

Game elicitation: exploring assistance in delayed-effect supply chain decision making

Patrick Jost

Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway
patrick.jost@ntnu.no

Monica Divitini

Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway
monica.divitini@ntnu.no

ABSTRACT

Socio-technical systems are creating work environments that are data-driven and real-time oriented. While algorithms can assist in this complex environment, decision making still predominantly relies on humans. An inadequate presentation of information and limited conception of decision outcomes are thereby potential sources of human error. Misperceptions of feedback and time-delayed effects, for example, contribute to the bullwhip effect observed in supply chains. In this case study, we applied the novel approach of game elicitation (GE) to explore human-centred assistance strategies for delayed-effect decision making. We designed a gaming simulation of a supply chain shortage incident to observe four logistics experts and four non-experts trying to balance the distribution system. Qualitative content analysis of thinking aloud protocols and reflective interviews yielded design suggestions for data presentation, monitoring, and learning regarding delayed-effect decisions. Findings suggest applicability in further domains of digital society, such as privacy decision making.

CCS CONCEPTS

• Human-centered computing; • Human-computer interaction (HCI); • Empirical studies in HCI;

KEYWORDS

game elicitation, gaming simulation, serious games, decision making, usability, human-centred design

ACM Reference Format:

Patrick Jost and Monica Divitini. 2020. Game elicitation: exploring assistance in delayed-effect supply chain decision making. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*, October 25–29, 2020, Tallinn, Estonia. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3419249.3420154>

1 INTRODUCTION

Humans have become part of complex socio-technical environments where real-time processing of data and artificial computing are increasingly changing workflows and influencing decision making. Because of this digitalisation process, the basis for decisions – information – becomes essentially hidden in rapidly processed

data streams. Although advanced computing is aiding, decision making relies on humans either due to responsibility or private context. However, in complex situations, humans are tending towards heuristics for decision making, which is related to bounded rationality [46], instinctive behaviour [12] and decision making under uncertainty [32, 50]. Humans' limited cognitive and temporal resources [47] are leading to merely satisfactory rather than optimal decisions. A fact that is used deceptively, for example, by dark patterns interaction design to nudge people into accepting unfavourable privacy decisions [18]. Conversely, human-factors design strategies that assist with supportive information at the appropriate time in useful form can contribute to avoiding unreflective decision making or miscalculations. Exemplary, useful design approaches for assistance in monitoring supply chain processes may help to prevent or reduce the bullwhip effect initially described by Forrester [14]. Originating from misperceptions of feedback and time delays, negative consequences of this persisting supply chain challenge are either inefficient production or excessive inventory.

Supply chain managers require appropriate information on the state of the complex system and predictive support for balancing the supply chain events. As Kochenderfer [32] points out in this regard, robust decision-making systems account for uncertainty in the current state and the future outcomes. The bullwhip effect has been demonstrated with simulation games called beer distribution game and several variations for research have been implemented. For example, to investigate human factors with multiple players/groups playing stakeholders [4] or varying goods [42]. As pointed out by Brauner et al. [4], the decision-making difficulty arises from the time delay between the submission of an order and its fulfilment. In this regard, Nienhaus et al. [41] confirmed that humans act as obstacles and increased information flow might reduce the bullwhip effect. However, which human-centred design strategies support delayed-effect decision making was not investigated.

In this case study, we explore assistive design strategies in delayed-effect decision making. With a two-step game elicitation (GE) approach, we investigated two research questions:

1. How can the GE artefact be improved for learning about decisions with a delayed outcome?
2. How can humans be supported in delayed-effect decision making in a supply chain task?

We describe the novel approach of game elicitation to combine and contrast the perspectives of logistics managers and non-experts on required decision support. Taking the design science (DS) approach proposed by Hevner [21], we constructed a gaming simulation that consists of the classical distribution chain manufacturer, distributor, supplier, retailer. In the rigour-design cycles, the game design and human-factors experts worked together with the domain

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

NordiCHI '20, October 25–29, 2020, Tallinn, Estonia

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-7579-5/20/10.

<https://doi.org/10.1145/3419249.3420154>

experts of the case company to implement a crisis event simulation scenario. Consequently, the game goal of holistic re-balancing corresponded to the role of supply chain event management and allowed identifying more generalised suggestions on delayed-effect decisions support. We contribute to the human-centred research knowledge base in two ways, practical and scientific [31]:

1. First, the created game artefact is intended to function as a proxy to the complex process and at the same time helping to unify perspectives of experts and novices. Based on our insights from qualitative content analysis (QCA) [36] of reflection-in-action during game sessions, we present design implications for future applications of game elicitation as a research approach.
2. Second, by QCA of in-depth debriefing interviews with experts and novices, we elicit human-centred design suggestions to support delayed-effect decision making.

2 RELATED RESEARCH

2.1 Human factors in decision making

Decision making science is oriented along different strands of research classifiable in normative, prescriptive and descriptive decision making [2]. Descriptive decision making is concerned with actions in real contexts and what conditions will influence the decisions related to human factors and cognition. Decisions in the context of complex systems are thereby underlying cognitive and system-related constraints. As suggested with the term bounded rationality [46], in these complex situations, people resort to heuristics [50] to decide in a "satisficing" rather than optimising way. From a systems perspective, a reduction of complexity can, therefore, improve decisions by reducing complexity and extraneous cognitive load [47] during the decision process.

Additionally, several biases are influencing human's decision making, including past experiences [26], regrets of previous decisions [44] and cognitive predispositions [3, 11, 48]. As pointed out by Evans [12] and supported by Greis et al. [19] cognitive abilities are also linked to reasoning and may lead to more automatic choices without reflection but higher risk-taking. Human-centred design improvements can be a way to reduce complexity for human actors in complex socio-technical systems (STS) [49]. A good example of a dynamic system that creates complexity through considerably delayed outcomes of decisions are modern supply chains [5]. A well-researched logistic phenomenon that originates from human information processing in supply chains is the bullwhip effect [14]. Separate calculation of demand and distortion of information, misperceptions of feedback and time delays are involved in causing disbalance and inefficiencies [20]. In relation to bounded rationality, human actors in the system would rely on heuristics in delayed-effect decision making specifically when the cost of collecting more information is too high. However, human factors studies have found that character traits such as conscientiousness and openness with stakeholders in the supply chain mitigate the impact of the effect [28]. In addition, researchers have shown that disasters or sudden emergency events disrupting the supply chain can trigger the bullwhip effect [7].

