CAN WE LEARN FROM SIMPLIFIED SIMULATION MODELS? AN EXPERIMENTAL STUDY ON USER LEARNING[1]

Dr. Antuela Tako [2]

Loughborough University, Epinal Way, Loughborough, LE11 3TU, Telephone: 01509228219; Email: a.takou@lboro.ac.uk

Mr. Naoum Tsioptsias

Loughborough University, Epinal Way, Loughborough, LE11 3TU, n.tsioptsias@lboro.ac.uk

Professor Stewart Robinson

Loughborough University, Epinal Way, Loughborough LE11 3TU, s.l.robinson@lboro.ac.uk

ABSTRACT

Simple models are considered useful, especially in supporting a group of stakeholders to consider options and identify solutions to their problems in facilitated modelling workshops. This paper describes an experimental study that investigates whether the level of model detail affects users' learning. More specifically we aim to establish whether the learning achieved when using a simplified versus a more complex simulation model, differs. Our subjects, undergraduate students, were asked to solve a resource utilization task for an ambulance service problem. The participants worked in groups under three different conditions based on the type of simulation model used (specifically a simple, adequate or no model at all) to support their analysis and to reach conclusions about the action to be taken. A before and after questionnaire and a group presentation capture the participants' individual and group attitudes towards the solution. Our results suggest that differences in learning from using the two different models were not significant, while simple model users demonstrated a better understanding of the problem. The outcomes and implications of our findings are discussed, alongside the limitations and future work.

Keywords: Discrete-Event Simulation, Simple models, Complexity, Learning, Behavioural Operational Research.

1. INTRODUCTION

---

[1] This paper is an extended version of the paper: Tsioptsias, Tako and Robinson (2018) *"Can we learn from wrong simulation models? A Preliminary experimental study on user learning"* in Proceedings of the Operational Research Society Simulation Workshop 2018 (SW18), Anagnostou, Meskarian, and Robertson (eds.)

[2] Corresponding author.

Simulation models are often used to support organisations in decision-making and problem-solving, both in industry and the public sector (Luoma, 2014; Pace, 2004; Wahlström, 1994). However, little is understood about the level of model detail required to help users understand their problems and make decisions. Understanding the use of simulation models from the users' point of view was studied in Tako and Robinson (2009). Since then, with the emergence of Behavioural Operational Research (BOR), more interest has risen in understanding model acceptance from the users' perspective (e.g. Hämäläinen et al., 2013; Gogi et al., 2016; Monks et al., 2016).

This paper explores learning from using simplified simulation models. We consider simple models as opposed to more complex ones, such models that have been simplified to represent a far abstraction of the real system (Robinson, 2015). We focus specifically on discrete-even simulation models. Literature suggests that modellers may often prefer to build more complex models that may yield less accurate results, than simplified ones (e.g. Lödding et al., 2003; Silver, 2004; Bahtiyar, 2005; Lee et al., 2011; Ward, 2014; Cárdenas Duarte et al., 2017). There is a general belief that simplified models may provide approximate or less accurate outcomes (Larroca and Rodríguez, 2014; Nwodo and Okoro, 2015) or users may find them unrealistic or not credible (Brooks and Tobias, 1996). Using simplified models can affect the clients' perception of model validity and more specifically it may affect their credibility in using the model and its outcomes as they may consider the model to be inadequate or plainly wrong. Yet, it is believed that we can still learn from using simplified models (Hodges, 1991; Bankes, 1998; Morecroft and Kunc 2007). In a more recent review of OR modelling in healthcare from a behavioral perspective, Kunc et al. (2018) discuss that simple models can prove useful to clarify stakeholder conflict. However, to the best of our knowledge, there is no evidence to demonstrate the effect of simulation models' level of detail on users' learning. More recently, Katsikopoulos et al. (2017) consider the benefits of using simple versus complex models to support decision making, albeit this is more relevant to models using multi-criteria decisions analysis methods. Existing OR and simulation literature on the uses and learning from simplified models is scarce.

