

Design methodology of analytical games for knowledge acquisition

Francesca de Rosa¹ and Alessandro De Gloria²

¹NATO STO Centre for Maritime Research and Experimentation,
francesca.derosa@cmre.nato.int

²Electrical, Electronics and Telecommunication Engineering and Naval Architecture
Department, University of Genoa, adg@elios.unige.it

Abstract

Analytical games explore a problem or a domain with a research purpose. Considerable research is ongoing to investigate improvements to analytical game design, execution and exploitation. Moreover, the fast-paced technological developments in many fields, such as artificial intelligence and virtual reality, make it even more compelling to account for the advantages and limitations of these new capabilities. In game design, the use of digital means is often regarded as a mere technical factor that relates to the platform selection, facilitator support and data recording processes. In this work a shift in perspective is proposed, to move from technology-oriented design selection criteria towards a broader assessment of the design choices. In fact, the introduction of technology (i.e., automation and autonomy) will not lead to a substitution of tasks, but will intrinsically change the game environment. This work introduces a framework to provide a structured guidance on the aspects to be factored in the different design phases of an analytical game, including the potential impact of the adoption of automation and autonomy. The proposed approach is based on previous research in the field of simulation-based serious gaming, model-driven engineering and human factors engineering. The framework is applied to a Knowledge Acquisition Analytical Game as a case study.

Keywords: Serious game, Analytical game, Knowledge acquisition, Design framework, Model-driven engineering, Human factors engineering

1 Introduction

The first definition of serious games appeared in [1] and referred to gaming for educational improvement. Along the years, several other definitions have been proposed [2] that identify them as games which main purpose is not entertainment. Although most research within the field of serious gaming has focused on games for learning [3], they can also serve other purposes. For example, in the Gameplay/Purpose/Scope (G/P/S) taxonomy for serious games [4] the purpose dimension allows classifying games on the basis of their function, namely message broadcasting, training or data exchange. Data exchange games have the specific purpose of “collecting information from [...] players” [4]. Some authors generically refer to games designed with the purpose of collecting data [4], while others distinguish games on the basis of the specific aim of collecting such data (e.g., for experimental, research or operational purposes) [5]. Following [6], we will use the term analytical games to indicate games with an experimental or research purpose. Therefore, analytical games can be included in the category



of data exchange serious games. Like other serious games, analytical games use simulation as a means to immerse the player into a simulated environment [7] that imitates at different levels of abstraction the reference system. The participants have a specific role and need to assess the situation, uncover the unknown and apply their knowledge and skills to solve the problems posed by the situation. The study of the processes (e.g. cognitive, behavioural and physical) involved into the game is the main aim of analytical games. A specific category of analytical games is the Knowledge Acquisition Analytical Games (K2AGs) [8], which support the knowledge engineering and knowledge acquisition (KA) for decision support and cognitive systems design. Knowledge acquisition techniques aim at collecting, structuring and organising expert knowledge. These techniques suffer from several drawbacks, such as time inefficiency, biases and prevented access to tacit knowledge (e.g., [9]). To overcome some of these limitations the knowledge engineering domain is exploring the use of games (e.g., [10]). Specifically, K2AGs investigate decision-making with a focus on information processing strategies, uncertain reasoning and the decision-making cycle (i.e., situational assessment, situational awareness and decision [11]). Game design has been widely discussed in game science research, but most of the work has focused on the design of the game artifact per se [12]. Recently, research has recognised the need to consider different levels of design, namely the design of the game artifact and the design of the game in relation to the socio-technical systems issues it tries to inform or support [12]. In order to provide a structured guidance to analytical game designers and users, in this paper we present a design framework, which builds on the approach proposed by [13] in the context of simulation-based serious gaming and complements the work on K2AG design formalisation, which focuses on the game artifact [14]. Specifically, the proposed solution aims at better formalising the technological and non-technological game design decisions, including aspects that go beyond the mere game platform. In fact, we propose to explicitly account for aspects related to automation and autonomy (A&A). Automation refers to “technique[s], methods[s] or system[s] of operating or controlling a process by [...] automatic means, as by electronic devices, reducing human intervention to a minimum” [15], while autonomy refer to the ability of machines (hardware and software) to perform independently under significant uncertain conditions for extended periods without external intervention [16]. A&A represent capabilities that might provide interesting opportunities and might significantly support analytical gaming. Examples are analytical games where players are not co-located in space (i.e., distributed games) and time, games with very complicated adjudication needs or games where non-playing characters (e.g., troops or adversaries) might be useful. In the last decades, the benefits of automation and later autonomy have been explored and clearly discussed (i.e., [17]). However, their use in gaming could negatively impact the quality of the collected data if human-system integration factors are not correctly addressed during the design phase. In fact, it has also been demonstrated how their introduction implies a trade-off between performances, workload and situational awareness (i.e., [18]). Introducing A&A in analytical games will intrinsically change the game environments. This might impact the way humans (e.g., players and facilitators) interact with such environments and the decision-making process under analysis. In order to build a common understanding of the issue and to provide initial guidance on the aspects to be factored in the preliminary design choices of analytical games, an automation and autonomy adoption assessment approach is included in the overall framework. Understanding such aspects will strongly contribute to the scientific rigour of the employed serious game, by fostering responsible and appropriate use of available technologies to further exploit analytical gaming capabilities. In fact, this will help make informed design decisions. The proposed framework builds on several decades of research conducted in the field of human factor and cognitive engineering. It is expected that this work will serve not only as a guideline for design and possibly for further research by the serious gaming community, but also as a bridge towards

other scientific domains.

The remainder of the paper is organised as follows: Section 2 provides some background on the use of serious games in the engineering practice and on game design formalisation; Section 3 discusses about A&A and serious gaming; Section 4 presents the proposed design cycle and summarises research in human factors and cognitive engineering relevant to analytical games; Section 5 shows a use case application of the design framework to K2AGs; and Section 6 reports the conclusions.

2 Related work

2.1 Serious games, analytical games and engineering

Analytical games can be used to experiment with complex systems in order to understand and predict their behaviours [7, 19, 20]. In the past years, different fields of engineering have started looking at games as supporting design tools. A literature review of serious games used in engineering research [3] reports how these games have been used to enable the sharing of data between collaborating designers or between researchers and subject matter experts (e.g., [21–25]) with the final goal of exploring decision-making in the engineering context. Some serious games specifically focused on the understanding of human problem-solving strategies. For example, Foldit [26] explores such strategies to inform computational algorithm optimisation in the context of protein structure design. The results obtained have demonstrated the potential of human-derived strategies elicited through a serious game used in conjunction with computational algorithms. K2AGs are analytical games designed to support the knowledge acquisition task for decision support system design. Specifically, K2AGs make use of a mix of simple game mechanics, game items (i.e., knowledge cards providing information to players) and data recording approaches that allow to elicit human reasoning schemes and uncertainty handling strategies. While focusing on the cognitive processes of the player (e.g., information processing, belief updates, attitudes and decisions) the games are able not only to collect domain knowledge in a more general sense, but also data to be readily used in the design of advance artificial intelligence algorithms [8]. The experiments conducted have demonstrated that these games can serve as innovative KA techniques. In fact, K2AGs have proven to be very efficient and effective in terms of knowledge elicitation, specifically with respect to time reduction, experiment simplicity and ability to extract the required qualitative and quantitative knowledge. The experts did not only consider the experimentation methods engaging, but regarded the games as operationally relevant and clearly supporting the overall stated goal. Several instances of K2AGs have been successfully developed and deployed in the last years, such as the Risk Game [27], the Reliability Game [28, 29], the MARitime Surveillance knowledge Acquisition (MARISA) Game [30] and the Maritime Unmanned System Trust (MUST) Game [31]). As a use case in this paper we consider the Reliability Game, which has been thoroughly analysed and was implemented both as board and computerised game. The Reliability Game has been developed with the purpose of experimenting with the impact of source factors (i.e., source type and source quality) on situational assessment and related situational awareness. The final goal of the game is to collect data to enrich the body of knowledge regarding human reasoning strategies handling partial reliable sources. This knowledge is relevant in order to inform the design of multi-source fusion algorithms. Specifically, it is a single-player game in which the participant is presented with the anomalous behaviour of a vessel and is asked to assess what is happening out of a set of collectively exhaustive and mutually exclusive hypotheses. The information and possible meta-data on the source of information are provided through cards that have to be positioned on the game board. These positions reflect the players beliefs regarding the current situation. The data collected through