Previously, the effect has been demonstrated on several accounts with online and offline supply chain games mostly involving multiple disjunct players based on the beer distribution game. It has been adapted to teach about systems thinking [17], disruptions in manufacturing or retail [45] and to demonstrate that humans represent obstacles in the information flow [41]. In this study, we extend the concept of the supply chain game to a GE approach to elicit human-centred design suggestions for decision making support. We design a gaming simulation representing a disruption scenario that needs to be balanced holistically across the entire distribution channel by making goods distribution and routing decisions with delayed results/feedback.

2.2 Gaming simulations as a research approach

Simulating complex systems is by no means a new approach. The benefits of simulated system behaviour have been described by researchers beginning from the 1970ies while emphasising abstraction and design [9] and application as a research approach for social systems [1]. However, game studies as a new strand of science concerned with the act of human playing [23] essentially evolved at the beginnings of the millennium around the debate of the connections of narration and simulation [15, 27, 40] as the defining character of games.

Klabbers [30] thereby proposes the view of the design-in-the-small for the game and design-in-the-large for the referenced STS. Similarly, Rongas et al. [43] refer to the distinction as the game layer (GL) and the simulation layer (SL) while emphasising the advantage of the inclusion of human factors. The game layer entails inherent possibilities to study otherwise unpredictable behaviour of human/organisational actors in a complex system when looking at a combined gaming simulation as human-factors research-site. [1, 43]. Correspondingly, Duke and Geurts [10] refer to gaming simulation as ". . . a method in which human participants enact a role in a simulated environment".

A gaming simulation can, therefore, be a way to elicit information from participants on the referenced STS exploratively. Similar to the method of picture elicitation [6] the simulation game thereby functions as a proxy to elicit narration, for example, of supportive mechanisms for decision making. Identified benefits in elicitation approaches supported with visual media include enhanced participation and control of interviewees, richer data and focus on what was essential to participants. Most importantly, for an explorative study, however, is the suggested effect for emerging unexpected matters [39]. Nonetheless, game elicitation differs from other media-oriented interviewing by adding interactivity and activity-orientation to the evaluation process. As pointed out by Klabbers [31], games in research follow along the three dimensions of human cognition awareness, activity and articulation. While awareness/articulation during play are affected by interaction and consequently, interaction design, the experiential level of players needs to be considered when trying to elicit tacit knowledge in the STS. Klabbers [30] proposes a mapping between GL and SL (the STS reference) with actors, rules and resources. We suggest that this mapping represents what is understood in balancing game goals to domain goals.

Although not all games are simulations (i.e. there might be no reference system), the balancing process applies to every game utilised for game elicitation. Reasons for this are the inherent interaction design and visualisation components of a game that influence humans' cognitive resources during the play activity [25]. Hence, another layer – the interaction layer (IL) – emerges in the design process as mapping GL to SL requires designed abstraction from real-world interaction in an STS. Or expressed in the form of an illustrative example: How is pressing down an iron lever in the real-world designed in the game simulation? These influences from designed abstraction need to be addressed thoroughly in design/rigour cycles to reduce interaction difficulties (i.e. optimise usability) during playing with the simulation. Moreover, this also raises design questions on what is part of the game layers user interface – the IL – and what is part of the reflected STS.

Additional questions are originating from the time domain. In several cases, it is neither possible nor feasible to have a gaming simulation represent the real-time STS frame. A good example is the beer distribution game that either must inform how much time would have passed in reality between a pick-up from factory and delivery to retail on the IL or distort space (i.e. shrink distances) and/or time (speed up) of the reflected STS. In any case, researchers that apply game elicitation have to be aware of the balancing factors in the mapping. By collaborating with domain experts already in the ideation phase of a gaming simulation, distorting or disrupting mappings between game goals and domain goals can be reduced.

Aside from laying bare the inextricable design science relationship of game elicitation, the outlined aspects also influence knowledge creation. Observation and thinking aloud during gameplay can reveal comments on the game interface, the utility of the game as a vessel for content and effectivity and can inspire the player in discovering and constructing explicit knowledge. However, thereof originating knowledge (reflection-in-action) contributes to the knowledge base on the application level and the design of the artefact for utility. For contributions to the analytical/scientific knowledge base, the elicitation context has to be set on the referenced STS. Although players recreate the STS structure by playing [31], the artificialities in the game layer need to be considered. Thus, a thorough debriefing with the participant is required to set perspective on the referenced STS [16, 22, 31, 34] and allow/support tacit knowledge extraction for social problem solving (reflection-on-action) [31]. Consequently, the duality of knowledge creation, practical and scientific, requires appropriate conduct when applying GE but at the same time represents a benefit of the design science approach.

Recent applications of gaming simulations in qualitative research approaches include studying of nursing education [51], railway innovations [37] and humanitarian logistics [38]. The studies are unified in their emphasis of structured debriefing as also pointed out by research on gaming simulation methodology [22, 34]. We follow the recommendations of previous studies with a design-empirical cycle process based on the question of delayed-effect decisions in supply chains and a debriefing session in the GE approach. In addition, we propose that GE can segment knowledge creation to practical knowledge on the game layer and analytical knowledge with reference to the STS.

3 SETTINGS AND PROCESS

3.1 Case company

The case company based in Europe is a worldwide operating freight forwarding company with several thousand employees. Multiple international branches are thereby faced with complex supply chain scenarios in land, sea, and air transport logistics. The business includes departments that are concerned with user experience, e-learning and the corresponding design of services and content. Sub-groups of the enterprise are focusing on optimising support for humans during work and the decision processes involved. Because of their continuous improvements in the work environment, they are frequently awarded as an excellent place of work. To improve human support in delayed-effect decisions that are spread between different branches of the enterprise (e.g. warehouse and retail delivery), a unifying lens was sought for eliciting possible design solutions for integration in the distribution process. Moreover, perspectives of process experts and non-experts/entry-level employees are sought to be explored to identify usability and learning-oriented improvement of decisions with delayed effects.