This paper aims to establish whether the level of model complexity affects the learning achieved by users of discrete event simulation (DES) models. We present an experimental study carried out with undergraduate students at Loughborough University to identify differences in participants' learning as a result of using a simplified versus a more complex version of the same model and/or no model at all. Our study aims to provide evidence on the usefulness of simple models. It is part of a wider study looking at the uses of wrong models and their role in supporting learning and decision-making. This work contributes to the existing behavioural operational research and simulation literature, as it provides evidence towards identifying the value of using simple models in supporting users from a learning point of view. This can be especially relevant to the existing facilitated modelling practice (e.g. Franco and Montibeller, 2010; Robinson et al., 2014; Tako and Kotiadis, 2015), where simplified models are normally used to interact with the clients and it may not be possible to revisit and rebuild a model at the workshop.

The paper is outlined as follows: Section 2 summarises existing work on the choice of level of model detail and on the evaluation of learning from using models in Operational

Research (OR) and Simulation. Section 3 presents the methodology, aims and hypothesis of the study, followed by the case study and the process followed, while Section 4 reports the results. Section 5 provides a discussion of the findings followed by the conclusions.

2. CHOOSING THE LEVEL OF MODEL DETAIL AND LEARNING

This section provides an overview of the existing OR and simulation literature discussing the level of detail represented in models and the impact of using simple versus complex models. We then review existing studies that evaluate learning as a result of using simulation models.

2.3 Level of model detail in OR and Simulation

The level of model detail can be represented by two mutually exhaustive concepts that of model simplicity and its opposite, complexity, which are two sides of the same coin. They are used interchangeably depending on the aims of the model and its intended use. Different authors consider one facet or the other, yet, the main argument is the importance of incorporating the adequate level of detail (simplicity or complexity) in the model (Innis and Rexstad, 1983). Modellers and/or clients may have different preferences either for a simple or a more complex model (Goldberg et al., 1990; Bruno and Halpern, 1999; Posada, 2003).

Simplification is the process of refining a model's level of detail by removing unnecessary elements (Innis and Rexstad, 1983; Carvalho et al., 2014; van der Zee, 2019). Simplification is considered a fundamental modelling activity in simulation (Salt 1993; Shannon 1998). It helps in creating simulation models that are useful and feasible, whereby the model focuses on system elements that matter as well as it avoids unnecessary study efforts. Many authors observe that model simplification has received little attention in DES (Sevinc 1991; Chwif et al. 2000; Brooks and Tobias 2000; Robinson 2006; Van der Zee et al., 2011, van der Zee 2019), evidenced by the low number of articles published on this topic. This is confirmed in a recent review on approaches for simulation model simplification (Van der Zee, 2017; 2019). Furthermore, van der Zee et al. (2018) identify that simulation educational support for mastering associated modelling skills is limited.

The principle of Occam's razor, also known as the law of parsimony, suggests that the simplest model possible should be created (Blumer et al., 1987; Shalizi, 2003; Baker, 2006).This is a guideline modellers are advised to use for developing good models (Nwodo and Okoro, 2015). Existing literature furthermore suggests that modelling should iterate, starting from oversimplified models to more refined ones (Brooks and Tobias, 1996; Schulze et al., 2017, Pidd 2010). Chwif et al. (2006) propose a complexity reduction technique that can be utilized during conceptual modelling in DES.