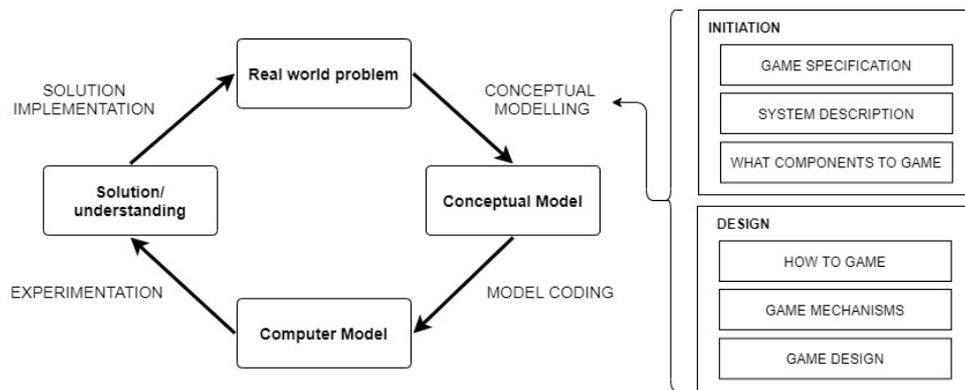


Figure 1: Design framework for simulation adapted to simulation-based serious gaming [13]

this game has been used to train a Bayesian network replicating the expert reasoning during the game [29]. One of the objectives was to learn how the players handled source reliability, which was a latent mental construct in this game. Another successful example of K2AG is the MARISA Game, which explores mental constructs related to illegal activities. The data collected contributed to the design of a multi-source dynamic Bayesian network for behavioural analysis, which has been successfully evaluated in operational trials by military operators.

2.2 Game design

Several existing frameworks, models and methodologies related to serious game design have been proposed. They often present different, yet complementary, perspectives on game design. For example, the Mechanics, Dynamics and Aesthetics (MDA) framework [32], proposes three perspectives that look at the actual implementation (mechanics), the design goal (dynamics) and the player experience (aesthetics). The Hierarchical Activity-Based Scenario (HABS) framework [33], presents a way to factor human-computer interaction (HCI) aspects into the design to drive scenario and narrative definition. While the human component of the game ecosystem is considered in many of these frameworks, the focus is mainly on player experience, immersion, the desirable emotional responses or on the desired learning outcome in the case of serious games for learning (e.g., [34]).

One of the main design frameworks for simulation games is the one proposed in [35], which introduces an overall design concept based on the following steps: initialisation, design, construction and use. The proposed model has been further extended by [36]. Here the design process is described as a set of stages: Stage I - setting objectives and parameters; Stage II - model development; Stage III - decision about representation; Stage IV - Construction and modification; and Stage V - preparation for use by others. Focusing on simulation-based serious games, in [13] this process has been mapped to the simulation design process proposed by [37]. This allows accounting for the iterative nature of the development of simulation-based projects. The resulting framework is reported in Figure 1. The main processes in this case are identified as the conceptual modelling, the model coding, the experimentation and the implementation. From the figure it is possible to observe how the initialisation and design processes proposed in [35] need to feed the conceptual modelling. Van der Zee et al. [38] further refined the approach providing a detailed guidance on how to approach conceptual modelling in simulation-based serious games with a learning objective.

Following [37], a conceptual model can be defined as “a non-software specific description of the computer simulation model (that will be, is or has been developed) describing the objectives, inputs, outputs, content, assumptions and simplifications of the model”. The term

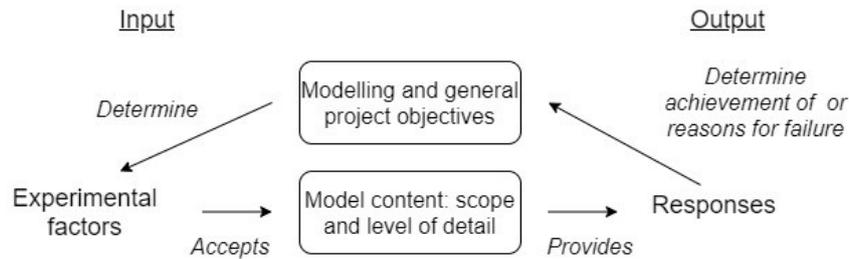


Figure 2: The conceptual model within a simulation project life-cycle as proposed in [37]

objective does not only refer to the overall purpose of the model, but also to the project objectives, which include the time-scale, nature of the model and the model use. The inputs, also defined as experimental factors, are the factors that are altered within the simulation to obtain the desired observations. The model content, instead, include the scope of the model, namely its boundaries, and the level of detail. Finally, the outputs serve a double scope: (i) understanding if the modelling objectives are met and (ii) in case of a negative outcome the reason for not meeting them (Figure 2). It has to be considered that, although the conceptual model is software independent, it might have to be adjusted because of the choices made in the subsequent steps of the design process. Van der Zee et al. [38] reports a list of all the activities involved in the conceptual modelling. These activities include both elements related to the problem at hand (e.g., the requirements of the project in terms of learning objectives and the understanding of the learning environment) and the elements that relate to the serious game (e.g., decision about the platform or the model use with respect to visualisation, player interaction and responsiveness).

In [3] the use of van der Zee's design framework is proposed to guide the design of serious games to study human decision-making in engineering. However, no specific guidance on how to adapt the generic framework to the specific context is available. In the work presented in this paper we aim at providing a specific guidance to designers of analytical games. The attention is drawn on the need to ensure to design the game following an experimental (or quasi-experimental) approach, controlling the intervening factors that might impact the players' reasoning schemes and decisions, which needs to be captured through the game. A different management process is proposed for analytical wargames (e.g., [39, 40]). This management life-cycle includes the following phases: define, design, plan support, develop, rehearse, execute and communicate results. These phases include several steps to be systematically followed in order to correctly manage the creation and deployment of the game. Differently than other design frameworks, this process is strongly focused on the actors involved along the cycle (i.e., sponsor, designer, facilitator, data collector and analyst). However, the focal point is on their role and the tasks they should perform. In the method proposed in this paper, instead, the focus is on the game as well as its interaction with the human component. In fact, we highlight how design choices should be cognisant of such interaction (Section 3). Therefore, we propose to explicitly account for human-system integration factors.

3 Automation, autonomy and gaming

3.1 Game platforms

Several typologies and taxonomies of games based on different aspects have been proposed [12]. One of these is the game platform, which is a fundamental factor from a design perspective. The game platform can be defined as the support through which the game is deployed

and played. For example, [41] presents a distinction between digital and non-digital games, where digital games include the ones requiring a computer, smartphone, tablet or other tools [42]. The non-digital ones, instead, make use of boards, cards, pencils and papers [43]. This classification proposes only two alternatives, namely full digitalisation or none. A different perspective considers these as two extremes of a continuum, where in-between states are possible [44]. Specifically, the distinction between manual (or analogue), computer-assisted and computerised games is made.

A relevant typology is the classification framework for computerised business simulation gaming based on the concepts of control over the task and interaction [45]. More specifically, following this framework a game can be: (i) Computer-based; (ii) Computer-controlled; or (iii) Computer-assisted. Computer-based games present a high computer-participant interaction and high participant control. Computer-controlled games are characterised by a high participant-participant interaction and high computer control. Finally, computer-assisted games exhibit a high participant-participant interaction and high participant control. Control over the task and interaction are important elements with respect to the design choices (i.e., assessing the potential impact). Although this work was proposed for business simulation games, it can be extended to other simulation games as well.

In analytical games the use of technology becomes a means through which facilitation is supported and elements of the game are managed. The final aim is to facilitate, collect, analyse and display relevant aspects of human decision-making processes. As highlighted in [46] we can distinguish between decision tools (“capable of moving, displaying, synthesizing, and manipulating information”), scenario tools (“capable of representing the effects of the decisions and events”), adjudication tools (“capable of [...] the management of results”) and synthesis tools (“which facilitate visualization and collaboration”).