3.2 Method

Game elicitation was chosen to bridge the gap and bring together the perspectives of logistics experts and novices over complex system decisions. By modelling time-delay in a supply chain disaster scenario with a gaming simulation (see following section 3.3), we created an accessible approach to interact with the problem of decisions with delayed feedback. Participants could engage and reflect on suggestions for support during the decision process. After the initial design science cycles [21] to create the gaming simulation with the logistics experts of the case company, we followed the ideas of Klabbers [31] by segmenting the research approach in two phases representing reflection-in-action and reflection-on-action.

Therefore, the GE process can elicit practically oriented findings for applying the game as a tool for research or learning by thinking aloud and observation during game sessions. On the other hand, when following a thorough debriefing protocol and setting the context to the referenced STS [16, 22, 34], semi-structured interviews

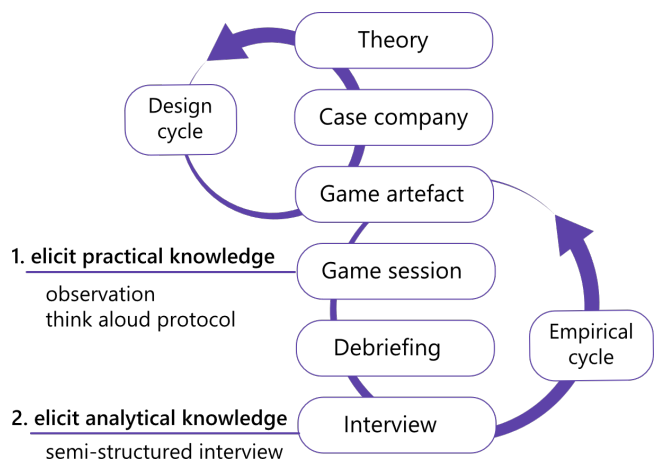


Figure 1: Two-step game elicitation process

with the participants can elicit analytical insights to improvements regarding the real-world processes [31, 38]. Figure 1 illustrates the two-step GE process. The research followed the mixed methods approach suggested by Kuckartz [35, 36] and Kohlbacher [33]. An inductive-deductive coding (see 3.5) was applied to analyse both, practical and analytical findings.

The grounding in the knowledge base of the case company and supply chain effects theory supports the generalisation of practical and analytical findings [13]. Since the research interest was on support in the process and on the referenced STS, we were not evaluating the game result, although another possible asset of GE for future studies.

3.3 Mapping the game layer to the simulation layer

We started the design cycle of the simulation layer according to the literature of the beer game and participatory ideation sessions. During these sessions, experts from the case company, human-factors specialists and game designers were analysing a part of the local supply chain from producers to consumers. Based on the first sessions with the logistics experts' factors of delays (e.g. the production cycles) were analysed and integrated into the SL. In the next phase game, ideation cycles were conducted on how the GL will be mapped on the simulation layer and what would fit best for a balancing challenge. Over the course of 4 months, we conducted two iterative design/evaluation cycles to map the game elements to the disrupted supply chain scenario, develop the roles/actors, resources and rules [29] as described next. The GL to SL mapping process was guided by the MDA framework [24].

Scenario. As suggested by literature that a disaster scenario induced ripple effect can cause the bullwhip effect [8], a scenario of disruption in the supply chain was designed. The case company had several producers in one country and retailers in a neighbouring country. The main route was leading through a massive mountain chain and was only passable through a traffic tunnel or a long detour over a mountain pass. Consequently, the disaster that caused the ripple effect was a massive rockslide that blocked the main route traffic tunnel between town A and town B, as illustrated in Figure 2. As a result, the supply chain was getting in a state of the bullwhip effect with excessive stock at the factory side and empty stock at the retail side. The challenge for the player was now to balance the stock levels by loading and sending trucks between distribution and producers, between distribution and wholesale over the mountain pass, and between wholesale and retail.

Roles/stakeholders. Therefore, the involved roles were given balance bars that indicated if the specific stakeholder is happy according to their stock levels. Since the supply chain was in disbalance due to a disruption, producers and warehouses' stock levels were excessive and retailers stock empty, all stakeholders were unhappy at the beginning of the gaming simulation. The case company was responsible for the transport between distribution and wholesale. The two roles were, therefore, combined in the GL. Finally, three stakeholders were displayed on the GL each having a "mood-indicator" corresponding to stock: producers, company staff (distributor and wholesale), and retailers.

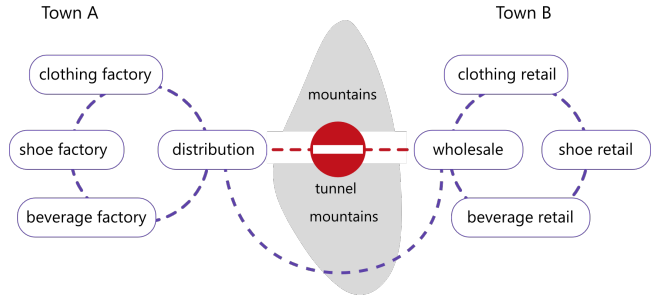


Figure 2: Scenario of the gaming simulation

Resources. The case company handles all kinds of different goods, including liquids and frozen goods. To regard that diversity and resulting complexity, three different types of goods – clothing, footwear, and beverages – were included in the scenario and two trucks of different sizes. The goods were colour-coded to track mixed truckloads and the stock levels of each good. The trucks could be loaded with all three goods. The smaller truck had half the loading space while being twice as fast as bigger trucks. Additionally, trucks had running costs when deployed. The production rate of the factories was decided to stay fixed to maintain a manageable difficulty.

Rules. The game goal was set to balance the mood of all stakeholders (i.e. make them happy) which was reflecting the domain goal as the mood was linked to the stock levels. Therefore, the player could check all stakeholders stock levels of the three goods and balance them over the supply chain by loading and sending trucks. Trucks could be deployed and discharged at the distributor and wholesaler facilities. The planned route of a truck could be changed at any time of the gaming simulation. The main aim was to solve the bullwhip effect state and restore balance as fast and economically as possible by making decisions on cargo loads and routes. Therefore, timer/running costs were shown in the GL and trucks could be followed on the screen in a 3D map. There was a maximum of 12 deployable trucks per town. The game goal was reached when all stakeholders were happy for at least two full minutes. In playtests with domain experts, it took averagely about 20 minutes to balance the game with optimal decision making. It was expected that participants would need more time since they were told to reflect during play and were partly not familiar with the domain.

