not eliminate uncertainty in their quantitative operationalisation.

Rather than replacing qualitative systems models, quantitative SD models should therefore be understood as a complementary extension for addressing specific research questions. For the systemic decision-making model discussed above, for example, it may be reasonable to investigate selected processes within the decision-making structure in greater depth using elaborated SD or agent-based modelling approaches. The key point is that theory-based structural representation and empirical parameterisation address different methodological questions: theory can justify the choice and organisation of behavioural mechanisms, while empirical evidence is required where numerical relationships and parameters are to be estimated.

Conclusion & Outlook

Representing human decision-making and behavioural change remains a central challenge for systems models intended to support sustainability transformations. Behavioural and decision-making theories provide rich and complementary explanations of how individuals and groups respond to interventions, but they are rarely transferred directly into systems models as complete theoretical frameworks. This paper has therefore focused specifically on theory-based behavioural representation and has examined how established behavioural theories can inform systems models at different degrees of structural integration.

The analysis distinguishes three forms of theory-based representation in systems models. *Theory-derived behavioural mechanisms* incorporate selected constructs or causal mechanisms into a broader systems model. *Theory-based behavioural modules* combine several theoretically derived mechanisms into a coherent behavioural substructure embedded within a broader SES, STS or SETS model. At the highest degree of structural integration*, theory-based systemic decision-making models* use behavioural or decision-making theory to provide a substantial part of the model's structural architecture, making the decision-making or behavioural-change process itself the central object of analysis. These forms should not be understood as a hierarchy in which greater integration is necessarily preferable. Rather, they represent different modelling choices, with the appropriate degree of integration depending on the modelling purpose, system boundary and behavioural processes that are relevant to the policy question.

This distinction also clarifies the relationship between theoretical grounding and quantitative operationalisation. *Qualitative systems models* can represent theory-derived behavioural *mechanisms, their interactions and feedback structures* without requiring their immediate translation into numerical parameters. Where research questions concern temporal dynamics, policy scenarios or emergent system behaviour, such representations may subsequently be developed into quantitative System Dynamics or agent-based models. However, increasing the quantitative detail of a behavioural representation does not in itself increase its theoretical validity. *Quantification requires additional assumptions* about measurable constructs, functional relationships, parameters and, where relevant, population heterogeneity. Behavioural variables such as perceived social norms, self-efficacy, perceived complexity or motivation are often context-dependent and difficult to operationalise consistently. Consequently, theory-based structural grounding and empirical parameterisation address different methodological questions: theory can justify which behavioural mechanisms are represented and how they are structurally related, whereas empirical evidence is required where numerical relationships and parameters are estimated.

The paper also highlights an important boundary of theory-based systems modelling. Established behavioural theories provide structured explanations of decision-making and behaviour, but they do not eliminate the heterogeneity, contextual dependence and uncertainty inherent in social systems. Nor does a more detailed behavioural representation necessarily produce a more reliable model. In particular, the translation of theories developed at the individual level into aggregated populations or broader SES/STS/SETS models requires explicit *assumptions about aggregation, heterogeneity and the relationship between individual decisions and*

collective system dynamics. The choice of behavioural representation should therefore be driven by what the model needs to explain, rather than by the ambition to incorporate as much behavioural theory as possible.

This paper has deliberately focused on theory-based behavioural representation. *Empirically grounded, stakeholder-informed and locally informed behavioural representations* constitute important complementary approaches, but involve different methodological questions concerning observation, contextual knowledge and the construction of locally valid behavioural assumptions. They were therefore acknowledged but *not systematically examined here*. Future research should investigate how these different knowledge bases can be combined while maintaining transparency about which elements of a behavioural representation are theoretically derived, empirically supported or grounded in local and stakeholder knowledge.