Developing extremely simple models suggests that vital components are excluded from a model's description, and the outcome is an oversimplified model[3]. Kaplan (1986) makes a distinction between simplifying to create a more elegant model as opposed to neglecting important factors. Robinson (2015) talks about "far abstraction" as a distinct -

---

[3] https://psychologydictionary.org/oversimplification/

excessive - simplification when creating the conceptual model that leads to a simulation model that does not resemble the real world. Since the conceptual model describes the way a system is abstracted, then a far abstraction could lead to lack of credibility (Robinson, 2015). A simpler model may by itself be sufficient to address a problem (Robinson, 2015). Building a simpler model suggests that simpler means shorter, more transparent, and, more efficient (Innis and Rexstad, 1983). It yields certain advantages, such as better understanding, faster analysis and modification, run time, which can result in improved implementation (Brooks and Tobias, 1996; Robinson, 2015). On the other hand, an oversimplified model may lack validity as the introduction of abstractions and different assumptions may make the reduced problem unrealistic (Brooks and Tobias, 1996).

Complexity on the other hand may refer to the processes, the data, the parameters, the outputs, the variables or the coding of a system (Innis and Rexstad, 1983), or it may as well refer to a system consisting of many parts that are dependent of each other (Shalizi, 2003). Complexity refers to the described length of a set of attributes of an entity (Gell-Mann, 1995; Gell-Mann and Lloyd, 2003). Adding complexity to a model means that it enables models to predict a wider range of outputs, while simplicity restricts the variety of model outcomes (Vanpaemel, 2009).

While the discussion so far is primarily opinion-based, with authors providing their opinions about the benefits of introducing simplifications or complexity into the model, to the best of our knowledge there is little empirical evidence addressing the impact of using simplified and/or complex models on users learning. For instance, a rather old comparison between a simplified DES version and a complex real-life computer system showed unexpected similarities in model results (Dickert and Zanakis, 1979). More recently, Akpan and Brooks (2014) compare the use of a simple 2D and 3D visualisation of a DES model of assembly and service operations and conclude that error spotting is easier with the 2D visualization. It is noted however that the latter paper focuses on the users' capability of spotting errors between the two models rather than comparing how these models affect the users' perspective on accepting and making decisions based on each model's level of detail. Studies in OR mainly in the field of forecasting, explore the idea of simple versus complex models, focusing mainly on which produces better model results (e.g. Green and Armstrong, 2015). Katsikopoulos et al. (2017) consider simple and complex models in OR, focusing primarily on multi-criteria decision analysis models. They provide guidelines to help decide when simple models should be used. They suggest the use of simple models for frequently repeated rather than one-off strategic decisions and conclude that simplified models can approximate complex models, so long as they include the appropriate level of simplification.

A similar discussion on simple and complex models is also found in SD literature. For example, Elsawah et al. (2017) report cases of simplified and complex models from SD and highlight that a model should be as simple as possible and as complex as needed. They make suggestions on how to handle this balance when working on SD models and refer to the recurring criticisms of unnecessary complexity used in models. Similarly, Newell and Siri (2016) discuss how complex problems may not be easily addressed by traditional approaches but instead require the mapping of a simpler example in the form of a conceptual source that would allow a "*conceptual metaphor*". They introduce low-order

SD models as a potential method for modelling and understanding the impacts of such complex decision-making domains. Saleh et al. (2010) use a simplified version of a manufacturing model in SD and find that the results are similar to those obtained from the original model that had extensive experimental runs. Soo et al. (2019) take a similar view on simplicity. Homer and Hirsch (2006) use a simple SD model to identify possible strategies for disease prevention in a health environment. They however argue that the model would require more details and factors added to identify more specific actions for decision-making. Similarly, Morecroft and Kunc (2007) suggest using "back-of-the-envelope" models to facilitate decision making in a simple way.

In this section we establish that current OR and simulation literature on the use of simplified models is limited. The use of simplified models is however supported, so long as the simplifications do not affect the credibility of the model or its results from the users' point of view. We next consider how learning is evaluated in existing simulation literature.

2.4 Learning from using simulation models

We start from the premise that models, regardless of their level of detail, are believed to provide learning opportunities for their users (Hodges, 1991; Bankes, 1998). Learning produces a change in behaviour by affecting someone's observable action (Schacter et al., 2014). Based on Argyris and Schön (1996), one can deduce that learning is achieved if a change in users' existing knowledge, attitude or decisions occurs as a result of interacting with a model.

