Testing Changes in the Railway System Through Gaming Simulation: How Different Types of Innovations Affect Operators' Mental Models

Julia C. Lo[a], Jop van den Hoogen[a] and Sebastiaan A. Meijer[b]

[a]Faculty of Technology, Policy and Management
Delft University of Technology
Delft, 2628 BX, The Netherlands

[b]Department of Transport Science
KTH Royal Institute of Technology
Stockholm, 100 44, Sweden

ABSTRACT

Simulation allows innovation managers to manipulate otherwise unchangeable parameters of a railway system and by doing so enable them search for more radical innovative solutions. To validly simulate a sociotechnical system the simulation needs to do justice to both the technical and social complexities. Gaming simulation provides such an opportunity by incorporating real human elements of the system into the simulation run. However, manipulating system elements might have validity-threatening effects on game player's mental models as we assume that real life mental models build up over long periods of interacting with an relatively inert system. This paper studied the relation between the concepts of innovation and mental models by showing the interplay between the different dimensions of both innovation and mental models. We measured the impact of the change(s) on mental models by looking qualitatively at proxies such as erroneous decisions, ambiguity and questions about the introduced change and discussion between players, and applied this to three gaming simulations we ran at ProRail, the organization responsible for managing the Dutch network and one gaming simulation at Network Rail, the British railway infrastructure organization. Our paper ends with a crude proposition: testing innovations that focus on procedures are more cumbersome, especially when not accompanied with an innovation that needs an update of declarative knowledge such as infrastructural changes.

Keywords: Gaming Simulation; Innovation; Mental models; Railway traffic and network control

INTRODUCTION

Innovations in complex sociotechnical systems have two unique features (Markard, 2011). Firstly, path dependent processes and non-linear relations between system elements cause these systems to sometimes reach undesirable lock-ins in local optima. Secondly, radical innovations, innovations intended to defy this path dependence, inherently impact social elements of the system such as operator behavior, rules, roles and procedures. Acknowledging that human input is vital to the validity of simulation experiments, ProRail, the Dutch railway infrastructure manager, started to employ gaming simulation. For them, gaming simulation seemed to do justice to

https://openaccess.cms-conferences.org/#/publications/book/978-1-4951-2099-2

both the technical and social complexities of railway innovations and allowed the organization to freely manipulate system parameters in a shielded and experimental environment. Innovations that were subject to gaming simulation tests are seen as possible changes in the railway system and have involved new traffic control procedures and disruption management principles, new time tables, new safety systems and new infrastructure layouts of nationwide corridors.

Gaming simulations are herein defined as the simulation of a system with participants, in which gaming methods and elements such as game design principles and levels of immersion, are applied. Different types of gaming simulations have been employed in the railway sector, such as individual human-in-the-loop simulators for train traffic operators, multi-actor board games and hybrid versions, in which multiple roles take part from both railway and passenger transport control centers on regional and national level. To test new railway configurations in these gaming simulations, railway traffic and train controllers need to play their own roles and subsequently take decisions as in their real work environment. Thus, the prerequisite to test these innovations is for participants to have a similar mental model and situation awareness of the railway traffic system in the simulated environment, as in the reference environment. However, in testing radically new innovations, one must wonder to what extent experienced operators as game players are able to realistically portray their everyday behavior given their current mental models and resulting situation awareness.

This paper wishes to combine insights from two separated areas of research: innovation and human factors in complex systems and seeks to find out to what extent railway transport innovations can be validly tested using human participants in a gaming simulation. In other words, the authors aim to understand which types of innovations can cause a change in the cognition (c.q. mental models) of operators which might lead to a validity threatening change in in-game decision-making. Thus, the questions that this paper attempts to answer are: *How can railway innovations and mental models be operationalized? How do different types of innovations relate to different conceptualizations of mental models? What are the implications of the relation between these different conceptualizations for the realism of operators' decision-making in a simulated environment?* The approach of this paper is first and foremost a theoretical one, reviewing relevant literature on innovation and human factors of operating complex systems. We build a framework that bridges the different disciplinary areas of research, by which we provide an operationalization of both the concepts of innovation and mental models. An empirical addition will be to investigate the interdependence of these concepts. For this purpose, we investigate four multi-actor tabletop gaming simulations and look how mental models were affected.