A further research direction emerging from this perspective is the development of reusable behavioural building blocks, or "*behavioural molecules*". Rather than requiring complete behavioural theories to be integrated into every systems model, behavioural molecules could capture selected, theoretically grounded decision-making or behavioural change mechanisms in a form that can be adapted and combined across different modelling contexts. Such building blocks could potentially support systems modellers in representing human behaviour more systematically, while also making the assumptions and complexity underlying behavioural responses more explicit to policy designers. An important research task will be to determine what constitutes a behavioural molecule, how its theoretical and empirical grounding can be established, how it can be adapted to different contexts and levels of aggregation, and how such building blocks can be integrated into qualitative and quantitative systems modelling approaches.

Taken together, the framework proposed here suggests that advancing behavioural systems modelling does not primarily require choosing a single "best" behavioural theory or transferring complete theories into systems models. Rather, it requires a more explicit process of selecting, translating and structurally integrating the behavioural mechanisms that are relevant to the modelling purpose. Developing systematic and reusable approaches for this translation could help *bridge behavioural science and systems modelling* while retaining the contextual complexity and uncertainty that characterise human responses to transformative sustainability policies.

Use of AI Tools

In the preparation of this article, AI-based language assistance was employed to support the writing, editing, and structuring of the text (Claude, Mistral, ChatGPT). Specifically, the AI was used to help formulate clear, coherent sentences, improve the flow of sections, and summarize complex concepts in decision-making and behavioural science for systems modelling. All intellectual content, scientific arguments, analyses, and interpretations were developed and verified by the authors. The AI did not generate original research, data analysis, or conclusions; its role was solely as a writing and editorial support tool.

Funding

This research was funded by the Leibniz Institute of Ecological and Regional Development (IOER) within the scope of the IOER project 'DEKASYS: Decarbonisation of the German Residential Building Stock—Qualitative and Quantitative System Modelling of the Societal Motivation Processes to Increase the Energetic Renovation Dynamics and the Effectiveness of Policy Instruments'.

Literature

Ahlborg, H., Ruiz-Mercado, I., Molander, S., & Masera, O. (2019). Bringing Technology into Social-Ecological Systems Research—Motivations for a Socio-Technical-Ecological Systems Approach. *Sustainability*, *11*(7), 2009. https://doi.org/10.3390/su11072009

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211. https://doi.org/https://doi.org/10.1016/0749-5978(91)90020-T

Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Bd. 1). Prentice hall Englewood Cliffs, NJ. https://doi.org/10.1177/10596011770020031

Berkes, F., Colding, J., & Folke, C. (Hrsg.). (2001). *Navigating Social-Ecological Systems: Building Resilience for Complexity and Change* (1. Aufl.). Cambridge University Press. https://doi.org/10.1017/CBO9780511541957

Burns, R. (1985). The Letters of Robert Burns. In R. Burns, *The Letters of Robert Burns* (S. 138–138). Oxford University Press. https://doi.org/10.1093/oseo/instance.00033657

Dold, M., Harper, D., Rajagopalan, S., & Whitman, G. (2025). Process, rationality and human wellbeing. *Behavioural Public Policy*, *9*(3), 485–494. https://doi.org/10.1017/bpp.2024.60

Gambardella, P. J., & Lounsbury, D. W. (2020). *Annotated Bibliography: Modeling Psychological and Sociological Dynamics*.

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems. *Research Policy*, *33*(6–7), 897–920. https://doi.org/10.1016/j.respol.2004.01.015

Haddad, C. R., Nakić, V., Bergek, A., & Hellsmark, H. (2022). Transformative innovation policy: A systematic review. *Environmental Innovation and Societal Transitions*, *43*, 14–40. https://doi.org/10.1016/j.eist.2022.03.002

Heckelman, J. C. (2019). Collective Action. In J. C. Heckelman, *The Oxford Handbook of Public Choice, Volume 1* (S. 466–488). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190469733.013.23

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion*, *11*(2), 101–120. https://doi.org/10.1007/BF00992338

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Kaufman, S., Saeri, A., Raven, R., Malekpour, S., & Smith, L. (2021). Behaviour in sustainability transitions: A mixed methods literature review. *Environmental Innovation and Societal Transitions*, *40*, 586–608. https://doi.org/10.1016/j.eist.2021.10.010