Few studies in DES explore learning resulting from using simulations. Tako and Robinson (2009) first referred to the learning achieved from using simulation models, providing evidence that both discrete event simulation and system dynamics models can equally aid instrumental learning. Monks et al. (2014) test differences in learning when using an existing model compared to when using a model in which clients were involved in building. They find that the participants who used an existing model learnt more compared to those involved in building and using that model because the latter had less opportunity to experiment with different scenarios. They also observe that it is difficult to establish the learning occurred as a result of using simulation models (Monks et al., 2016). Gogi et al. (2016) test whether learning insights can be gained from using discrete event simulation. Through their study they look to establish whether statistical results or animation of simulation models have a higher impact on the generation of insight. Their results show that a higher frequency of insights is gained from using simulation model statistics.

In the System dynamics (SD) field, the concept of learning is widely used and models are considered as learning tools. Business flight simulators are also called 'learning laboratories' that can help managers gain insights about their business operations (Forrester 1961, Morecroft and Sterman 1994). For example, Thompson et al. (2016) explore the critical learning incident that occurs as a result of the mental models that clients and consultants were engaged in developing. They report that clients did not recognise any critical learning incident despite the fact that their confidence increased due to their engagement with the modelling process, resulting in mental model change. Elsawah et al. (2017) focus on increasingly complex SD management flight simulators to explore again critical learning incidents. They find that simulators helped in understanding the SD model, subject to several limitations.

From the work discussed above, it can be surmised that few studies explore learning from using simulations. Of these we have not identified any studies focusing on learning as a result of the model's level of detail, referring to either complex or simple models. More recently Kunc et al. (2018) discuss OR modelling in healthcare from a behavioural perspective, making a call for more behavioural studies that measure the impact of changes in behaviour as a result of interacting with models and taking part in the modelling process (Kunc et al., 2018). The current study compares simulation users' learning and credibility based on treatment conditions related to the level of model detail. This in turn helps to develop empirical evidence about the value of simple models in practice as we believe that this is an important aspect that the simulation modelling community would benefit from. We note that we focus more particularly on instrumental learning, considering the model users' understanding and change of attitude on a particular managerial problem. We do not visit double loop learning defined as the ability to transfer acquired learning or skills in equivalent situations (Argyris & Schön, 1996), nor learning from a pedagogical perspective. For more details of developing simulation models for educational purposes the reader is referred to Van der Zee et al. (2012), who provide detailed guidance on conceptual modelling for simulation-based serious games for training on logistics type of problems.

2.5 Summary on level of model detail and learning

In summary, based on the above, existing literature comparing the use of simplified versus more complex models is limited, focusing primarily on general guidance suggesting that simple models are useful. There is furthermore the argument that by reducing the level of model detail (simplicity) or increasing the level of model detail (complexity), may affect the level of credibility placed into that model from the users' point of view. Credibility is the perception from the clients' point of view that the model or its results are sufficiently accurate for the purpose at hand (Gass, 1993; Pace, 2004; Robinson, 2014). This is one aspect of model validation (Balci 1994; Gass, 1993; Pace, 2004), where the model is checked to ensure that the client and users accept the model and its results and that it is representative of their perception of the real system (Pace, 2004).

To the best of our knowledge, there is no empirical evidence to suggest that using simplified or less complex models can achieve better learning outcomes. To address the current gap in the literature, our study is looking to establish whether learning can be achieved from using simple models and whether the level of model detail can affect learning and credibility from the users' point of view.

3. METHODOLOGY AND EXPERIMENTAL DESIGN

This section describes the experimental study, including the research questions and hypothesis, study design, the case study, and the simulation models used.