HOW COMPLEX SYSTEMS EVOLVE

When a system comprises of many elements that interact in non-simple ways, this system can said to be complex (Simon, 1962). Non-simplicity or complexity involves the interdependence of system elements such as agents and subsystems, that are asynchronous, non-linear and have many feedback and feed forward loops (De Bruijn & Herder, 2009). These products are characterized by a complex interplay of technical components that map into functions (Saviotti and Metcalfe, 1984). In performing their functions, elements in complex systems only work in conjunction with other elements, similar to how in living systems genes only result in certain traits in conjunction with other genes. (Frenken, 2006). This conjunction, called epistatis, cause complex systems to have many local optima (Kauffman, 1993; Frenken, 2006): only certain combinations of system elements have a higher fitness, i.e. they perform and function better than combinations that are only slightly different. Epistatis is crucial to understanding how systems are made up, perform and evolve over time. Insights from biology have inspired scholars of technology and innovation to portray technological systems as sets of elements in epistatic relation to each other. Kauffman (1993) was one of the first to formally model complex systems using his NK-model and introduced the fitness landscape concept from biology to the study of systems in general. A system can have N elements that have K epistatic relations with each other. Since system elements have a finite set of states, the system has a finite set of possible combinations called the possibility space or design space (Frenken, 2006). K then relates to the ruggedness of the fitness landscape: for more complex systems hold that more local optima exist.

Whereas a general system is just a collection of elements that together perform some relevant functions, the notion of 'sociotechnicality' refers to the distinction within these systems between the animate and the inanimate

(Bonen, 1981) or simply the technical and the social (Trist, 1981). Sociotechnical systems are systems that consist of inanimate technical elements and of animate human actors that control, steer, manipulate and operate the technical elements. Later conceptions of sociotechnical systems implicitly addressed a third part of these systems: the rules and institutions that govern how social elements behave in relation to other social elements and in relation to the technical elements. The SHELL model for instance portrayed these systems as having a software part, hardware part, an environment and a collection of liveware, hence the acronym (see Carayon, 2006). In studies on the transitions of these systems, similar notions were used wherein the system itself influences and is influenced by the human actors and the rules and institutions that apply (Geels, 2004). Only in conjunction of both inanimate and animate elements is such a system able to perform its function. Wilson et al. (2007: 102) describe railway systems as a "purposeful system that is open to influences from, and in turn influences, the environment (technical, social, economic, demographic, political, legal, etc.); the people within it must collaborate to make it work properly; and success in implementation of change and in its operation depends upon as near as possible jointly optimizing its technical, social, and economic factors". Railway systems therefore are seen as large socio-technical systems (Markard, 2011).

The fact that systems develop according to evolutionary principles has been acknowledged by the likes of Simon (1962), Bonen (1981) and Nelson and Winter (1982) although on what level the principles of variation, selection and retention precisely work remains equivocal. Evolutionary principles have been applied to designs (Abernathy and Utterback, 1978), routines (Nelson and Winter, 1982), artifacts (Simon, 1962; Frenken, 2006), and sociotechnical systems in general (Kemp et al., 1998; Geels, 2004). Whatever the exact locus of evolutionary processes, all theorists agree on somewhat the same patterns that emerge. Firstly, evolutionary processes create highly path dependent trajectories where a design, artifact or system tends to stabilize into a regime of co-evolved system elements. System evolution tends to gradually grow towards some local optimum. We thus see dominance of one specific configuration over others termed 'dominant designs' (Abernathy and Utterback, 1978), technological regimes (Nelson and Winter, 1982), or sociotechnical regimes (Geels, 2004). Secondly, whereas incremental innovations only reproduce the regime, radical innovations inherently defy this reproduction pattern, i.e. path dependence, and shift the development of the system to another technological trajectory. However, since elements have continually co-evolved with each other, regimes are inherently hostile towards radical innovations. Only when certain environmental conditions apply, radical innovations may be introduced to the system. We therefore see that radical innovations do occur but their occurrence is rare compared to incremental innovations. In the evolution of large systems we therefore see punctuated equilibria (Tushman and Romanelli, 1984) where large time periods of incremental change are combined with short time windows in which regime shifts take place.

Whereas variation in biological systems is caused by random mutations in genes, variation in technological systems occurs in the first place because human actors search purposefully for different system configurations and designing as an activity can be described as a complex optimization problem (Frenken, 2006). Van den Hoogen and Meijer (2014) looked at how these different search strategies are supported by gaming simulation and stated that more distant, i.e. radical, configurations might pose validity threats because of game players not being able to apply their mental models to the new system. However the study neglected, as do others, the fact that certain element changes might have bigger impacts on game players mental models than other element changes. Returning to the sociotechnical framework, systems design can either focus on the system itself, i.e. the hardware, the rules and institutions, i.e. the software and human actors, i.e. the liveware. Similar notions are found in the P-S-I framework (Subrahmanian et al., forthcoming), where design is portrayed as taking place on the level of a product (what is being designed), the level of the social sphere (who designs) and the level of institutions (what rules govern the design process). If we were to translate this to the design *content* of a railway system instead of to the design *process*, the P-S-I framework would frame design of a system as either changing the product space (P), the social space (S) or the institutional space (I). Both frameworks however are relatively abstract and allow many different systems to be studied. Specifically for railway systems, Goverde (2005) typified the dimensions that are the focus of designing a railway system. In his model he states that this design process is of a hierarchical nature, where hardware elements such as tracks, signaling and timetables are designed first and subsequently serve as input for downstream design efforts on for instance crew scheduling and traffic control procedures.