3.4 Data collection and analysis

For getting a comprehensive insight on decision support suggestions, the GE approach was used to bring together the perspectives of experts and non-experts. We invited four non-experts with no more than 6 months of experience in logistics from the local university and experts with at least 10 years of experience from worldwide operating logistics companies to take part in the study. The GE process was conducted, as outlined in Figure 1.

First, participants were informed about the project and the tasks involved, data processing and signed informed consent before filling out a questionnaire on demographic data and her/his experience

with logistics processes. Participants were then introduced in the gaming simulation interface, rules, roles, resources, and

game goal in a five-minute briefing. Additionally, they were encouraged to think-aloud and discuss their perceptions/suggestions for the game and its potential for learning supply chain decisions. Subsequently, participants started to play while the think-aloud process was taped. The duration of the gaming sessions lasted between 25 to 35 minutes.

Second, a short break of 5 minutes was initiating the debriefing phase of 10 minutes were questions on the participants' experience with the game followed by transitional questions on how the game and reality are connected [22, 34]. Finally, a 40-minute semi-structured interview with a focus on decision making in the referenced STS followed. Open-ended questions such as: "What would help you to make decisions on the orders and to plan the routes?" or "What information/interfaces would support you in your decision making when managing the supply chain?" were asked and taped.

For analysis, the two audio recordings were transcribed verbatim and coded separated to the reflection-in-action and the reflection-on-action parts in an inductive-hermeneutic QCA approach described by Kuckartz [36]. A categorisation scheme was developed by iterative coding. All transcripts were coded by one researcher, while 70% were additionally coded by another researcher to improve data rigour. The main demographic distinction between the two participating groups – 4 experts and 4 non-experts – (Figure 3) was the experience in logistics processes with the experts having averagely 17,5 years of professional experience compared to 0,4 years in the non-expert group. At 36.8 years, the average age of the non-experts was also lower than the average age of the expert group at 49.5 years.



Figure 3: Logistics expert playing the gaming simulation

We first analysed observation protocols and the comments from thinking aloud during the game session to explore the emerging categories for improving the game elicitation approach by QCA. Knowledge elicited through this reflection-in-action was rich on comments regarding the interaction layer and game layer. When comparing experts with non-experts, the experts specifically commented on game-layer additions for learning and suggestions for

reality abstraction. Non-experts, on the other hand, were more commenting on improvements for navigating through different views of the involved roles for keeping an overview of the game process.

Second, we analysed the semi-structured interviews conducted after debriefing with the same mixed-methods QCA approach to elicit human-factors design suggestions for delayed-effect decision making. Knowledge elicited through this reflection-on-action was clustering around suggestions for presenting data during the decision process and real-time monitoring of stock and vehicles.

4 FINDINGS

The following sections present these findings on the practical level (reflection-in-action/RQ1) and scientific level (reflection-on-action/RQ2) by indicating the code quantities of the found suggestions and subsequent contrasting of expert and non-expert views.

4.1 Reflection-in-action

Mapping – interface artificialities: During game session reflection, experts did primarily comment on improvements of game goal/interface artificialities such as the balancing bars that were showing if the stakeholders are happy. On the other hand, non-experts did make more suggestions to the navigational concept and the different views in the game (Figure 4).

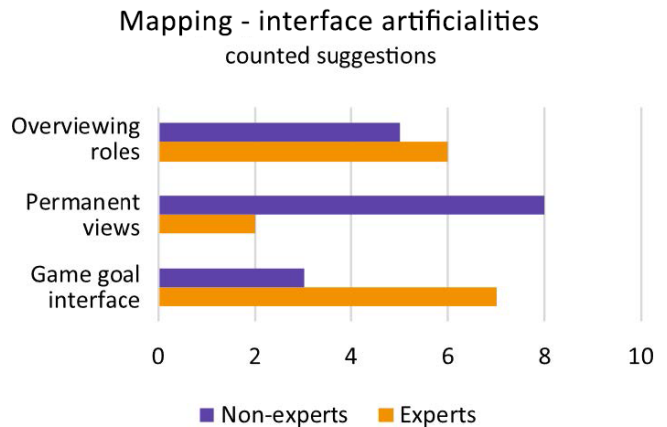


Figure 4: Interface artificialities suggested by experts and non-experts (RQ1)

Finding 1.1: Overlooking more than one role requires comparative interface views to reduce cognitive load.

Combining different roles in one agency (i.e. one player supervises many roles) creates interface artificialities such as multiple views of warehouses. Experts suggested comparative views for managing the roles: "I need a comparative overview of the different stores. Otherwise, it's a challenge to keep in mind what's going on from production to retail." (E-3). Similarly, non-experts addressed the multiple stakeholder views with commenting on cognitive load: "I have many different camps - production, warehouse and retail – and then I'm supposed to remember everything." (NE-1)

Finding 1.2: Permanent views are preferred by less experienced users.

Connected with comparative views for multiple stakeholders, both groups also suggested permanent views over switching views. However, non-experts more frequently expressed the wish for permanent views: "You always have to look and then switch back. Some time has passed in the meantime, and the trucks have arrived; instead of having a permanent view." (NE-2), "Well, there could be something to constantly show the stock level instead of clicking-through somehow." (NE-3)

Finding 1.3 Game goal artificialities are impacting focus and require a detailed pre-game briefing.

Introduced game goal additions impacted the focus on the complex decision processes. For example, the colour scheme for a conceptual distinction of the three different goods in the supply chain did create confusion within experts and non-experts: "The different colours are not clear to me, they should be explained before." (E-3), "I do have difficulties with the colour-coded goods above, probably if I practice it before a few times, I'll do better." (E-4), "Yes, the happy bars are there below, but that is only the game idea, isn't it? I am searching for the inventory of products to solve the problem." (NE-3). Experts commented more on this suggestion than non-experts.

Mapping – learning artificialities: Similarly, but to a much greater extent, experts reflected more on learning additions (Figure 5). Learning artificialities that were suggested included video-introductions, assistance during the decision process, such as a mentor player character and debriefing in the form of chart analysis that can be used to discuss with a trainer.