Although modern technologies offer a wide range of gaming solutions, ranging from low cost and low technology to highly advanced ones, in most cases organisations resort to a mixed approach, made of manual games, software-heavy tools and fully computerised approaches [46]. The reason for which organisations still adopt manual games, regardless of the level of resources available, is that these have significant pragmatic advantages, such as cost, flexibility, adaptability, adjudication transparency, large libraries of game mechanics and ease of use when dealing with classified topics [46]. The same game could be implemented on one or more platforms, depending on the analytical objectives and requirements. Several factors influence this choice. For example, the design time, the cost, the need to observe the game execution to ensure that the game is a valid experimentation method and the need to gain insights into the players reasoning from face-to-face interaction. On the other hand, the need to enrich the data sample (i.e., number and location of players to involve) could lead to the selection of digital games. The concept of game platform is strongly related to the automation and autonomous technologies that can be adopted within the game.

3.2 *Defining the concepts*

There is confusion around the terms automation and autonomy, which often are used interchangeably. The confusion might arise from the fact that often the concepts are presented as part of a continuum, without clear definitions (e.g., [47, 48]). From the perspective of progressive development stages, autonomy can be interpreted as highly advanced automation, capable of adapting to a wider set of conditions [49, 50]. Although some of the implications of the adoption of automated and autonomous technologies might be similar (Section 3.3 and Section 3.4), they actually imply technical differences and a paradigm shift. Understanding such difference is important in order to make informed design decisions.

More in detail, automation can be understood as the ability of the system to perform a

specific pre-defined task, leading to a deterministic result [51]. Automated systems are not able to operate in situations and scenarios that are unexpected and not fully codified. On the contrary, autonomy can be understood as the ability of the system to independently perform tasks, even in not fully anticipated scenarios. Autonomous systems employ technologies that allow learning, reasoning and self-adaptation, possibly resulting in non-deterministic results [51]. Full-autonomy at system level has not yet been achieved and in most cases even sub-systems act at some level of semi-autonomy. In fact, these systems still lack the ability to autonomously operate in each unexpected scenario or condition.

3.3 Automation

When automation is discussed, one key element is represented by the concept of LOA (i.e., [52–55]). Many taxonomies of levels of automation (LOA) have been proposed. An important aspect that emerges from the different LOA taxonomies is the distinction between the concept of mechanisation and computerisation. Mechanisation refers primarily to the automation of physical tasks [56] and material aspects, while computerisation links to the automation of information handling and control. This distinction was born and is easily understandable in the manufacturing domain, while in gaming it becomes less evident. In fact, here also the physical tasks and material aspects are moved to digital tools. The use of the term digital game, instead of computerised game, should overcome potential confusion with respect to mechanisation and computerisation as defined in other scientific domains.

The different LOA taxonomies are built around different perspectives [56]. Some refer to the decision-making cycle steps and the roles within this cycle (i.e., monitoring, option generation, selecting and implementing) [52, 57]. Others refer to the type of system output (i.e. notification, suggestion or action) [58]. LOAs can also be characterised along the decision-making stages by different degrees of automation and different combinations of human and system inputs (i.e., [57]). An important taxonomy to report is the one proposed in [54], where automation is analysed with respect to information processing stages (i.e., sensory processing, perception/working memory, decision making and response selection) [59]. Here the LOA is not unique, but rather articulated along four main dimensions, namely information acquisition, information analysis, decision selection and action implementation. Each LOA dimension corresponds to one of the information processing stages. Specifically: (i) information acquisition automation: corresponds to automation in support of human sensory processes; (ii) information analysis automation: corresponds to automation in support to working memory and inferential processes (i.e., information fusion, predictions); (iii) decision automation: corresponds to automation of the selection of decisions and actions to be performed; and (iv) action automation: corresponds to actual implementation of the selected action.

These dimensions are similar to the ones proposed in [57], however, they focus more on information processing. Moreover, the LOA is not explicitly defined as a combinations of levels along these dimensions. In fact, no specific scale of LOA is proposed. The reasons behind this choice are to be found in the difficulty to map each specific technology in a specific level and to clearly define the benefits and limits of each LOA. Although many criticisms to the use of LOA exist (e.g., empirical evidence for LOA predictions) [60], we align with the view expressed in [50, 61] that they should be used as a means to guide the discussions around specific aspects of automation design. These aspects include, but are not limited to, the impact on performance, situational awareness and workload. We also acknowledges the relevance of the perspective presented by [62], which besides the understanding of the potential tasks allocation between humans and machines and how to achieve it, underlines the importance of understanding the connections between elements of the system that might be automated.

3.4 *Autonomy and artificial intelligence*

It appears that the ability of systems to perform tasks autonomously will have a significant impact in terms of increased productivity and efficiency [63]. The current advances of the state-of-the-art of AI, which is a key enabling capability of autonomy, are key factors leading to the progressive shift from automation to autonomy. These advancements are based on the improvements in the fields of machine learning, big data and graphical processing units. However, the shift is in its early stages, as the above-mentioned advancements have been mainly observed in narrow AI [64], which is the field of AI that looks at handling a singular or limited task. In this phase, humans and the autonomous systems will continue to interact and the research on human-automation interaction will serve as foundation for the human-autonomy teaming one. However, the shift from automation to autonomy will also lead to a paradigm shift. In fact, the concept of operations of these systems will change from a one-way information flow between the system and the human, to a two-way information flow and consequent team-like behaviours (i.e. collaboration, coordination and joint planning) [50].

Artificial intelligence (AI) and autonomy are gaining increased attention in the gaming community. Most research on the use of AI in games has been performed in the field of commercial games and a considerable body of knowledge on game AI [65] is available. Although game industry has claimed for over a decade to have introduced AI in games, academic AI researchers have warned that the techniques actually adopted in most cases are fully scripted routines (i.e., automated) and based on old approaches. Nevertheless, steps towards the use of more sophisticated techniques (e.g., neural networks) are observed [66]. Moreover, game AI has started moving from its traditional focus on non player character (NPC) behavior, towards investigating its potential with respect to player-experience modeling, procedural-content generation and massive data mining on user behaviour [66]. A notable example is the multiplayer real-time strategy wargame powered by IBM Watson [67, 68].

One of the evaluation factors in successful commercial game AI has been identified as the level of integration of AI in the design of the game [69]. In fact, it is recognised that the lack of such integration might lead to “unjustifiable NPC behaviours” [66], which might negatively impact to the suspension of disbelief [70]. Suspension of disbelief, is the status on which a person assisting or acting into a narrative believes what is not there [71]. The importance of suspension of disbelief in games has been discussed in details by [72]. The unjustifiable NPC behaviours also link to concepts such as AI transparency and AI explainability, which are currently researched by the AI academic community. These are only some of the factors that might impact the human component of the analytical game ecosystem (Section 4.4). In order to provide guidance to the designers of analytical games the proposed design approach explicitly includes such aspects into the framework.

From an analytical game perspective full autonomy will never be achieved as they always implies a human-in-the-loop. In fact, following [73] at least one human player needs to be involved in a game. Moreover, as highlighted in the definitions of wargames [73] a game has to be seen not as the mere set of game items (i.e., game boards, maps, counters), game mechanics and rules, but rather as an ecosystem composed by the game artifact and a human component. This shift of perspective also aligns with the game research trend that recognises the need to consider different levels of game design (Section 1).

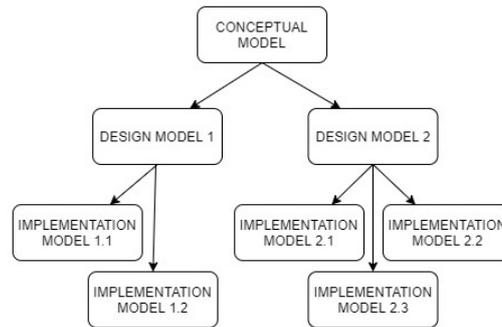


Figure 3: *The Model-Driven Engineering modelling process for modelling and simulation [74]*

4 Analytical games for knowledge acquisition design framework

4.1 Model-Driven Engineering and serious games

Van der Zee's framework for simulation-based serious games merges into the conceptual modelling task both the modelling activities that relate to the real world situation (i.e., learning environment and objectives) and the ones that relate to their link to the game, pushing some gaming related decisions to the first design phases. When it comes to analytical games it might be beneficial to adopt a step-wise approach that differentiates between the development of a model of the analytical problem to be investigated and the way in which it will be implemented within the game. To achieve this distinction we propose to apply Model-Driven Engineering (MDE) theoretical principles to the design of analytical games.