Table 1: Three conceptions of the focus of intended design changes and illustrations specific to the railway system

Geels (2004)	Subrahmanian et al. (forthcoming)	Goverde (2005) slightly adapted	Example
Sociotechnical system	Product Space (P)	Railway network (adapted to include traffic control (1) hardware-part)	Connections, doubling tracks, power supply, signaling, switches, control panels, information systems
		Line planning	Frequencies, service patterns and connections
		Timetable	Detailed planning, clock-face planning
		Rolling stock circulation	Composition of trains, length of service
Humans	Social Space (S)	Crew scheduling	Changing crew schedule
		Traffic control (2) liveware-part	Roles, responsibilities, knowledge
Institutions	Institutional Space (I)	Traffic control (3) software-part	Procedures for handling disruptions, rules, coordination mechanisms

The level of epistasis of that part of the system a designer intends to change is directly related to the position of that part in the hierarchical ladder of railway planning. Thus in designing a system, added tracks, new signaling schemes or removed railway switches might render current line plans and timetables useless or need new ways of controlling traffic. However, in our opinion this cascading effect is not just a matter of where on the hierarchical ladder the change is focused. In addition, certain changes still allow downstream elements to remain unchanged. Adding tracks for instance, might lead to changes downstream (to utilize the added capacity) but does not force them to. On the other hand, removing tracks will certainly lead to new timetables as downstream element are designed in such a way as to maximally utilize the possibilities given by upstream elements. This restrictive cascading property of design changes, next to the hierarchical level of the focal element, both influence the impact of an innovation. We introduce a term from graph theory to depict the impact of a change: centrality.

Insights from evolutionary and complex systems perspectives on innovation have provided us with two dimensions of an intended design change. In Table 2 we summarize the two dimensions of innovations in sociotechnical systems.

Table 2: Two dimensions of innovation

Dimension	Definition	Metrics	Explanation
Centrality	Interdependence of innovation	High vs. low	The extent to which the proposed change will force subsequent changes of other system elements

Focus	Focus of change	Product – Humans - Institutions	Qualitative description of the changed element

CONCEPTUALIZATIONS OF MENTAL MODELS

Mental models are deemed important as they serve as knowledge structures, in which an individual's representation of a physical system can be described (Endsley, 2000a; Klimoski and Mohammed, 1994; Mathieu, Heffner, Goodwin, Salas and Cannon-Bowers, 2000). They have been frequently used by researchers to explain individuals' cognitive functioning and performance (Salas, Stout and Cannon-Bowers, 1994). Moreover, mental models are assumed to be a fundamental mechanism for the establishment of situation awareness (SA), as without a well formulated mental model, attention is not directed to certain cues and thereby operators might oversee certain elements in the environment (level 1 SA), operators are not able to establish a good comprehension of the situation (level 2 SA), nor able to make a good prediction of future states (level 3 SA) (e.g. Endsley, 1988). Subsequently, a high situation awareness is expected to be a predictor of good decision-making in operational settings.

However, operators in complex socio-technical systems hardly execute tasks in a solitary environment. On the contrary, the high interdependence between actors often define and shape the complexities of these systems, emphasizing the importance of team knowledge for operators' cognition. Team mental models are also known as e.g. common cause maps, teamwork schemas, shared frames, socio-cognition, transactive memory (Klimoski and Mohammed, 1994). Mental models in teams have been predominantly measured through the accuracy and similarity between team members (Mohammed, Ferzandi and Hamilton, 2010). As a result of the need to compare individual mental models, the mental model construct has been operationalized to task-team types of mental models and knowledge structure models.

One type of conceptualization of mental models can be distinguished in terms of technology/equipment, job/task, team-interaction and team (ETTT) models (Cannon-Bowers, Salas and Converse, 1993; Lim and Klein, 2006; Matthieu et al., 2000). The technology/equipment model is related to the technology and equipment that is used to execute tasks in a team. This also involves indirect interaction, such as changing the direction of railway switches through computerized systems. The technology/equipment model is the most stable of the four types as the components from this model do not often change. The task model is related to the perception and understanding of procedures, strategies and so forth, in which operators need to understand the ways how to accomplish their task, e.g. necessary information and procedures. The influence of environmental conditions on the task and task demands, such as changed weather conditions or sudden peaks in passenger flows, are also part of the job/task model. Thirdly, the understanding of the responsibilities, norms and interaction patterns of other team members is part of the team-interaction model. Procedures, such as which team members need to interact with each other, what kind of particular