Kern, F., Rogge, K. S., & Howlett, M. (2019). Policy mixes for sustainability transitions: New approaches and insights through bridging innovation and policy studies. *Research Policy*, *48*(10), 103832. https://doi.org/10.1016/j.respol.2019.103832

Kwakkel, J. H., & Pruyt, E. (2013). Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change*, *80*(3), 419–431. https://doi.org/10.1016/j.techfore.2012.10.005

Lane, D. C., & Rouwette, E. A. J. A. (2023). Towards a behavioural system dynamics: Exploring its scope and delineating its promise. *European Journal of Operational Research*, *306*(2), 777–794. https://doi.org/10.1016/j.ejor.2022.08.017

Larson, R. B. (2019). Controlling social desirability bias. *International Journal of Market Research*, *61*(5), 534–547. https://doi.org/10.1177/1470785318805305

Lipman, B. L. (1995). Information Processing and Bounded Rationality: A Survey. *The Canadian Journal of Economics*, *28*(1), 42. https://doi.org/10.2307/136022

Macy, M. W., & Willer, R. (2002). From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*, *28*(1), 143–166. https://doi.org/10.1146/annurev.soc.28.110601.141117

Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1), 42. https://doi.org/10.1186/1748-5908-6-42

Müller-Hansen, F., Schlüter, M., Mäs, M., Donges, J. F., Kolb, J. J., Thonicke, K., & Heitzig, J. (2017). Towards representing human behavior and decision making in Earth system models – an overview of techniques and approaches. *Earth System Dynamics*, *8*(4), 977–1007. https://doi.org/10.5194/esd-8-977-2017

Prochaska, J. O., & Velicer, W. F. (1997). The Transtheoretical Model of Health Behavior Change. *American Journal of Health Promotion*, *12*(1), 38–48. https://doi.org/10.4278/0890-1171-12.1.38

Raven, R., & Walrave, B. (2020). Overcoming transformational failures through policy mixes in the dynamics of technological innovation systems. *Technological Forecasting and Social Change*, *153*, 119297. https://doi.org/10.1016/j.techfore.2018.05.008

Reckwitz, A. (2002). Toward a Theory of Social Practices: A Development in Culturalist Theorizing. *European Journal of Social Theory*, *5*(2), 243–263. https://doi.org/10.1177/13684310222225432

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Rogge, K. S., & Reichardt, K. (2016). Policy mixes for sustainability transitions: An extended concept and framework for analysis. *Research Policy*, *45*(8), 1620–1635. https://doi.org/10.1016/j.respol.2016.04.004

Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., Janssen, M. A., McAllister, R. R. J., Müller, B., Orach, K., Schwarz, N., & Wijermans, N. (2017). A framework for mapping and comparing behavioural theories in models of social-ecological systems. *Ecological Economics*, *131*, 21–35. https://doi.org/10.1016/j.ecolecon.2016.08.008

Schrieks, T., Botzen, W. J. W., Wens, M., Haer, T., & Aerts, J. C. J. H. (2021). Integrating Behavioral Theories in Agent-Based Models for Agricultural Drought Risk Assessments. *Frontiers in Water*, *3*, 686329. https://doi.org/10.3389/frwa.2021.686329

Schünemann, C., Dshemuchadse, M., & Scherbaum, S. (2025). A systemic decision-making model for energy retrofits. *Buildings & Cities*, *6*(1), 511–526. https://doi.org/10.5334/bc.534

Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society*, *1*(1), 25–39. https://doi.org/10.1007/BF02512227

Strassheim, H. (2021). Behavioural mechanisms and public policy design: Preventing failures in behavioural public policy. *Public Policy and Administration*, *36*(2), 187–204. https://doi.org/10.1177/0952076719827062

Thaler, R. H. (2018). From Cashews to Nudges: The Evolution of Behavioral Economics. *American Economic Review*, *108*(6), 1265–1287. https://doi.org/10.1257/aer.108.6.1265

Verplanken, B., & Orbell, S. (2003). Reflections on Past Behavior: A Self-Report Index of Habit Strength[1]. *Journal of Applied Social Psychology*, *33*(6), 1313–1330. https://doi.org/10.1111/j.1559-1816.2003.tb01951.x