MDE differentiates between three different models. The first one, the domain model or conceptual model, is a solution-independent model of the simulation resulting from the analysis phase. The second one, namely the platform-independent design model, results from the design phase and describes a solution. Finally, the platform-specific implementation model is the output of the implementation phase. As described in [74], which applies MDE to simulation engineering, a one-to-many relation exists between (i) the conceptual model and the design models and (ii) each design model and the implementation models (Figure 3).

The interpretation of the meaning of conceptual model might vary between simulation engineering, software engineering and gaming. Nevertheless, the distinction of the three different kinds of models is very important. In fact, this perspective proposes to consider in a first instance a solution or computation independent model. In this phase the system design choices are not included and the model focus is on the application domain and experts' perspectives. Starting from the conceptual model several platform-independent computational solutions can be derived (i.e., design models). These can be implemented through several implementation models, corresponding to different platforms. The final selection will be guided by the appropriate design choices, such as the ones on architectural style, non-functional criteria, performances, adaptability and technological tools available.

4.2 The extended design framework

The MDE distinction between the three models is here applied to the analytical game design framework. Some work on the application of MDE to serious games for learning has been proposed (e.g., [75]), but the scope was to improve and speed up the development of the game

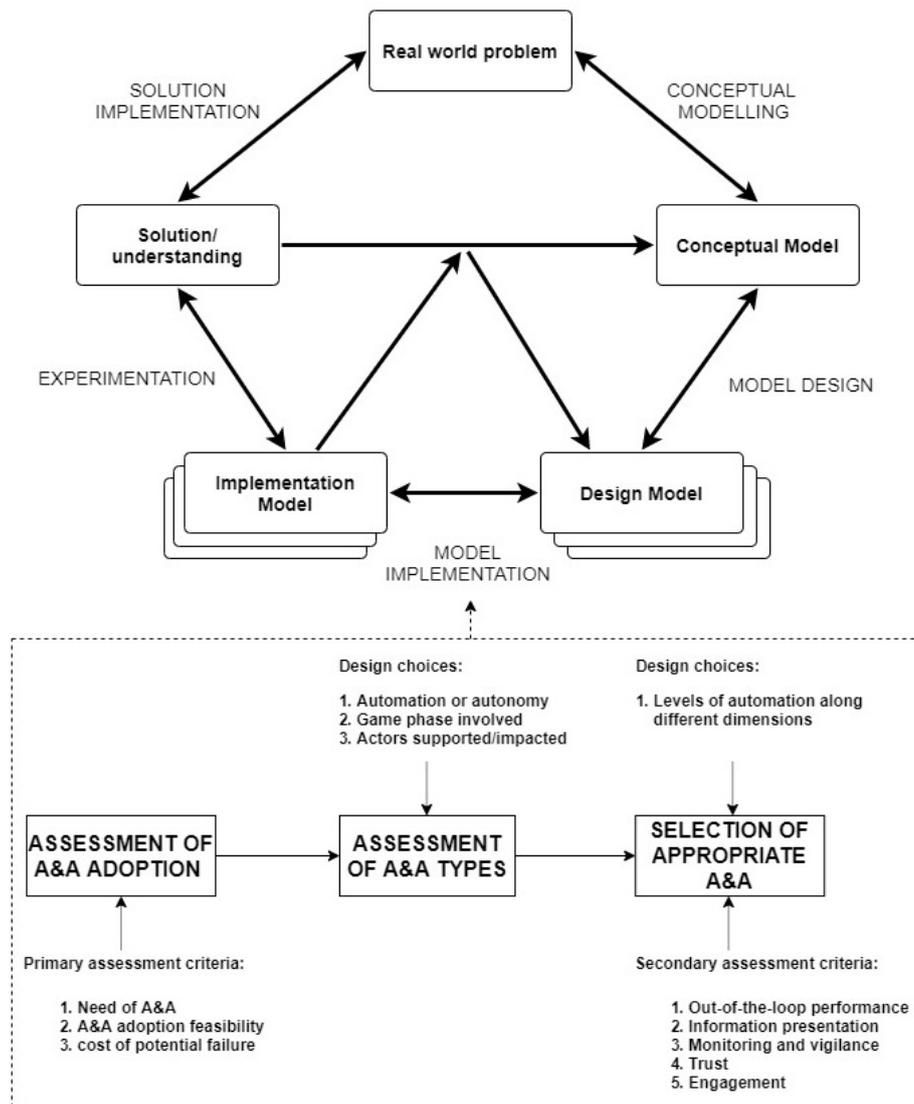


Figure 4: Overall analytical game design approach

from an artifact perspective. While this is not in contrast with this framework, here the introduction of MDE is a means to include a design step in which the conceptual model purely focuses on the cognitive factors (e.g., experimental factors) that need to be handled in the research. In fact, a sound understanding of their status and relations enables the assessment of the potential impact of the design choices in the subsequent design phases. Figure 4 illustrates the adaptation and extension of van der Zee’s design framework [13] to analytical games. We can observe how the design phase is actually composed by the three distinct steps of conceptual modelling, design modelling and the implementation modelling. Given that analytical games are simulation games that can be either digitalised or not, we will refer to an adaptation of the definition by [37], where the conceptual model of a analytical game is “a solution and platform independent description of the analytical problem, describing the objectives, inputs, outputs, content, assumptions and simplifications of the simulation”.

The conceptual modelling process encompasses the analysis of the research objectives and questions, in order to define the information to be collected, and the assessment on how such information should be analysed. More specifically, these aspects correspond to the definition

of the experimental factors [37]. For example, in the case of K2AGs, this is the phase in which the formal link with the knowledge engineering problem is established and formalised. Conceptual modelling should culminate in a conceptual model, possibly translated into an analytical plan. Figure 4 highlights that the design could lead to different design models and how for each design model there could be more implementation models. For example, this is the case when a game has both a digital and a board game version. The outcomes for the different implementation models might not only inform directly the solution/understanding step, but also other implementation, design or conceptual models. While van der Zee's model [13] presents a cycle between the four identified steps, the analytical game design cycle proposed in this paper includes bidirectional and additional arrows to illustrate the different feedback loops within the cycle (Figure 4). For example, a feedback loop between the solution/understanding and the conceptual model has been introduced in addition to the ones in van der Zee's model. This feedback loop represent cases in which further refinements to the experiment might be needed before implementing the solution or cases in which a different model might need to be implemented (i.e., a game may need a digital version to collect additional data). Therefore, a new design cycle iteration might have to be started. The same applies to the feedback loop introduced between the implementation models and the conceptual model. Finally, the design cycle is enriched with the assessment framework for A&A (Section 4.3). Often the considerations regarding the introduction of A&A are technology oriented and design choices are driven by engineering factors related to resources and technical feasibility. Although important, these engineering decision factors need be complemented by design considerations that are able to provide the full spectrum of the impact of A&A adoption, including the ones on the human component (e.g., players, facilitators and adjudicators). Often the argument on the importance of such factor is made within domains characterised by safety concerns (e.g., safety in automated vehicles). Although less evident in the case of analytical games, the importance is still very high. In fact, it might impact the analytical rigour and the quality of the collected data. Figure 4 positions the A&A assessment framework with respect to the overall analytical serious game design cycle. It shows how this assessment should inform the design choices with respect to the implementation model and should be driven by the analysis of the factors in the conceptual model.