3.1 Study objective and hypothesis

The aim of this research is to identify whether the level of model detail affects learning and credibility from the users’ point of view. We expect that less complexity can affect users' perception of the model. However, increasing a model's complexity doesn't necessarily mean that the model is more accurate (Robinson, 2014). Indeed, Green and Armstrong (2015) reviewed studies that compare simple and complex methods and found that complexity is preferred for reasons other than improving models' forecasting accuracy. Instead they suggest that complexity seems to increase the occurrence of errors. Katsikopoulos et al. (2017) suggest the use of simple models in certain cases such as frequently repeated rather than one-off strategic decisions, still, even in the latter case, they can approximate complex models.

We compare the impact of using a simplified model versus an adequate (slightly more complex) model on the learning and credibility from the users’ point of view. We assume that learning occurs as a result of a change in people's attitudes towards a belief. This can be demonstrated by examining their perceptions about the solution to a problem. As such we look to test the following study hypothesis:

"The use of simple and complex simulation models offers the same learning outcome."

We expect that users may find simplified models easier to use and hence we look to establish whether users gain a better understanding of the problem and solutions. Our aim is to test statements found in the literature (Katsikopoulos et al. 2017; Green and Armstrong, 2015) and to provide further evidence relevant more specifically to simulation models.

3.2 Study design

An experimental study is developed to test the study hypothesis. Final year undergraduate students, 58 in total, taking a simulation module "Simulation for Decision Support" at Loughborough University took part in the experiment. The experiment took place as part of a three-hour lecture. All students were familiar with basic simulation modelling and had undertaken a placement year in a company in their third year of studies. They were randomly allocated into groups of 6-7 students, forming 9 groups in total. Each group was assigned randomly to one of three treatment conditions, based on the type of model used:

- adequate model (AM) - a relatively complete or requisite simulation model,
- simple model (SM) - a simplified simulation model, and,
- no model (NM) – no model was provided, instead this group was asked to create a conceptual model of the problem, consisting the control groups.

The participants were not aware of any differences in the model or condition they were allocated. Table 1 summarises the assignment of the 9 groups into treatment conditions:

Condition	Explanations	Groups assigned
Adequate Model (AM)	Adequate model	2, 3, 4
Simple Model (SM)	Simplified model with approximate outputs for time targets	1, 5, 6
No Model (NM)	No model is used, participants develop a conceptual model of the problem	7, 8, 9

Table 1 *Group assignment to treatment conditions*

The process took place in the following order. An overview of the problem and tasks was provided. The participants were then given the case study (section 3.3) to read along with a pre-test questionnaire to complete individually that captures participants' initial attitude towards the problem and the managerial decision (potential solution). Next, they were randomly split into groups using systematic sampling. Groups 1 to 6 were each provided with a notebook computer with the assigned version of the model (simple or adequate). Further paper-based instructions were given to each group based on the treatment condition they were assigned to. The students were asked to work in groups and to use the PartiSim tools and guidance (Kotiadis and Tako 2010) to support their group discussions and the facilitation process (Tako and Kotiadis 2015). The students were then asked to work on the task for approximately 1.5hrs and to prepare a presentation with their recommendations towards the available option for the managerial decision (more details are provided in Section 3.4). During this time, the researchers offered clarifications about the task to individual groups, as required. At the end the students re-assembled in the lecture room and were asked to complete an individual post-test questionnaire, in order to capture attitude changes towards the problem and the solution. The questionnaire was the same for all groups, with the exception that the questionnaire for the No Model groups did not included two questions related to the model. Lastly, the students presented their group findings to the whole class. Questions were asked by the researchers to clarify their recommendations towards the decision and the prioritisation of targets.