Table 3: Conceptualization of mental models in terms of equipment-task-team interaction-team types of models (ETTT mental model conceptualization) (Cannon-Bowers et al., 1993; Matthieu et al., 2000)

Type	**Knowledge contents**	**Railway knowledge components**	**Stability of the model contents**
Technology/ equipment model	Equipment functioning, operating procedures, equipment limitations, likely failures	Network layout, such as railway tracks, switches, signals; Computerized systems, such as the PRL (train traffic flow) system, dynamic timetable interface	High
Job/ task model	Task procedures, task strategies, environmental conditions, likely contingencies, likely scenarios	Task procedures, such as the role dependent operating procedure, TAD (train order protocol); Environmental conditions, such as the weather,	Moderate

Team interaction model	Roles/ responsibilities, role interdependencies, information patterns, information sources, communication patterns	Roles as defined in the operating procedure	Moderate
Team model	Knowledge over teammates' knowledge, skills, abilities, preferences, tendencies	Team configurations as in planned working shifts	Low

information is needed, but also knowledge when to help team members are also knowledge contents related to the team-interaction model. Finally, the team model is related to the understanding of knowledge, preferences, skills, attitudes, strength and weaknesses of other team members. The team model has a rather low model content stability due to frequent changes in teams, e.g. as railway traffic operators work in shifts, they often need to collaborate with different colleagues. Thus, as they might not work together in the same team configuration for a long period of time, team members develop their knowledge about the abilities, preferences etc. of their colleagues more slowly. In Table 3, an illustration is provided with knowledge components from the railway domain.

A second type of operationalization of mental models can be realized through knowledge structures, i.e. declarative, procedural and strategic (DPS) mental models (see Table 4) (Mohammed et al., 2010; Salas et al., 1994). Declarative models refer to knowledge of facts, rules and relationships (knowledge of what). Procedural models refer to the timing and sequential type of knowledge (knowledge of how). Strategic models refer to knowledge that forms the basis for problem solving (knowledge of the concept and contingency plans). In relation to the ETTT mental model conceptualization, these three types of knowledge can be applied to one single knowledge content. For example, declarative knowledge can be related to facts and rules of a railway switch (e.g. single slip, double slip, outside slip). Procedural knowledge of a railway switch is related to how a railway switch works and how it can be operated. An illustration of strategic knowledge is using a specific switch to reroute a train to a different railway track if the railway track for that train is blocked, and knowing that no other train is currently using the alternative railway track.

Table 4: Conceptualization of mental models in terms of knowledge structures (DPS mental model conceptualization) (Mohammed et al., 2010; Salas et al., 1994)

Type	Definition	Knowledge contents	Example
Declarative model	Information about concepts and elements, and their relationship	Facts, rules, relationship, knowledge about the overall system task goals, the relation among system components, equipment/hardware, position/roles, and the team members themselves	Umbrella: size, shape, function, knowledge that an umbrella is used to keep yourself dry
	Knowledge of what/ knowing that		
Procedural model	Sequential and timing type of knowledge	Task action/goal relationship, and external influences on this relationship	Use of an umbrella
	Knowledge of how/ knowing how		
Strategic model	Information that is the basis of problem solving	Action plans to meet specific goals, knowledge of the context in which procedures should be implemented, actions to be taken if a proposed solution fails, and how to respond if necessary information is absent	Applied use of an umbrella not only for rain, but also sun, sandstorms etc.
	Knowledge of what and how and applied to the context		

The current paper does not specifically investigate the impact of different railway changes on team mental models, but looks into the implications of a changed system on the strength of the changes on operators' mental models. The following section focuses on the connection of the different conceptualizations of mental models with the different conceptualizations of innovations.

LINKING INNOVATIONS TO MENTAL MODELS

Our exploration of literature on innovation and mental models resulted in dimensions that described both concepts. To understand how different dimensions of innovations impact different dimensions of mental models we have to link them (see Table 5). Firstly, we see a striking resemblance between centrality and the mental model conceptualization. As the level of centrality relates directly to the stability of a system element and the hierarchical planning model of a railway system looks at first glance quite similar to how mental models of equipment, task, team interaction and team are portrayed, we link these two dimensions. Thus, we expect that changes that have high centrality will have more problems in changing the mental models of game players in a game. Elements found higher up the hierarchical ladder such a technical system elements are more stable and mental models are built upon this premise. Accordingly, any change focused on these elements will lead game players to have more problems in quickly updating this part of their mental model. The second proposition is that any type of change, independent of its centrality can have changes in declarative, procedural and/or strategic mental models. A railway switch can look and function differently (change in declarative knowledge), it could also change in its way of operation (procedural knowledge) and its role within the railway system (strategic knowledge). Thirdly, we see similarities between product-related changes to equipment mental model type, between social changes and the team mental model type and between institutional changes and the task and team interaction mental model types. Finally, whether a change is focused on the product, social or institutional level will directly relate to that part of the mental model that either focuses on declarative, procedural or strategic knowledge. Product-related changes will have higher impacts on declarative knowledge as tasks and procedures remain the same but the working of the system with which game players interact will change. Social-related changes such as new roles and responsibilities will change partly the working of the system and partly the procedures that need to be applied. Institutional changes specifically address how operators need to interact with the system by describing operating rules, interaction rules and new incentive structures. Whereas the technical system works the same, knowledge on how to steer, operate and control this system has to be updated. We therefore expect for these kinds of changes that procedural mental models need to be updated. Proposition 1 and 4 will be assessed through four case studies, while propositions 2 and 3 will be herein assumed.