4.3 Assessment of automation and autonomy adoption for analytical games

When assessing what functions of the analytical game should involve A&A, an evaluation framework should be applied, in order to ensure a correct assessment of the potential implications and impacts. In this context, we could consider an extension of the approach proposed by [54]. This work presents an iterative process that includes the following steps:

- (i) the assessment of what should be automated;
- (ii) the identification of the potential types of automation, namely the automation dimensions involved (i.e. information acquisition, information analysis, decision selection and action implementation);
- (iii) the identification of the potential level of automation for each dimension (Section 3.3);
- (iv) the application of primary evaluative criteria;
- (v) the initial definition of types and levels of automation after the application of the primary evaluation criteria;
- (vi) the application of secondary evaluative criteria;

(vii) the final definition of types and levels of automation.

The primary evaluative criteria include the consequences on human performance with a focus on situational awareness, mental workload, complacency and skill degradation, while the secondary ones include consideration on automation reliability and costs associated to decision and implementation outcomes.

For analytical games a similar assessment framework, referring both to automation and autonomy, could be adopted (4). However, the first aspect to take into consideration would be the assessment of whether to introduce A&A at all. The related primary assessment criteria should take into account: (i) the need for A&A; (ii) the feasibility of its adoption; and (iii) the cost of a potential A&A failure. In fact, as clearly highlighted by [50], automation should be adopted only when needed in order to avoid unnecessary system complexity, decision biases and issues to performances (Section 4.4). Moreover, an initial assessment of resources and time available should be performed. One particular element to take into consideration from the start are the potential classification restrictions both on information technology systems to be used and data to be collected. The cost of a potential failure, instead, represents the consequences of mismatches between the game results and the analytical needs and objectives. Therefore, it includes not only the consequences of not collecting data due to technical failure, but also the cost of poor scientific rigour of the analytical game and its results, which might be further used in the design of algorithms and systems. After the assessment on the A&A adoption on the basis of the primary criteria, the assessment on the type of A&A should take place. This assessment aims at deciding whether to employ automation or autonomy, in which phase of the game (e.g., which task) and which actors are supported or impacted. Finally, the secondary assessment criteria (i.e., out-of-the-loop performance, information presentation, monitoring and vigilance, trust, learning and engagement) should inform the evaluation of LOA along the different dimensions (i.e., information acquisition, information analysis, decision and action). Out-of-the-loop (OOTL) performance refers to the concept that performances might be decreased by removing the human “from a control loop due to allocation of system functions to an automated/computer controller” [76]. The secondary criteria would lead to the assessment of the A&A impact on the game life cycle, which can be further decomposed in the following components: (i) the impact on the design (from a designer perspective); (ii) the impact on the execution (both from a facilitators, adjudicator and player perspective); and (iii) the impact on data collection.

4.4 *Discussions on secondary criteria*

An overview of the literature reveals that gaming research investigating aspects related to the secondary criteria is limited (e.g., [77]). While work exploring these elements in context (i.e., analytical gaming) would be beneficial, initial guidelines can be found in the human factor and cognitive engineering literature. This research has investigated the impact of A&A on human performance, exploring it along the different factors that build up such performance (e.g., situational awareness or workload). Relevant results of these studies have been summarised by [50], which proposes a model for human-autonomy system oversight. This work and other related literature contains detailed discussion on this topic. This article will focus on aspects that are relevant to analytical games.

OOTL performance is a key element related to human performance in the light of A&A use. Specifically, studies have identified A&A brittleness (i.e., [78]) as an issue due to the difficulty posed to humans that might be slow at realising the malfunctioning of A&A and the reasons for such malfunctioning [79]. Additional issues are posed by unexpected mode transition, when A&A fails. In fact, resuming manual control might be difficult. This aspects give rise to the phenomenon known as OOTL performance, which has been correlated to loss

of situational awareness, due to changes in information presentation, human monitoring activities, human vigilance, trust and engagement [11]. The brittleness and OOTL performance might severely impact analytical gaming. For example, facilitators might have difficulty in resuming manual control (if possible) when the technological component of the game fails. Issues related to information presentation might also significantly impact analytical games, given the implications on the game actors' situational awareness. The information presentation issue might be correlated to the low levels of transparency and the inability of the A&A interfaces to meet all the information needs that lead to an appropriate level of situational awareness [50].

With respect to vigilance and A&A complacency, it has been demonstrated that increases in trust might lead to a decreased level of human monitoring and attention allocation to critical information [80]. Trust in A&A has been found to be influenced by individual factors, situational factors (e.g. time constraints and effort) and system factors, such as system reliability and performances [81, 82]. The impact of trust, vigilance and complacency also applies to game actors' performances, whose attention to critical information might be decreased.

A loss of situational awareness has also been observed in studies that linked it to low engagement and loss of information cues due to humans not in direct control of tasks [83]. In fact, passive monitoring of A&A might lead to low cognitive engagement, with a decrease of the understanding and information retention. Cognitive engagement is of primary importance in gaming, as highlighted in several game domains (e.g., [72, 84, 85]). Moreover, cognitive science has shown the relationship between cognitive engagement, motivation to think and information processing behaviours (e.g., [86]). Specifically, it has been demonstrated that engagement enhances the individual motivation to elaborate information (e.g., [87]). The decision-making and the underpinning information processing are at the core of the studies conducted through analytical games.

A&A complexity is another important aspect to account for. In fact, it might impact situational awareness through a reduction of system predictability, due to the addition of features, modes and interactions between subsystems [88]. Moreover, the increased complexity and unfamiliarity with technology, might negatively impact the players' learning curve, therefore requiring longer introduction and familiarisation sessions at the beginning of the game execution. This might have a negative impact on the overall game as experts' time availability is costly and often limited.

Finally, A&A might induce skill degradation [54]. This is probably going to be a negligible aspect from a player perspective, however, it might have significant consequences on the facilitators', data collectors' and adjudicators' skills in the long term. Therefore, organisations may need to consider this implication in their analysis.

5 *K2AG use case*

K2AGs focus on collecting data on decision-making at cognitive level, making potential perturbation factors of cognitive processes a critical aspect of K2AG design. Therefore, perturbation factors correlated to A&A adoption in K2AGs should be carefully assessed. As for each assessment related to the design of a system, trade-offs drive the design choices. Such trade-offs might change over time. Therefore, when iterative design is part of the system life-cycle, reassessments of the initial choices might occur. The framework for analytical games design presented in this paper is here applied to the Reliability Game (Section 2.1) as a case study. In the interest of brevity we summarise the salient aspect of the design cycle and refer the reader to [28, 29] for further details.

The tasking for the Reliability Game was to develop a K2AG for an upcoming table-top

exercise (TTX) to explore the impact of uncertainty on human information processing and to collect data to inform the development of advanced artificial intelligence algorithms. There were three main requirements: (i) a fast design, development and deployment of the game; (ii) very limited dedicated resources; and (iii) the capability to collect a significant data sample for further analysis.

As the first design step, the researchers started defining the objectives, scope and content of the model. Specifically, it was decided to focus on source reliability and its impact on situational awareness. Literature has widely discussed how reliability is actually the result of several underpinning factors [89]. The source factors selected as experimental factors were two, namely source type and source quality. Care had to be given to the understanding on how to represent the factors (i.e., rating scales) and their variations. Moreover, how to capture their impact on human situational assessment had to be defined. The impact was recorded in the forms of human belief changes. These beliefs can then be modelled as subjective probabilities or belief functions and analysed. After the conceptual model has been created it was decided to proceed with the experimentation through the K2AG approach. Therefore, the conceptual model has been turned into a design model. Here it was decided to take advantage of messages to be used as a means to convey the experimental factors as well as other information. Moreover, the means of capturing beliefs was established. Specifically, the use of geometrical shapes was selected as the main data collecting approach [8]. Once the design model had been established platform specific design choices entered into play. For the first design iteration the application of the primary criteria led to the choice of implementing the Reliability Game on a non-digital platform (i.e., board with the belief data gathering area and cards with messages). In fact, the need for A&A was estimated as low as the players for the TTX event would be physically present on premises. Moreover, for game validation purposes a direct observation on how the player interacted with the game was given paramount importance. The feasibility of A&A adoption was estimated as low as the time available for design, development and testing was very limited, as well as the personnel resources. The risk associated with a potential A&A failure was not neglectable, not only due to the reduced development and testing time, but also in relation to the associated costs, which included a reputation cost as well as the inability to collect a useful data-set for further research and development activities. The first deployment of the Reliability Game enabled the collection of a useful data sample that supported the validation of the game (including an analysis of the effectiveness of the game mechanics) and an initial qualitative analysis on the collected data.