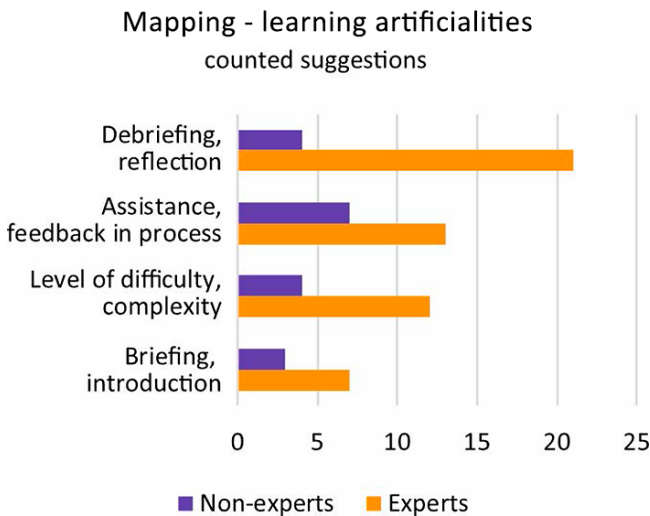


Figure 5: Learning artificialities suggested by experts and non-experts (RQ1)

Finding 1.4: Debriefing transmits knowledge of the underlying complex system decisions.

Experts strongly suggested debriefing as a strategy to promote learning with the game including charts and recorded decisions for reflection between a trainer and trainees after the game: "I think a question is what evaluation tools, for example, trajectories, are there for the debriefing to help the trainer analysing decisions made in the supply chain system and then discuss with trainees." (E-3). Non-experts agreed on looking at recorded data mainly after the gaming

simulation to reflect on the decisions: "For analysis, charts would be good in this process, but less for active use of the game, more for evaluation afterwards." (NE-3)

Finding 1.5: Assistance and feedback while playing supports decision reflection.

Other experts suggested a mentor character inside the game or a multiplayer approach to guide learners: "You'd have to map the decision process out and support with an avatar during the process somewhere, like a reflective boss. Or you can do that through a player community." (E-2). Expert 3 thereby pointed out that: "Feedback assessed during the game, like short multiple-choice questions would help to reflect on the decisions and the delay effect with a trainer afterwards."

Finding 1.6: Adaptive difficulty/complexity helps to learn.

Further learning additions frequently suggested by experts and some non-experts was the dynamic adaption of difficulty/complexity during play: "I think the process is quite easy to understand. But it could help to learn different situations if there is another level of complexity, where there are perhaps more routes." (E-1). One non-expert suggested having unsolvable tasks included: "Because if it is really difficult or maybe even an unsolvable task, then you would learn what to look out for." (NE-1)

Mapping – reality abstraction: Abstractions from reality were, unsurprisingly, addressed mostly by the expert group (Figure 6).

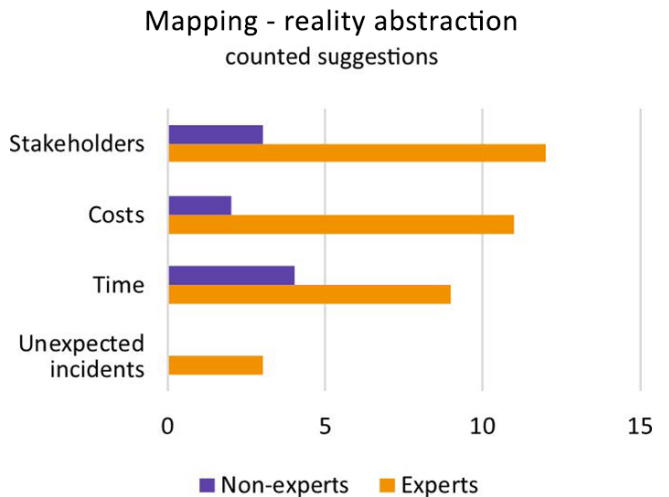


Figure 6: Reality abstractions suggested by experts and non-experts (RQ1)

Finding 1.7: Stakeholder events can improve reality mapping to the referenced complex system.

Thereby stakeholder mapping was suggested to include surprising events or interruptions according to the players game progress: "When things are going badly, annoying or unfriendly phone calls from clients could be introduced to represent more of the reality." (E-2)

Finding 1.8: Cost ratios/fixe time rhythms and time simulations demonstrate decision influences.

Likewise, experts did suggest fewer cost and time abstractions and more regard to economic events such as avoidance of empty

trips: "For the trucks, there is also a cost-benefit ratio. It would help to have this economic perspective somehow. For example, empty trips that could be avoided." (E-2), "Some procedures are in a daily rhythm. Fixed departure times should be considered for influences on decisions." (E-4). Non-experts, on the other hand, suggested having simulation capabilities: "If I could simulate different scenarios in advance, it would help me to see what the influences on the costs are." (NE-4)

Finding 1.9: Unexpected events can improve both game experience and reality mapping.

Finally, for results on the practical application of GE, experts suggested to improve realism but also game experience by including unexpected events in the supply chain mapping: "It would be exciting to have certain events. Unforeseen events somewhere such as tire damage that could also happen in reality." (E-2), "If the goods arrive late at a warehouse, then you could provide workarounds and maybe send them as express delivery to the retailer." (E-4)

4.2 Reflection-on-action

Presenting data: Both, experts and non-experts were strongly suggesting data aggregation in interfaces for support (Figure 7).

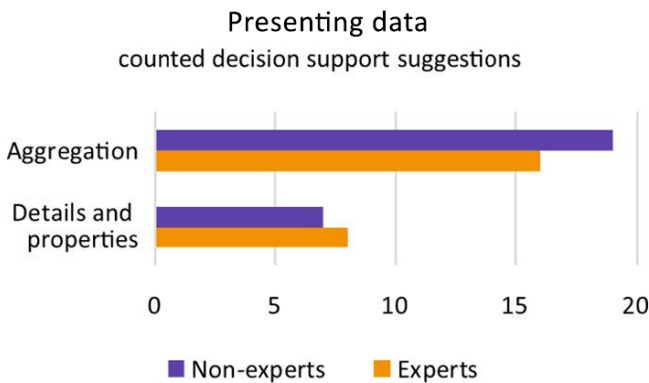


Figure 7: Data presentation suggestions by experts and non-experts (RQ2)

Finding 2.1: Aggregation of information in interfaces supports the decision-making process.