3.3 The case study

The case study is a resource utilisation task for an ambulance service adapted from a master's dissertation project (Puntambekar, 2016). The case initially describes the problem and the call cycle. Incoming calls are classified as emergencies (life-threatening) or urgent (non-life-threatening). When answering a call, the operators assess the severity of patients' condition and decide on the route to be followed. Regardless of call type, a proportion of calls is redirected to the Clinical Assessment Team (CAT) for re-evaluation. The majority of calls result in the patients being transported to the local accident and emergency department (ED) or to alternative pathways (e.g. community care services). In some cases, the ambulance crew provides clinical treatment on scene and the patient may not need to be conveyed to ED. Treatment may be also provided over the phone by clinical advisors. This helps to avoid the dispatch of ambulances to patients who do not require an ambulance. During the winter months, the ambulance service faces a higher number of calls, which affects the service's ability to deal with incoming calls within specified time targets.

Management think that a lot of patients are unnecessarily being taken to A and E and are examining the option of increasing CAT intervention with the view to reducing use of ambulances and to release resources for patients that require emergency transportation. Three alternative solutions under consideration by management were provided: keeping the percentage the same to 30%, increasing it to 40% or further to 50%. Either option was described as possible in the case study. The task asked participants to comment on this managerial decision and recommend the percentage of patient calls that should be redirected for re-evaluation to the CAT team. To check which answer would be the most suitable, certain targets were set that the participants would have to meet and prioritise. These targets (in order of importance) were:

- time targets for non-life and life-threatening patients - % of calls responded within the required time (8 mins for urgent and 30 mins for emergency calls)
- cost expenditures for personnel and ambulances – in the form of a maximum total
- ratio of number of patients treated in ED versus those treated in alternative pathways.

The full case study is available by the authors upon request.

3.4 The two variations of the model and possible solutions

The model used was adapted based on a model created as part of an MSc dissertation project in Business Analytics and Consulting through facilitated workshops with employees of the ambulance service (Puntambekar, 2016). The model was modified and financial variables added. The Simul8 software (SIMUL8 Corporation) was used to develop the models. Two main variations of that same model were selected: a more complex model that is considered requisite and termed here as "adequate" (Figure 1) and a simplified one termed here as "simple" (Figure 2). It should be noted that the adequate version of the model used for the experiment, has been developed working closely with ambulance personnel in real life facilitated workshops, which has been accepted and considered valid by the ambulance service concerned (Puntambekar 2016). Hence, the adequate model is considered the reference point from a validation point of view.

The simplified model was built from the adequate model. From a comparison of Figs 1 and 2, the reader can notice that the simplified model includes less details. A number of simplifications were introduced to the simplified model, where some variables, parameters and working stations have been taken out and the routes in the call cycle are simplified. For example, transport of patients to an alternative care provider in the adequate model contains an activity centre called "alternative request" that routes patients out to transport to alternative care or returns to ED, which is omitted in the simple model, connecting directly patient transportation to ED or alternative care. The types of available resources were also reduced from three (CAT personnel, First Vehicles on Scene, and Last Vehicles on Scene) in the AM to two (CAT personnel, and Vehicles) in the SM. As a result of these simplifications, the numerical outcomes approximated those in the adequate model, but were not exactly the same.