In order to collect significant data samples to perform a quantitative analysis, several additional sessions of the game have been played along one year. The long time to collect the additional data samples was due to the player availability and the need to have the player and the facilitator physically co-located. Moreover, it required a considerable amount of hours of availability of the facilitator, which in this case was also in charge of the analysis. Therefore, a second A&A adoption assessment has been performed. In this second iteration, the feasibility of A&A adoption was regarded as medium. In fact, although the resources available were still limited, there were no more specific time constraints for development and testing. Moreover, the designers had a good understanding of the game mechanics effectiveness and efficiency. Finally, the cost of a potential A&A failure in this case was low as the reputation cost could be balanced by appropriate testing before deployment. The assessment led to the development of a small digital prototype of the Reliability Game, in the form of a computer stand-alone game. Specifically, it automates the means through which the game is delivered (i.e., a software instead of a board game) and the facilitation tasks.

With respect to the secondary assessment criteria for this design solution, the OOTL performance was considered a minor issue, as the game session length is short and the kind of activities performed by the player would remain unaltered as only the support platform would

Table 1: Automation and autonomy (A&A) adoption criteria applied to the two design iterations of the Reliability Game

Iteration	Criterion	Impact estimate	Motivation
First iteration			
Primary criteria	Need for A&A	Low	- TTX planned as a physical event - Need to observe the players-game interaction
	Cost of failure	Medium	- Reputation cost and absence of a data-set for further research
	Adoption feasibility	Low	- Short time for design, development and testing - Limited resources in terms of personnel
Second iteration			
Primary criteria	Need for A&A	Medium	- Increase data-sample and reduce facilitation burned
	Cost of failure	Low	- Reputation cost, diminishable with appropriate testing before deployment
	Adoption feasibility	Medium	- No specific time constraints for development and testing; - Good understanding of the game mechanics effectiveness and efficiency - Limited resources in terms of personnel
Secondary criteria	OOTL performance	Low	- The overall game session is short; - Kind of activities performed by the player would remain unaltered
	Information presentation	Low	- Similar game mechanics as the non-digital game; however, an appropriate evaluation of the game interface would be needed
	Monitoring and vigilance	Medium	- The absence of interaction with the facilitator might reduce the attention allocation to critical information to be processed
	Trust	Medium	- It has been observed how trust and involvement of the players grow when they were able to understand the final goal of the game
	Learning	Medium	- Similar game mechanics as the non-digital game; however, due to the absence of the facilitator appropriate guidance to the player is needed within the game
	Engagement	Medium	- Social interaction is an important component of the games - Players error management and recovery need to be well articulated

change from analogue to digital. The information presentation factor was assessed as having a medium impact. In fact, the information presentation would follow similar game mechanics as the non-digital game (i.e., messages provided through cards). However, an appropriate evaluation of the game interface would be needed to ensure a correct human-system integration and user experience. Monitoring and vigilance, trust and engagement are considered the factors that required further evaluation in this second design iteration. In fact, the absence of social interaction could influence engagement. The lack of a facilitator might impact player's monitoring and vigilance, reducing the attention allocation to critical information to be processed, which is a pillar of this specific game. Finally, the automation of the facilitation task could negatively impact player trust. Specifically, observations reveal that trust and involvement of the players grew when they were able to interact with the facilitator to better understand the

final goal of the game. Also the learning curve of the players could be negatively influenced in this second design iteration. In fact, the simple mechanics used in the previous design were retained, but the absence of a facilitator explaining and monitoring how players would use the data collection method proposed could represent an obstacle. Therefore, to minimise the impact a correct strategy to provide appropriate guidance to the player on how to navigate in the game is needed. Furthermore, players error management and recovery need to be well articulated to avoid frustrations or misinterpretations, that might decrease overall trust and engagement. Table 1 summarises the application of the A&A adoption criteria in the two Reliability Game design iterations.

6 Conclusion

This work presents a design framework for analytical games, which aims at guiding designers through the different design phases, which include conceptual modelling, design modelling and implementation modelling activities. Automation and autonomy technology might be a relevant support in the advancement of analytical gaming practice. The proposed framework provides a structured approach for thinking through the implications of introducing A&A and which factors to consider during analysis, in order to ensure the needed scientific rigour in the use of analytical games. The framework includes a set of preliminary assessment criteria, which build on several decades of research conducted in the field of human factor and cognitive engineering. The primary criteria identified for such assessment are: need for A&A, its adoption feasibility and the cost of a potential failure. The results of this first assessment would support the design choices related to the A&A type selection (i.e., automation or autonomy), the game phases where A&A might be introduced and the actors impacted by A&A. Finally, secondary assessment criteria should inform the selection of the appropriate level of A&A. These criteria relate to out-of-the-loop performances, information presentation, monitoring and vigilance, trust and engagement. Understanding such aspects will strongly contribute to the scientific rigour of the games employed. The use case of the Reliability Game is here reported to show how the proposed criteria provided the rationale behind many of the design choices. The application of this framework to K2AGs is contributing to the development of a complete and flexible knowledge acquisition toolkit for K2AG designers and users. Nevertheless, the summarised results support the need to build a stronger bridge between the serious game research field and other scientific areas that have been investigating these issues. This will allow taking advantage of the extensive body of knowledge available and their potential future developments. Further research should investigate the factors identified specifically within the context of analytical games automation and autonomy adoption, possibly defining relevant evaluation criteria and metrics.

Acknowledgments

This research received partial support by NATO Allied Command Transformation (NATO-ACT) through the Data Knowledge & Operational Effectiveness (DKOE) programme of work of the NATO STO Centre for Maritime Research and Experimentation.

References

- [1] C. Abt, *Serious Games*. New York: The Viking Press, 1970.



- [2] “Serious game,” retrieved on 24 April 2018, <https://www.igi-global.com/dictionary/serious-games/26549>.
- [3] S. Vermillion, R. Malak, R. Smallman, B. Becker, M. Sferra, and S. Fields, “An investigation on using serious gaming to study human decision-making in engineering contexts,” *Design Science*, vol. 3, no. 15, 2017. doi: 10.1017/dsj.2017.14
- [4] D. Djaouti, J. Alvarez, and J.-P. Jesse, “Classifying serious games: the G/P/S model,” *Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches*, vol. 1, pp. 118–136, 2011. doi: 10.4018/978-1-60960-495-0.ch006
- [5] I. Ståhl, “What is operational gaming?” in *Operational Gaming*, ser. Frontiers of Operational Research and Applied Systems Analysis, I. Ståhl, Ed. Oxford: Pergamon Press, 1983.
- [6] E. M. Bartels, “Building better games for national security policy analysis: Towards a social scientific approach,” Ph.D. dissertation, 2020.
- [7] R. Duke and J. Geurts, *Policy games for strategic management*. Amsterdam, The Netherlands: Dutch University Press, 2004.
- [8] F. de Rosa, “Knowledge acquisition analytical games: games for cognitive systems design,” Ph.D. dissertation, University of Genoa, Italy, 2020.
- [9] D. Bauer and I. Wolff, “Biases in beliefs: Experimental evidence,” Thurgauer Wirtschaftsinstitut, Universität Konstanz, TWI Research Paper Series 109, 2018.
- [10] S. Hoppenbrouwers, B. Schotten, and P. Lucas, “Towards games for knowledge acquisition and modeling,” *International Journal of Gaming and Computer-Mediated Simulations*, vol. 2, no. 4, pp. 48–66, 2010. doi: 10.4018/jgcms.2010100104
- [11] R. M. Endsley, “Toward a theory of situation awareness in dynamic systems,” *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995. doi: 10.1518/001872095779049543
- [12] J. H. G. Klabbers, *The Magic Circle: Principles of Gaming & Simulation*, third and revised ed. Sense Publishers, 2009.
- [13] D.-J. van der Zee and B. Holkenborg, “Conceptual modelling for simulation-based serious gaming,” in *Proceedings of the Winter Simulation Conference*. Winter Simulation Conference, 2010. doi: 10.1109/WSC.2010.5679133 pp. 522–534.
- [14] F. de Rosa, A.-L. Joussetme, and A. De Gloria, “The Knowledge Acquisition Analytical Game framework for cognitive systems design,” in *Advances in Intelligent Systems and Computing*, D. Russo, T. Ahram, i. W. Karwowski, G. Di Bucchianico, and R. Taiar, Eds., vol. 1322. Springer, Cham., 2021. doi: 10.1007/978-3-030-68017-6_82
- [15] “Automation,” retrieved on 14 February 2021, <https://www.dictionary.com/browse/automation>.
- [16] U. Krogmann, “From automation to autonomy: Trends towards autonomous combat systems,” in *Advances in Vehicle Systems Concepts and Integration (RTO MP-44)*. Neuilly Sur-Seine Cedex, France: NATO Research and Technology Organization, 1999.
- [17] P. M. Fitts, *Human engineering for an effective air-navigation and traffic-control system*. Washington, DC: National Research Council, 1951.
- [18] L. Onnasch, C. Wickens, H. Li, and D. Manzey, “Human performance consequences of stages and levels of automation: an integrated meta-analysis,” *Human Factors*, vol. 56, no. 3, pp. 476–488, 2014. doi: 10.1177/0018720813501549
- [19] R. Duke, “A paradigm for game design,” *Simulation & Gaming*, vol. 11, no. 3, pp. 364–377, 1980. doi: 10.1177/104687818001100308
- [20] S. Kurapati, I. Kourounioti, H. Lukosch, L. Tavasszy, and A. Verbraeck, “Fostering sustainable transportation operations through corridor management: A simulation gaming approach,” *Sustainability*, vol. 10, no. 2, pp. 364–377, 2018. doi: 10.3390/su10020455