Experts were requesting flexibility to create their own decision overviews: "I would prefer a sort of dashboard that I could assemble myself for the decision-relevant data in the system." (E-2). Other experts pointed out that aside from the goods on stock they also would want an aggregated view on goods inflow, outflow and on the road: "A comparative view of current stock levels, outgoing, incoming and goods in transit would help to identify the delay." (E-3). The consensus for configurable aggregated views is also present in the comments of non-experts: "What would help making decisions is a permanent stock view of all the warehouses I am interested in, which can perhaps be configured." (NE-3)

Finding 2.2: Indicators such as average load, cost/benefit factors or fast-moving goods support decisions.

On the other hand, several suggestions concerning detailed views on properties were also found: "It would be helpful to have an efficiency factor cost/time displayed for trucks and the routes". (E-2), "An average load indicator of the truck can give some clues for more

long-term decisions". (NE-2), "Important are the fast-moving items, it would be useful to get detailed information on those fast-movers". (E-1)

Monitoring – stock movement:

Finding 2.3: Reaction triggers and continuous line-charts support process monitoring.

When asked about support for delayed-effect decisions, stock monitoring suggestions were made by experts and non-experts alike (Figure 8): "It would be great to have a continuous chart on stock levels with minimum/maximum reaction lines." (E-1), "What I would need is an inflow/outflow curve diagram." (E-3), "A warning message, such as an exclamation mark before a stock level drops to a low point, would help to monitor." (NE-2)

Monitoring - stock movement
counted decision support suggestions

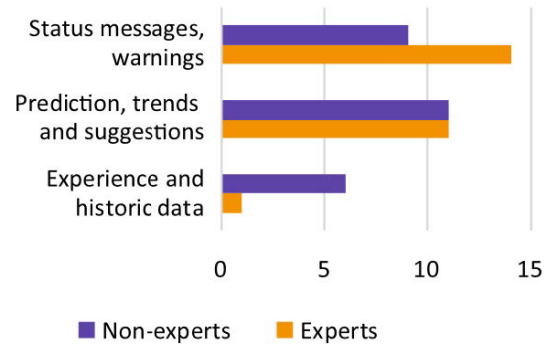


Figure 8: Stock monitoring suggestions by experts and non-experts (RQ2)

Finding 2.4: Predictive trend indicators either as trend charts or numerical assist decision-making.

Predictive support for deciding on deliveries was suggested equally frequent by experts and non-experts: "A continuously calculated graphical trend curve that takes into account predictions on goods movements would be great to have as a tool for decisions." (E-1). However, predictions would not have to be necessarily illustrative: "It would help me if the individual stock movement would be in a table with the predicted time of arrival." (E-4), "A prognosis of retail sales volume would help to make decisions based on the sales cycles." (NE-2)

Finding 2.5: Historical data of retail trade and warehouse turnover can facilitate decisions.

Finally, some of the non-expert group found that historical data would help their decisions: "To decide on deliveries it would be good to know how much the retail trade has sold in the last days or weeks." (NE-1). The same interviewee did further suggest a historical overview of individual warehouses where the timing was critical in the past.

Monitoring – vehicle movement:

Finding 2.6: Estimations on time-of-arrival and progress help with vehicle routing decisions.

Non-experts made considerably more suggestions on monitoring of vehicles than interviewed experts (Figure 9).

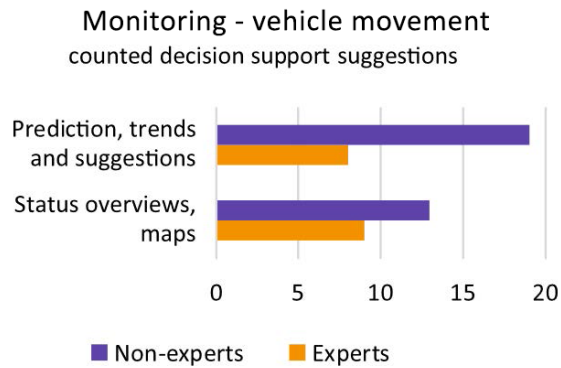


Figure 9: Vehicle monitoring suggestions by experts and non-experts (RQ2)

The form of recommended support thereby focused on prediction and status overviews. Concerning predictions, the form of suggested assistance ranged from estimations of time-of-arrival as a plain number, to maps with colour coded vehicles to monitor if they are in time: *"To decide on routes a map would be helpful where truck movements are shown and coloured when estimations are critical."* (NE-2), *"I would prefer just a timeline and a dot to set in relation the estimated time-of-arrival of a truck."* (E-2).

Finding 2.7: Map-oriented status overviews of vehicles can improve the feeling for delays and distances

Similarly, support suggestions making routing decisions were found that included real-time maps to improve the feeling for distances and delays: *"A map would help. Then you could get a feeling for the distances between the warehouses and retail."* (NE-2). Several comments suggested that the map can be abstract with only moving dots but should also additionally have the stock levels displayed: *"It would help to have a bird's eye view on all the vehicles on the road and also to have the stock levels displayed on it."* (NE-3)

5 DISCUSSION

The GE research approach demonstrated its duality of knowledge creation in this case study. It shows its strengths in bringing together different perspectives on complex reference systems. Non-expert participants were able to engage in a rich discussion about support in supply chain decisions and reflect on practical improvements for learning through a gaming simulation.

As regards the *first research question* and thus the *practical findings* from analysing *reflection-in-action*, the different viewpoints and experience levels of the groups generated rich insights for abstractions and artificialities in the mapping between IL, GL and SL. While experts focused on identifying what learning additions could be supportive and on improvements of reality mapping, non-experts expressed more their suggestions for overview and control of the complex process. In other words, GE has demonstrated that it can detect and elicit practical improvements related to participants' expertise. The summarised suggestions provide *indications on how to improve the mapping of game simulations* for GE or learning purposes regarding decision making in complex systems:

1. Overviewing several roles introduces cognitive load but can be facilitated by comparative views
2. Reducing game goal interface elements support focus and shorten pre-game briefing
3. Feedback in-game and debriefing helps to transmit knowledge about decision making
4. Dynamic adaption of difficulty/complexity can help to learn from different decision scenarios
5. Unexpected incidents originating from stakeholders or environment can improve reality mapping and game experience

The benefits of having insights on novices and experts' perceptions became particularly obvious when looking first at answering the *second research question* and the *corresponding analytical results* of *reflection-on-action interviews after debriefing*. Although lacking experience in the referenced complex system, non-experts' suggestions on decision making improvements confirmed the recommendations of the domain experts. Both groups recommended aggregated, comparative data presentation for supporting delayed-effect decisions. Similarly, for decision support related to stock movement predictive trend indicators either as charts or numerical assistance were endorsed by both groups equally frequent.