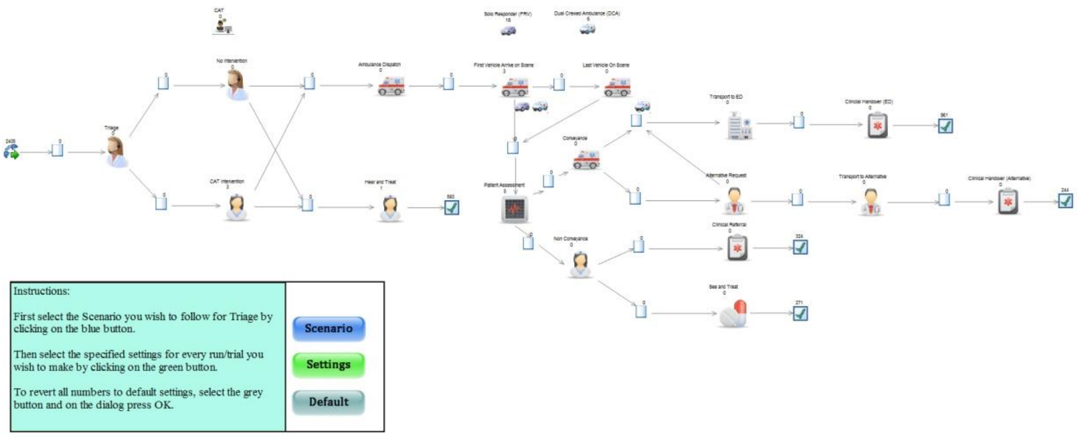

Figure 1 The adequate simulation model

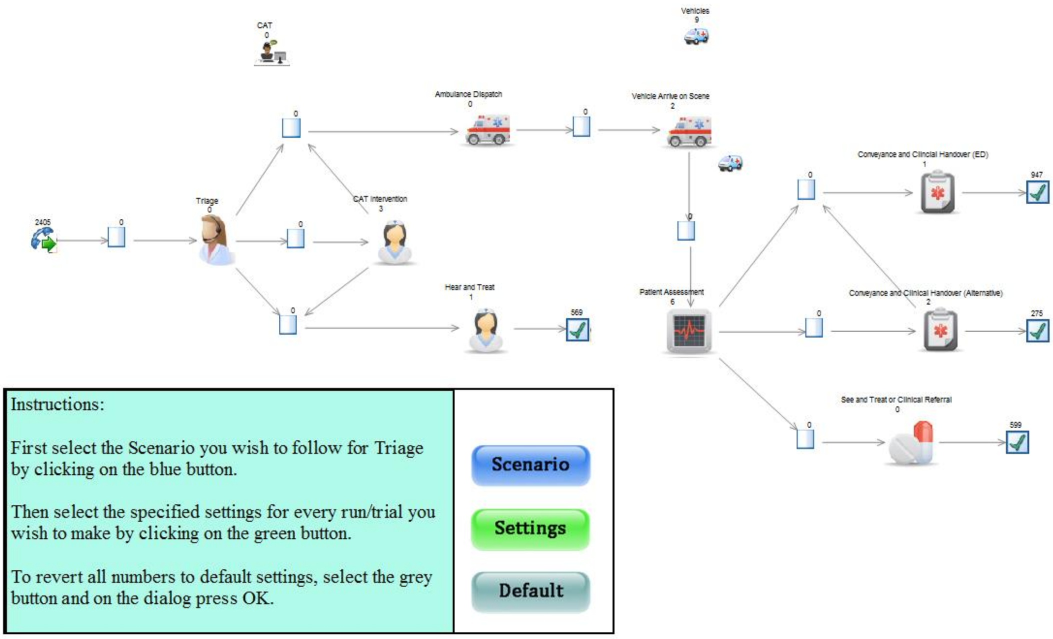

Figure 2 The simplified version of the simulation model

To demonstrate a change of attitude (leading to learning), the participants would need to understand and choose after experimentation with the model that the initial setting (30%) is the answer that offers a solution to the model without extreme effort. They did not necessarily need to meet all the targets set, however they would need to be aware that some were more crucial to meet such as percentage of calls responded to within the 8 minutes target.

Table 2 presents the set up of the models for all the alternative options , alongside our expectation as to whether a solution can be found(column “Solvable”), the expected time to reach this conclusion (column "Time required") and an interpretation as to how a solution could be derived in each case. The latter two columns are related to our expectation as to whether a solution is expected to be provided by the participants in column “Solution Expected?”. Wherever a low amount of time is needed to find a solution, an answer is expected (noted Yes in the table), if a high amount of time is expected than an answer is not expected (noted NO in the table). This in turn can help us identify whether a correct solution is reached by the participants, either as a result of experimenting with the model or through guessing or sheer luck. This is noted in the column “Interpretation”. The rationale for reaching this conclusion is next explained.