Table 5: Propositions on the relation between the innovation and mental model concept

Innovation concept	**Mental model concept**	**Proposition**
Centrality	Equipment-Task-Team interaction -Team	1. High centrality of a change invokes changes in the equipment mental model type, while a low centrality of a change will invoke changes in the team mental model type
Centrality	Declarative-Procedural-Strategic	2. Declarative, procedural and strategic mental model types are not related to the degree of centrality and therefore equally relevant to all types of changes
P-S-I	Equipment-Task-Team interaction -Team	3. Product-related changes can be related to the equipment mental model type, social changes can be related to the team mental model type. Institutional changes can be related to task and team interaction mental model types
P-S-I	Declarative-Procedural-Strategic	4. Product-related changes will mostly need changes in declarative mental models, Institutional changes will need changes in procedural changes whereas social changes need updates in both declarative and procedural changes

GAMING SIMULATIONS IN THE RAILWAY SECTOR

In essence, gaming simulations are operating models of reality (Ryan, 2000) to which gaming elements are added (Meijer, 2012). Thus our exercises are first and foremost a simulation of reality. Adding gaming elements is only

https://openaccess.cms-conferences.org/#/publications/book/978-1-4951-2099-2

needed to make the simulation playable and it allows game players to interact with the model. Different from computer simulations as closed systems were human behavior is translated into algorithms, we leave this part of the model open. By doing so, we account for the sociotechnicality and complexity of railway systems: technical and human elements together determine how the system behaves and their interdependence is two-way and nonlinear.

Table 6: Characteristics of four research gaming simulations

	NAU	**1st phase**	**OV-SAAL**	**Leeds**
Purpose	Testing improvements in resilience and robustness when introducing a new control concept for Utrecht Central station	Studying the impact of current and alternative procedures for the improvement of speed and settlement of railway infrastructure disruption mitigation	Testing four railway infrastructure changes on their impact with different disruptions on the trajectory Schiphol-Amsterdam-Almere-Lelystad (SAAL)	Testing redefined roles of train traffic operators with two disruptions around the station of Leeds
Type of change	Railway switches, roles, procedures	Procedures	Railway tracks	Roles
Scenarios	Two: 1. “old” way, 2. new mechanism	Two: 1. current procedure, 2. alternative procedure	Four: adding infrastructure in: 1. Amsterdam south axis 2. Weesp, 3. Duivendrecht, 4. Almere	Two: 1. new roles with light disruption, 2. different light disruption
Simulated world	Detailed infrastructure for Utrecht Central; detailed current timetabling; face-to-face communication lines; stylized planning tools	Railway system between Amsterdam Central and Alkmaar Station. Facilitated information system (partial automation and function of computer information system operated by facilitators), colocation by room separation	Detailed infrastructure for certain parts of the SAAL trajectory; detailed track occupancy on stations; train face-to-face communication lines; stylized planning tools	Detailed infrastructure for Leeds Station; detailed current timetabling; face-to-face communication lines; stylized planning tools
# of partici-pants and roles	Nine: train traffic controllers (3), regional network controller (1), driver re-scheduling (1), rolling-stock re-scheduling (1), platform coordinator (1), network controller (1), service controller (1)	Twelve: train traffic controller (4), regional network controller (1), national network controller (1), regional passenger traffic monitor (1), regional passenger traffic junction coordinator (1), regional passenger traffic material and passenger coordinator (1), national passenger traffic controller (1), passenger information dispatcher (2)	Nine: train traffic controller (1), regional network controllers (3), national network controller (1), regional passenger traffic monitor (1), regional passenger traffic junction coordinator (1), regional passenger coordinator (1), national passenger traffic control center (1),	Six: dispatcher (1), incident manager (1), service delivery planner (1), timetable planners (1), service and infrastructure manager (1), information controller (1)
Type of role and objectives	Execution of tasks – similar as to in their daily work, only in the second s	Execution of tasks – similar as to in their daily work	Senior roles, execution of tasks – similar as to in their daily work	Execution of tasks – as to their new roles
Time model	Continuous	Continuous	Step-wise	Continuous