On the other hand, our findings have also created arguments for including domain experts in a GE investigation. Non-experts suggested historical data as decision support much more often than experts. It is a good indication for viewing this suggestion with caution as it can promote experiential bias [26] and does not necessarily lead to better choices with delayed outcomes. The perspective of experience is further noticeable in the comments on decision support originating from vehicle monitoring. While non-experts were suggesting maps and overviews, experts were emphasising reduced visuals with dots on a timeline. A challenging, complex system already has considerable amounts of information to display. Reduced visual representations, therefore, can help to avoid extraneous workload and consequently improve reasoning and reflective decision making [12]. Conclusively, when synthesising the study findings, *humans are supported most in delayed-effect decision making by:*

1. Aggregated comparative views in user interfaces that allow customisation
2. Detailed indicators on decision-relevant quotas
3. Reaction triggers and continuous line-charts for process monitoring
4. Predictive trend indicators in line-chart or numerical form
5. Reduced visual aids to convey a sense of distances and delays

By applying GE as a research approach, distinctive strengths were identified that provide research trajectories for further investigation. In essence, GE is a successful technique for bridging experts and non-experts' perspectives in complex system investigations and creating a duality of knowledge with practical and analytical insights. Additionally, it creates an improved game artefact that can be applied for further research and educational purposes. Further analytical insights could also arise from including the game session outcome (e.g. timing, score) in the analysis.

However, these advantages are accompanied by *constraints* primarily originating from an intricate research process. To apply GE,

a resource-intensive, iterative development with contextual and theoretical grounding is required. Aside from the challenges of artefact development, data collection and analysis are demanding with two separately collected and analysed data sets. The high expenditure of time is thereby also affecting the participants. Finally, inherent dependencies between the three layers of interaction, game, and simulation need further investigation to facilitate analysis and interpretation of findings.

6 CONCLUSIONS AND FUTURE WORK

In this paper, we applied the novel approach of game elicitation to explore human-centred support strategies in delayed effect decision making. The two-step research process proved valuable in bridging perspectives of non-experts and experts while generating suggestions for improvements on a practical and analytical level. The recommendations were triggered through a supply chain activity with a disaster scenario disbalance. Identified support for decision making with delayed feedback, however, is oriented on the reduction of cognitive load as the underlying mechanism. Consequently, this suggests applicability in other domains of decision making with delayed effects or associated uncertainty such as privacy and data sharing. Nonetheless, further research is required to learn more about the influences on analytic findings from dependencies in the three layers interaction, game, and simulation. Ultimately, we propose to advance GE as a research approach by application in further domains and include collaborative scenarios for eliciting knowledge emerging from dialogical interaction.

ACKNOWLEDGMENTS

This research was supported by the Research Council of Norway (Norges Forskningsråd) by funding the IKTPLUSS project ALerT, #270969.