- [21] N. J. Habraken and M. D. Gross, “Concept design games,” *Design Studies*, vol. 9, no. 3, pp. 150–158, 1988. doi: 10.1016/0142-694X(88)90044-0
- [22] E. Brandt and J. Messeter, “Facilitating collaboration through design games,” in *8th Conference on Participatory Design*, Toronto, ON, 2004. doi: 10.1145/1011870.1011885
- [23] M. Fumarola, J.-P. van Staalduinen, and A. Verbraeck, “A ten-step design method for simulation games in logistics management,” *Journal of Computing and Information Science in Engineering*, vol. 12, no. 1, pp. 150–158, 2011. doi: 10.1115/1.3617440
- [24] Z. Kosmadoudi, T. Lim, J. Ritchie, S. Louchart, Y. Liu, and R. Sung, “Engineering design using game-enhanced cad: the potential to augment the user experience with game elements,” *Computer-Aided Design*, vol. 45, pp. 777–795, 2013. doi: 10.1016/j.cad.2012.08.001
- [25] J. van den Hoogen and S. Meijer, “Gaming and simulation for railway innovation: a case study of the dutch railway system,” *Simulation & Gaming*, vol. 46, no. 5, pp. 489–511, 1980. doi: 10.1177/1046878114549001
- [26] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popovic, and F. players, “Predicting protein structures with a multiplayer online game,” *Nature*, vol. 466, pp. 756–760, 2010. doi: 10.1038/nature09304
- [27] A.-L. Joussetme, G. Pallotta, and J. Locke, “Risk Game: Capturing impact of information quality on human belief assessment and decision making,” *International Journal of Serious Games*, vol. 5, no. 4, pp. 23–44, 2018. doi: 10.17083/ijsg.v5i4.258
- [28] F. de Rosa, A.-L. Joussetme, and A. De Gloria, “A Reliability Game for source factors and situational awareness experimentation,” *International Journal of Serious Games*, vol. 5, no. 2, pp. 45–64, 2018. doi: 10.17083/ijsg.v5i2.243
- [29] F. de Rosa, A. De Gloria, and A.-L. Joussetme, “Analytical Games for Knowledge Engineering of Expert Systems in Support to Situational Awareness: the Reliability Game case study,” *Expert Systems with Applications*, vol. 138, pp. 112 800–112 811, 2019. doi: 10.1016/j.eswa.2019.07.017
- [30] F. de Rosa and A. De Gloria, “An analytical game for knowledge acquisition for maritime behavioral analysis systems,” *Applied Science*, vol. 10, no. 2, p. 591, 2020. doi: 10.3390/app10020591
- [31] F. de Rosa and C. Strode, *Lessons learned from Autonomous ASW TTX*. NATO Science & Technology Organisation, Centre for Maritime Research and Experimentation, in press, no. MR.2020.013.
- [32] R. Hunicke, M. Leblanc, and R. Zubek, “MDA: A formal approach to game design and game research,” in *Proceedings of the AAAI Workshop on Challenges in Game AI*, 2004, pp. 1–8.
- [33] T. Marsh, “Activity-based scenario design, development, and assessment in serious games,” in *Gaming and Cognition: Theories and Practice from the Learning Sciences*, 2010, pp. 213–225.
- [34] M. B. Carvalho, F. Bellotti, R. Berta, A. De Gloria, C. I. Sedano, J. Baalsrud Hauge, J. Hu, and M. Rauterberg, “An activity theory-based model for serious games analysis and conceptual design,” *Computers and Education*, vol. 87, pp. 166–181, 2015. doi: 10.1016/j.compedu.2015.03.023
- [35] C. S. Greenblat and R. D. Duke, *Principles and practice of gaming simulation*. London: Sage Publications, 1981.
- [36] C. S. Greenblat, *Designing Games and Simulations: An Illustrated Handbook*. Newbury Park, Calif: Sage Publications, 1988.
- [37] S. Robinson, “Conceptual modelling for simulation Part I: definition and requirements,” *Journal of the Operational Research Society*, vol. 59, no. 3, pp. 278–290, 2008. doi: 10.1057/palgrave.jors.2602368

- [38] D. van der Zee, B. Holkenborg, and S. Robinson, "Conceptual modeling for simulation-based serious gaming," *Decision Support Systems*, vol. 54, no. 1, pp. 33–45, 12 2012. doi: 10.1016/j.dss.2012.03.006
- [39] S. Burns, Ed., *War Gamers' Handbook: A Guide for Professional War Gamers*. U.S. Naval War College Newport, United States, 2015.
- [40] United States Army War College, *Strategic Wargaming Series Handbook*, 2015.
- [41] P. Battistella and C. G. Wangenheim, "ENgAGED: Games development process for Computing Education," Brazilian Institute for Digital Convergence, Department of Informatics and Statistics, Federal University of Santa Catarina, Brazil, Tech. Rep. 01/2015, 2015.
- [42] T. Mitamura, Y. Suzuki, and T. Oohori, "Serious games for learning programming languages," in *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2012. doi: 10.1109/ICSMC.2012.6378001 pp. 1812–1817.
- [43] T. M. Connolly, M. Stansfield, and T. Hainey, "An application of games-based learning within software engineering," *British Journal of Educational Technology*, vol. 38, no. 3, pp. 416–428, 2007. doi: 10.1111/j.1467-8535.2007.00706.x
- [44] Development, Concepts and Doctrine Centre, UK Ministry of Defence Shrivenham, *Wargaming Handbook*, 2017.
- [45] P. Thavikulwat, "The architecture of computerized business gaming simulations," *Simulation & Gaming*, vol. 35, no. 2, pp. 242–269, 2004. doi: 10.1177/1046878104263545
- [46] Y. H. Wong, S. J. Bae, E. M. Bartels, and B. Smith, *Next-Generation Wargaming for the U.S. Marine Corps: Recommended Courses of Action*. Santa Monica, CA: RAND Corporation, 2019.
- [47] Society for Automotive Engineers, *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. SAE International, 2018.
- [48] C. D. Wickens, A. S. Mavor, R. Parasuraman, and J. P. McGee, *The Future of Air Traffic Control: Human Operators and Automation*. Washington, DC: National Academy Press, 1998.
- [49] P. Hancock, "Imposing limits to autonomous systems," *Ergonomics*, vol. 60, no. 2, pp. 1–8, 2016. doi: 10.1080/00140139.2016.1190035
- [50] M. Endsley, "From here to autonomy," *Human Factors*, vol. 59, no. 1, pp. 5–27, 2017, <https://doi.org/10.1177/0018720816681350>.
- [51] W. Xu, "From automation to autonomy and autonomous vehicles: challenges and opportunities for human-computer interaction," *ACM Interactions*, vol. 27, no. 6, 2020. doi: 10.1145/3434580
- [52] T. B. Sheridan, "Computer control and human alienation," *Technology Review*, vol. 83, no. 1, pp. 60–73, 1980.
- [53] M. R. Endsley, "Level of automation: Integrating humans and automated systems," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 41, no. 1, pp. 200–204, 1997. doi: 10.1177/107118139704100146
- [54] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 30, no. 3, pp. 286–297, 2000. doi: 10.1109/3468.844354
- [55] M. Groover, *Automation, Production Systems, and Computer-Integrated Manufacturing*. Upper Saddle River, N.J.: Prentice Hall, 2001.
- [56] J. Frohm, V. Lindström, J. Stahre, and M. Winroth, "Levels of automation in manufacturing," *Ergonomia - an International journal of ergonomics and human factors*, vol. 30, no. 3, 2008.