Gaming simulation has two big advantages over a direct trail-and-error approach in innovation. Firstly, system elements that are often spatially and temporally dispersed in real-life are brought closer to each other. In this way, innovation managers are better able to understand what causal patterns play a role in the system. In our gaming simulation exercises we bring operation centers together or we allow time to be compressed or slow down to better uncover the processes underlying the phenomena of interest. Secondly, since it is a controlled environment, parameters that are deemed highly unchangeable in real life can be manipulated in the experimental setup of a session. So adding a railway switch to the layout of a station might cost 100.000 euro but inside a game it is a matter of erasing a black line connecting two railway tracks. As such, gaming simulation allows alternative solutions to be tested in a safe environment (Kriz, 2003; Meijer, 2012).

In 2009, ProRail started to explore the value of using gaming simulation to test out innovations. Since the start of the Railway Gaming Suite (RGS), many different gaming simulations have been applied. This paper focuses on tabletop gaming simulations that involved multiple operating roles from the railway domain. Because of the need to quickly build gaming simulations that focused on specific parts of the railway network, we chose to use low-tech means to build our simulation environments. We found that as long as the processes are realistic, such as using real operating procedures, a real timetable and real infrastructure layout, operators soon recognized the system as theirs and high levels of immersion were reached: crucial for having a valid game (Meijer, 2012).

In table 6, we summarized the different gaming simulations that are used as case studies in this paper. Four instances of the use of gaming simulation are applied to relate different types of innovations to different types of mental models. Three of them were used for testing innovation in the Dutch network; Network Rail, the UK equivalent of ProRail, applied another one on testing new traffic control roles in the Leeds-Bradford area. Three gaming simulations each covered one of the three focal points of sociotechnical innovations: product, social and institutional elements, whereas a fourth one focused on all three simultaneously. Since the purpose of the paper is to investigate how valid mental models of operators can be ensured in a gaming simulation of a system that has changed parameters (e.g. new procedures or new railway tracks), it is necessary to look into methods how to assess the validity of operators' mental models. Changed mental models can be linked to the learning domain in terms of changed knowledge. For the current case studies, observational methods are used to assess the implications of the changed nature of the mental models during the gaming simulation. We selected the following indicators to identify an unstable, and thus not fully developed, mental model: discussion between participants, inability to make use of, mistakes related to, ambiguity and questions about the introduced changes. Additionally, during the debriefing, operators were asked about their experiences of the game.

RESEARCHING RAILWAY INNOVATIONS

To explore the two propositions as mentioned in Table 5, the four gaming simulations have first been assessed in accordance to their success, in which in the current approach success is indicated by the degree of change in the mental model, i.e. the amount of learning that is needed before the mental model more or less stabilized. Based on the observations, we see that the NAU game was most successful in that operators indicated and portrayed confidence in their task execution with the introduced changes in the railway system. This was followed by the OV-SAAL game, where operators did not show much struggles in the application of the new knowledge, but forgot to make use of some of the new options that were introduced. Operators that participated in the Leeds game showed some confidence, but also uncertainty in their newly introduced roles. In the fourth game, the 1st phase game, operators were troubled with the newly introduced changes in railway system and indicated doubts about their correct understanding of the procedure, which made it the least successful game based on these indicators.

In Table 7, the changed elements of the railway system in the different gaming simulations are mapped in accordance to the ETTT conceptualization of mental models. Adaption of operators' mental models requires a change in one or more of the four mental model types, i.e. equipment, task, team interaction and team. Although the design choices of the game may have an impact on the equipment model (stylized analogue interface system), the overview focuses on the changed elements. As such, railway switches are seen as part of the equipment mental model, whereas a change in the roles affect the assignment of the task (task model), what the roles and responsibilities are (team interaction model), but also the knowledge about the team members' preferences, abilities etc. (team model). Similarly, procedures impact the task as well as the team interaction model.

Table 7: Tested changes in the four multi-actor tabletop gaming simulations and relation with the task-team conceptualization of mental model

	Equipment	Task	Team interaction	Team
NAU	Railway switches (P)	Roles (S), procedure (I)	Roles (S), procedure (I)	Roles (S)
OV-SAAL	Railway tracks (P)			
Leeds		Roles (S)	Roles (S)	Roles (S)
1st phase		Procedure (I)	Procedure (I)	