REFERENCES

- [1] David C. Bell. 1975. Simulation Games: Three Research Paradigms. *Simulation & Games* 6, 3: 271–287.
- [2] David E. Bell, Howard Raiffa, and Amos Tversky. 1988. *Decision making: Descriptive, normative, and prescriptive interactions*. Cambridge University Press.
- [3] Hartmut Blank, Steffen Nestler, Gernot von Collani, and Volkhard Fischer. 2008. How many hindsight biases are there? *Cognition* 106, 3: 1408–1440.
- [4] Philipp Brauner, Simone Runge, Marcel Groten, Günther Schuh, and Martina Ziefle. 2013. Human factors in supply chain management. In *International Conference on Human Interface and the Management of Information*, 423–432.
- [5] Sunil Chopra and Peter Meindl. 2007. Supply chain management. Strategy, planning & operation. In *Das summa summarum des management*. Springer, 265–275.
- [6] John Collier and Malcolm Collier. 1986. *Visual anthropology: Photography as a research method*. UNM Press.
- [7] Alexandre Dolgui, Dmitry Ivanov, and Maxim Rozhkov. 2020. Does the ripple effect influence the bullwhip effect? An integrated analysis of structural and operational dynamics in the supply chain. *International Journal of Production Research* 58, 5: 1285–1301. <https://doi.org/10.1080/00207543.2019.1627438>
- [8] Alexandre Dolgui, Dmitry Ivanov, and Boris Sokolov. 2018. Ripple effect in the supply chain: an analysis and recent literature. *International Journal of Production Research* 56, 1–2: 414–430. <https://doi.org/10.1080/00207543.2017.1387680>
- [9] Richard D. Duke. 1974. Toward a general theory of gaming. *Simulation & Games* 5, 2: 131–146.
- [10] Richard D. Duke and Jac Geurts. 2004. *Policy games for strategic management*. Rozenberg Publishers.
- [11] J. St BT Evans, Julie L. Barston, and Paul Pollard. 1983. On the conflict between logic and belief in syllogistic reasoning. *Memory & cognition* 11, 3: 295–306.
- [12] Jonathan St BT Evans. 2003. In two minds: dual-process accounts of reasoning. *Trends in cognitive sciences* 7, 10: 454–459.
- [13] Bent Flyvbjerg. 2006. Five misunderstandings about case-study research. *Qualitative inquiry* 12, 2: 219–245.
- [14] Jay Wright Forrester. 1961. *Industrial dynamics*. M.I.T. Press.
- [15] Gonzalo Frasca. 1999. Ludology meets narratology: Similitude and differences between (video) games and narrative. *Ludology.org*.
- [16] Roxane Gardner. 2013. Introduction to debriefing. In *Seminars in perinatology*, 166–174.
- [17] Jack S. Goodwin and Stephen G. Franklin. 1994. The beer distribution game: using simulation to teach systems thinking. *Journal of Management Development*.
- [18] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14.
- [19] Miriam Greis, Passant El. Agroudy, Hendrik Schuff, Tonja Machulla, and Albrecht Schmidt. 2016. Decision-Making under Uncertainty: How the Amount of Presented Uncertainty Influences User Behavior. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordICHI '16)*, 1–4. <https://doi.org/10.1145/2971485.2971535>
- [20] Wilmjakob Herlyn. 2014. The Bullwhip Effect in expanded supply chains and the concept of cumulative quantities. In *Innovative Methods in Logistics and Supply Chain Management: Current Issues and Emerging Practices. Proceedings of the Hamburg International Conference of Logistics (HICL)*, Vol. 18, 515–530.
- [21] Alan R. Hevner. 2007. A three cycle view of design science research. *Scandinavian journal of information systems* 19, 2: 4.
- [22] Jop van den Hoogen, Julia Lo, and Sebastiaan Meijer. 2016. Debriefing research games: context, substance and method. *Simulation & Gaming* 47, 3: 368–388.
- [23] Johan Huizinga. 1980. *Homo ludens a study of the play-element in culture*. Routledge & Kegan Paul, London.
- [24] Robin Hunicke, Marc LeBlanc, and Robert Zubek. 2004. MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI*, 1722.
- [25] Patrick Jost, Sue Cobb, and Isabella Hämmerle. 2019. Reality-based interaction affecting mental workload in virtual reality mental arithmetic training. *Behaviour & Information Technology*: 1–17. <https://doi.org/10.1080/0144929X.2019.1641228>
- [26] E. Ásgeir Juliusson, Niklas Karlsson, and Tommy Gärling. 2005. Weighing the past and the future in decision making. *European Journal of Cognitive Psychology* 17, 4: 561–575.
- [27] Jesper Juul. 2005. *Half-real: video games between real rules and fictional worlds*. MIT Press, Cambridge, Mass.
- [28] Maaz Hasan Khan, Salma Ahmed, and Danish Hussain. 2019. Analysis of Bullwhip effect: A Behavioral Approach. *Supply Chain Forum: An International Journal* 20, 4: 310–331. <https://doi.org/10.1080/16258312.2019.1661756>
- [29] Jan HG Klabbers. 1999. Three easy pieces: A taxonomy on gaming. In *International Simulation and Gaming Research Yearbook*. Kogan Page, 16–33.
- [30] Jan HG Klabbers. 2009. Terminological ambiguity: Game and simulation. *Simulation & Gaming* 40, 4: 446–463.
- [31] Jan HG Klabbers. 2018. On the architecture of game science. *Simulation & Gaming* 49, 3: 207–245.
- [32] Mykel J. Kochenderfer. 2015. *Decision making under uncertainty: theory and application*. The MIT Press, Cambridge, Massachusetts.
- [33] Florian Kohlbacher. 2006. The use of qualitative content analysis in case study research. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 1–30.
- [34] Willy Christian Kriz and Jan Ulrich Hense. 2006. Theory-oriented evaluation for the design of and research in gaming and simulation. *Simulation & Gaming* 37, 2: 268–283.
- [35] Udo Kuckartz. 2014. *Mixed Methods*. Springer Fachmedien Wiesbaden, Wiesbaden.
- [36] Udo Kuckartz. 2014. *Qualitative Text Analysis: A Guide to Methods, Practice and Using Software*. SAGE.
- [37] Julia Lo, Jop Van den Hoogen, and Sebastiaan Meijer. 2013. Using gaming simulation experiments to test railway innovations: implications for validity. In *2013 Winter Simulations Conference (WSC)*, 1766–1777.
- [38] Heide Lukosch and Tina Comes. 2019. Gaming as a research method in humanitarian logistics. *Journal of Humanitarian Logistics and Supply Chain Management* 9, 3: 352–370. <https://doi.org/10.1108/JHLSCM-06-2018-0046>
- [39] Analia Inés Meo. 2010. Picturing Students' Habitus: The Advantages and Limitations of Photo-Elicitation Interviewing in a Qualitative Study in the City of Buenos Aires. *International Journal of Qualitative Methods* 9, 2: 149–171.
- [40] Janet H. Murray. 2005. The last word on ludology v narratology in game studies. In *International DiGRA Conference*.
- [41] J. Nienhaus, A. Ziegenbein, and P. Schoensleben. 2006. How human behaviour amplifies the bullwhip effect. A study based on the beer distribution game online. *Production Planning & Control* 17, 6: 547–557. <https://doi.org/10.1080/09537280600866587>
- [42] Christoph Roser, Mizuho Sato, and Masaru Nakano. 2020. Would you like some wine? Introducing variants to the beer game. *Production Planning & Control* 0, 0: 1–9. <https://doi.org/10.1080/09537287.2020.1742370>
- [43] Bill Roungas, Femke Bekius, Alexander Verbraeck, and Sebastiaan Meijer. 2020. Improving the decision-making qualities of gaming simulations. *Journal of Simulation*: 1–14.

- [44] Adi Sagi and Nehemia Friedland. 2007. The cost of richness: The effect of the size and diversity of decision sets on post-decision regret. *Journal of personality and social psychology* 93, 4: 515.
- [45] Sourish Sarkar and Sanjay Kumar. 2016. Demonstrating the effect of supply chain disruptions through an online beer distribution game. *Decision Sciences Journal of Innovative Education* 14, 1: 25–35.
- [46] Herbert A. Simon. 1972. Theories of bounded rationality. *Decision and organisation* 1, 1: 161–176.
- [47] John Sweller, Paul Ayres, and Slava Kalyuga. 2011. *Cognitive Load Theory*. Springer New York, New York, NY.
- [48] Maggie E. Toplak, Richard F. West, and Keith E. Stanovich. 2011. The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks.
- [49] Eric Trist. 1981. The evolution of socio-technical systems. *Occasional paper 2*: 1981.
- [50] Amos Tversky and Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *science* 185, 4157: 1124–1131.
- [51] Margaret Verkuyl, Lynda Atack, Krista Kamstra-Cooper, and Paula Mastrilli. 2020. Virtual Gaming Simulation: An Interview Study of Nurse Educators. *Simulation & Gaming* 51(4): 537-549.