- [57] M. R. Endsley and D. B. Kaber, "Level of automation effects on performance, situation awareness and workload in a dynamic control task," *Ergonomics*, vol. 42, no. 3, pp. 462–492, 1999. doi: 10.1080/001401399185595
- [58] B. Lorenz, F. Di Nocera, and R. Parasuraman, "Human performance during simulated space operations under varied levels of system autonomy," in *Proceedings of Bioastronautics Investigators' Workshop*, Galveston, Texas, USA, 2001.
- [59] D. Broadbent, *Perception and Communication*. Pergamon Press, 1958.
- [60] G. A. Jamieson and J. Gyrd Skraaning, "Levels of automation in human factors models for automation design: Why we might consider throwing the baby out with the bathwater," *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 42–49, 2018. doi: 10.1177/1555343417732856
- [61] D. B. Kaber, "Issues in human-automation interaction modeling: Presumptive aspects of frameworks of types and levels of automation," *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 7–24, 2018. doi: 10.1177/1555343417737203
- [62] J. D. Lee, "Perspectives on automotive automation and autonomy," *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 53–57, 2018. doi: 10.1177/1555343417726476
- [63] S. S. Tavallaey and C. Ganz, "Automation to autonomy," in *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2019. doi: 10.1109/ETFA.2019.8869329 pp. 31–34.
- [64] K. C. Morris, C. Schlenoff, and V. Srinivasan, "Guest editorial a remarkable resurgence of artificial intelligence and its impact on automation and autonomy," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 2, pp. 407–409, 2017. doi: 10.1109/TASE.2016.2640778
- [65] J. Laird and M. VanLent, "Human-level AI's killer application: Interactive computer games," *AI Magazine*, vol. 22, no. 2, p. 15, 2001. doi: 10.1609/aimag.v22i2.1558
- [66] G. N. Yannakakis, "Game AI revisited," in *Proceedings of the 9th Conference on Computing Frontiers*. New York, NY, USA: Association for Computing Machinery, 2012. doi: 10.1145/2212908.2212954 pp. 285–292.
- [67] P. Boinodiris, L. Sokol, R. R. Behrens, and H. Blejer, "Wargames Powered by IBM Watson: Multiplayer RealTime Strategy in 4D," IBM Redbooks Point-of-View, Tech. Rep. REDP-5128, 2014.
- [68] D. Stapleton, "War gaming supported by cognitive computing and time manipulation," in *Modelling and Simulation Support to Operational Tasks Including War Gaming, Logistics, Cyber Defence*, no. STO-MP-MSG-133. NATO Science and Technology Organisation, 2015.
- [69] S. Rabin, *AI Game Programming Wisdom*. USA: Charles River Media, Inc., 2002.
- [70] G. Calleja, *In-Game: From Immersion to Incorporation*. The MIT Press, 2011.
- [71] S. T. Coleridge, *Biographia Literaria*. London: Rest Fenner, 1817.
- [72] P. Perla and E. McGrady, "Why Wargaming Works," *Naval War College Review*, vol. 64, no. 3, 2011.
- [73] P. Perla, *The Art of Wargaming: A Guide to Professionals and Hobbyists*. Annapolis, Md: Naval institute Press, 1990.
- [74] G. Wagner, "Information and process modeling for simulation - Part I," *Journal of Simulation Engineering*, vol. 1, 2018.
- [75] K. M. L. Cooper and C. S. Longstreet, "Towards model-driven game engineering for serious educational games: Tailored use cases for game requirements," in *2012 17th International Conference on Computer Games (CGAMES)*, 2012. doi: 10.1109/CGames.2012.6314577 pp. 208–212.

- [76] D. B. Kaber and M. R. Endsley, “Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety,” *Process Safety Progress*, vol. 16, no. 3, pp. 126–131, 1997. doi: 10.1002/prs.680160304
- [77] S. Dorton, S. Tupper, and L. Maryeski, “Going digital: Consequences of increasing resolution of a wargaming tool for knowledge elicitation,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, no. 1, pp. 2037–2041, 2017. doi: 10.1177/1541931213601988
- [78] D. Woods and R. Cook, “Incidents-markers of resilience or brittleness?” in *Resilience Engineering: Concepts and precepts*, E. Hollnagel, D. Woods, and N. Levenson, Eds. Ashgate, UK: Aldershot, 2006, pp. 69–76.
- [79] C. D. Wickens and C. Kessel, “The effects of participatory mode and task workload on the detection of dynamic system failures,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 24–34, 1979. doi: 10.1109/TSMC.1979.4310070
- [80] S. Hergeth, L. Lorenz, R. Vilimek, and J. F. Krems, “Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving,” *Human Factors*, vol. 58, no. 3, pp. 509–519, 2016. doi: 10.1177/0018720815625744
- [81] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. de Visser, and R. Parasuraman, “A meta-analysis of factors affecting trust in human-robot interaction,” *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011. doi: 10.1177/0018720811417254
- [82] K. E. Schaefer, J. Y. C. Chen, J. L. Szalma, and P. A. Hancock, “A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems,” *Human Factors*, vol. 58, no. 3, pp. 377–400, 2016. doi: 10.1177/0018720816634228
- [83] D. B. Kaber, E. Onal, and M. R. Endsley, “Design of automation for telerobots and the effect on performance, operator situation awareness, and subjective workload,” *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 10, no. 4, pp. 409–430, 2000. doi: 10.1002/1520-6564(200023)10:4<409::AID-HFM4>3.0.CO;2-V
- [84] E. A. Boyle, T. Hainey, T. M. Connolly, G. Gray, J. Earp, M. Ott, T. Lim, M. Ninaus, C. Ribeiro, and J. Pereira, “An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games,” *Computers & Education*, vol. 94, pp. 178 – 192, 2016. doi: 10.1016/j.compedu.2015.11.003
- [85] I. Buil, S. Catalán, and E. Martínez, “Engagement in business simulation games: A self-system model of motivational development,” *British Journal of Educational Technology*, vol. 51, no. 1, pp. 297–311, 2020. doi: 10.1111/bjet.12762
- [86] M. Fleischhauer and A. Strobel, “Directly and indirectly assessed need for cognition differentially predict spontaneous and reflective information processing behavior,” *Journal of Individual Differences*, vol. 36, no. 2, pp. 101–109, 2015. doi: 1614-0001/a000161
- [87] J. Stephan and J. Brockner, “Spaced out in cyberspace? Evaluations of computer-based information,” *Journal of Applied Social Psychology*, vol. 37, no. 2, pp. 210–226, 2007. doi: 10.1111/j.0021-9029.2007.00157.x
- [88] M. R. Endsley, B. Bolte, and D. G. Jones, *Designing for Situation Awareness: An Approach to User-Centered Design*. London, UK: Taylor & Francis, 2003.
- [89] F. de Rosa and A.-L. Joussetme, *Critical review of uncertainty communication standards in support to Maritime Situational Awareness*. NATO STO Centre for Maritime Research and Experimentation, 2018, no. CMRE-FR-2018-010.