Following these observations, we identify that some types of changes have a widespread influence and affect multiple types of mental models, e.g. a change in the roles and responsibilities indirectly also affect the task responsibilities related to the roles and the implications on knowledge of team members' abilities for that new role. Our first working proposition in this paper is that changes of high centrality will invoke stronger changes in the mental models of game players. Relating these classifications to the success of the gaming simulations, we observe that a changed procedure is more difficult for operators to adapt to (1st phase game) than in the case of a sole change to the infrastructure (OV-SAAL) and roles (Leeds). This is opposed to our proposition, in which changes to components with a high centrality (building blocks of railway system, i.e. infrastructure) are expected to have a stronger impact on the change of the overall mental model. We have one possible explanation for this discrepancy: firstly, operators noted that in their daily work they often interact with railway systems that vary to great extent due to for instance signaling and railway switch malfunctioning. Our assumption that exactly these elements are inert therefore does not hold. On the other hand, we stated that centrality is not just a matter of where on the hierarchical ladder the element is placed. OV-SAAL and NAU both experimented with highly inert elements but where for NAU subsequently many downstream elements had to be changed as well, OV-SAAL could be played with the same timetable, the same operators and the same procedures. So in the latter example, operators interacted with a model using their current mental model, explaining potentially why they forgot to use additional means of solving disruptions that the design change gave them. We see that removing nodes and links from the networks tend to cause more downstream cascade effects and hence have a higher centrality than adding nodes or links.

In Table 8, an overview is provided of the different types of changes and relevant changes relating to DPS mental model conceptualization. In all gaming simulations, declarative knowledge needed to be adapted for the product-related (infrastructure) and social-related changes (roles), e.g. the location and amount of railway switches was changed in the NAU game, but not the function or ways of operating switches, while procedural mental models were mostly relevant for the changes in roles and institutional changes (procedures). Strategic mental models were influenced due to their relevance for the scenarios in the gaming simulations. In accordance with the proposition, product-related changes affected declarative mental models, institutional changes affected procedural mental models, and social changes needed updates in both declarative and procedural mental models. Additionally, we can deduct that the extent that a certain amount for learning is necessary might be related to what type of knowledge structure is changed and in which combination; in the NAU and Leeds games, both declarative and procedural knowledge were impacted, while in the OV-SAAL game and 1st game either declarative or procedural knowledge was changed. In the 1st phase game, procedures were more difficult to learn than the availability and impact of additional railway tracks. Thus, this leads to another proposition that changing solely procedural knowledge might be more difficult than when declarative and procedural knowledge are changed and we can subsequently state that experimenting with design changes that focus on both product, social and institutional elements of the system causes less problems for game player's mental models than experimenting solely with design changes in institutional elements.

Table 8: Tested changes in the four multi-actor tabletop gaming simulations and relation with the mental models operationalized as different knowledge structures

	Type of change	Declarative	Procedural	Strategic
NAU	Railway switches (P), roles (S), procedure (I)	Yes, removal of railway switches (P), redefined roles (S)	Yes, redefined roles (S) and new procedure (I)	Yes, impact of the changed availability of switches, redefined roles and new procedures in the scenarios
OV-SAAL	Railway tracks (P)	Yes, availability of tracks (P)		Yes, impact of the changed availability of switches in the scenarios
Leeds	Roles (S)	Yes, redefined roles (S)	Yes, redefined roles (S)	Yes, impact of the redefined roles in the scenarios
1st phase	Procedure (I)		Yes, new procedure (I)	Yes, impact of the new procedures in the scenarios

DISCUSSION AND CONCLUSION

Reviewing the four gaming simulations and looking at how operators coped with changes to the system, we see that changes in operating routines (procedures) were difficult and in the gaming simulation session where this type of change was solely tested, we found a longer period of learning. However, changing procedures might have a less invasive impact on the mental model development, when the change is introduced as well with declarative knowledge. Parallel implications can be drawn between this proposition and findings in other studies on learning in which is stated that learning to operate a new device (procedural) is improved by teaching the inner working of the device (declarative) (e.g. Kieras and Bovair, 1984). Additionally, a minimal amount of time is needed for players to familiarize themselves with the gaming simulation design, and the planned changes, but the amount and effort of training might be related to the use of certain types of changes. Thus, based on the case studies, we preliminary conclude that different types of innovations, i.e. infrastructure and procedural innovations, impact mental models differently and therefore pose different requirements for the design of a valid gaming simulation experiment. Experimental studies are necessary to validate the propositions, in which future research could look into testing the propositions through quantitative mental model measurements and a more structured approach for the observation of the indicators.

Regarding the analysis of case studies, a number of limitations can be subscribed to the current approach: firstly, the strength of a change in the mental models has been measured holistically. As such, possible differences on the impact of the changes on a single operator have not been taken into account in the observations. Furthermore, only two of the four propositions have been assessed in the case studies, as the remaining propositions were assumed based on their definition. A more in-depth literature review is needed to assess proposition two and three. Finally, the current analysis did not take into account what the impact was of the different gaming simulation design choices (e.g. analogue stylized interfaces for train traffic and network controllers) on changes in the mental model, e.g. especially with regards to the equipment model. However, it has been argued that the a reference system can be reduced to basic elements in the simulated system without putting high demands on the change of an game player's mental model (Dormans, 2011).

All in all, the current paper provides initial propositions for the design implications of a simulated environment that uses iconic representations with the purpose to validly test new innovations in the railway domain.

ACKNOWLEDGMENTS

This research was funded through the Railway Gaming Suite program, a joint project by ProRail and Delft University of Technology.

REFERENCES

Abernathy, W. J., & Utterback, J. M. (1978). Patterns of industrial innovation. *Technology Review*. 80(7), 254-28.

Bonen, Z. (1981). Evolutionary behavior of complex sociotechnical systems. *Research Policy*, 10(1), 26-44.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In J. John Castellan, N. (Ed.), *Individual and Group Decision Making: Current Issues*. Hillsdale, New Jersey: Lawrence Erlbaum

Carayon, P. (2006). Human factors of complex sociotechnical systems. *Applied Ergonomics*, 37(4), 525-535.

Dormans, J. (2011). Beyond iconic simulation. *Simulation & Gaming*, 42(5), 610–631.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 32(2), 97-101.

Endsley, M.R. (2000). Situation models: An avenue to the modeling of mental models, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44(1), 61-64.

Frenken, K. (2006). *Innovation, Evolution and Complexity Theory*, Cheltenham: Edward Elgar.

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research Policy*, 33(6), 897-920.

Goverde, R. M. (2005). *Punctuality of Railway Operations and Timetable Stability Analysis*. Delft: Netherlands TRAIL Research School.

Kauffman, S. A. (1993). *The Origins of Self-order*, New York, NY: Oxford University Press.

Kemp, R., Schot, J., & Hoogma, R. (1998). Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management. *Technology Analysis & Strategic Management*, 10(2), 175-198.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, 8, 255-273.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor?. *Journal of Management*, *20*(2), 403-437.

Kriz, W. C. (2003). Creating effective learning environments and learning organizations through gaming simulation design. *Simulation & Gaming*, 34(4), 495-511.

Lim, B. C., & Klein, K. J. (2006). Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior*, 27(4), 403-418.

Lo, J. C., Van Den Hoogen, J. & Meijer, S. A. (2014), Using Gaming Simulation Experiments to Test Railway Innovations: Implications for Validity, in: R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill & M. E. Kuhl (eds.), *Proceedings of the 2013 Winter Simulation Conference*, IEEE.

Markard, J. (2011). Transformation of infrastructures: sector characteristics and implications for fundamental change. *Journal of Infrastructure Systems*, 17(3), 107-117.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, 85(2), 273.

Meijer, S. A. (2012). Gaming Simulations For Railways: Lessons Learned From Modeling Six Games For The Dutch Infrastructure Management, in: Perpinya, X. (ed.), *Infrastructure Design, Signaling and Security in Railway*, Croatia: IntechOpen.

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010), Metaphor no more: A 15-year review of the team mental model construct, *Journal of Management*, 36(4), 876-910.

Nelson, R. R., & Winter, S. G. (1982) *An Evolutionary Theory of Economic Change*. Cambridge: Belknap.

Ryan, T. (2000). The role of simulation gaming in policy-making. *Systems Research and Behavioral Science*, 17(4), 359-364.

Salas, E., Stout, R. J., & Cannon-Bowers, J. A. (1994). The role of shared mental models in developing shared situational awareness. In R. D. Gilson, D. J. Garland & J. M. Koonce (Eds.), *Situational Awareness in Complex Systems*, 297-304. Daytona Beach, Florida: Embry-Riddle Aeronautical University Press.

Saviotti, P. P. & Metcalfe J. S. (1984). A theoretical approach to the construction of technological output indicators, *Research Policy*, 13(3), 141-151.

Simon, H. A. (1962). The Architecture of Complexity, *Proceedings of the American Philosophical Society*, 106(6), 467-482.

Subrahmanian, E., Reich, Y. & Meijer, S. A. (forthcoming). The game between design and institutions: the PSI framework.

Trist, E. (1981). The evolution of socio-technical systems. *Occasional Paper*, 2.

Tushman, M. L., & Romanelli, E. (1985). Organizational evolution: A metamorphosis model of convergence and reorientation. *Research in Organizational Behavior*, 7.

Van den Hoogen, J. & Meijer, S. A. (2014). Gaming Technology Landscapes, In: S.A. Meijer, R. Smeds (eds), *Frontiers in Gaming Simulation*, Springer Lecture Notes in Computer Science, 8264, 153-160.

Wilson, J. R., Farrington-Darby, T., Cox, G., Bye, R., & Hockey, G. R. J. (2007). The railway as a socio-technical system: human factors at the heart of successful rail engineering. *Proceedings of the Institution of Mechanical Engineers, Part F:*

Journal of Rail and Rapid Transit, 221(1), 101-115.