As it can be seen in table 2, for both models for the initial alternative decision (30%) a solution can be achieved, and it is possible to achieve all the targets within the available time. The alternative decisions of (40% and 50%) for the adequate model either requires too long to experiment with and to find a solution or it is not feasible to find a solution and achieve all the targets. Using the adequate model, the problem can be solved for these alternatives , but it also requires longer to consider all different parameters (settings), thus the column "Time required" is noted as high for reaching a solution. If a solution for these two alternatives is found by the participants, but key targets (time related targets) are not met, it is considered as reached either through guessing or by luck(column "Interpretation" in table 2). This is because finding the right answers within the available time is not feasible through this model set up.

On the other hand, using the simple model, it is not possible to solve the problem if the 50% alternative is chosen, and also key targets (time related targets) are not possible to meet. The 50% alternative can be solved through mere guessing as it is not possible to find a solution that achieves all the targets within the expected time frame. A solution for the 40% alternative can be reached by luck or through guessing, in the same way as for the case of the adequate model above.

Type of model	Alternative decision	Solvable	Time required	Solution expected?	Interpretation
Adequate	30%	Yes	Low	Yes	Through experimentation
	40%	Yes	High	No	Luck or guessing
	50%	Yes	High	No	Luck or guessing
Simple	30%	Yes	Low	Yes	Through experimentation
	40%	Yes	High	No	Luck or guessing
	50%	No	N/A	No	Guessing

Table 2 Expected solutions and interpretation of solutions for each type of model

3.5 Pilot study

To test the experiment's feasibility, 3 pilots were run individually with second-year PhD students in order to check the model, questionnaires and the instructions. None of them was familiar with simulation, however we did not consider this as an issue because we only expected them to use the models and further guidance would be provided by the researcher if needed. Each participant was randomly assigned to one of the two simulation groups, where 2 individuals worked with the simple and one with the adequate model.

The participants found the models functional and they showed a good understanding of the models and the case study. They used the models assigned to them and experimented with the different alternatives and parameters. No specific issues were reported. They also completed the pre- and post- questionnaires. I was observed that the overall process was intuitive and easy to follow. When comparing the participants' responses to the questionnaire, there was no change of attitude noted, however this is possible given the small sample used (n=3). In the conversations held subsequently, the participants seemed to comprehend that a change in their initial answer (thus keeping the initial 30% option) would make sense.

Based on the feedback received, a few minor changes were made in the description of the case study to further highlight key parts of the case study, such as the different options related to the managerial decision. Also the language used in the case study was further simplified and the text slightly reduced to make sure it did not take participants too long to read.

4. RESULTS

In total, 39 pre-test and 45 post-test questionnaires were returned from the 58 initial participants. The 6 additional post-test responses were excluded from the analysis, as we were not able to make a like for like comparison. Of these, the NM group had 10 students, SM 16 and the AM the remaining 13 students. All students were 21-23 years old (but one who was 25). The rest of the demographics are provided in Table 1. The distribution of abilities and marks is representative amongst the different treatment groups.

Groups	**Participants**	**Gender**		**Reported degree grades**		
		M	**F**	**1st**	**2:1**	**2:2**
No Model (NM) - control	10	15.4%	10.3%	7.7%	12.8%	5.1%
Simple Model (SM)	16	25.6%	15.4%	5.1%	28.2%	7.7%
Adequate Model (AM)	13	28.2%	5.1%	7.7%	20.6%	5.1%
Total	39	69.2%	30.8%	20.5%	61.6%	17.9%

Table 3 *Participant demographics based on those that handed over both questionnaires.*

The results are next analysed. We first present the results based on the solutions provided regarding the managerial decision (Section 4.1), then the analysis comparing the Likert scale questions of the pre- and post- questionnaires (Section 4.2), and finally the outcomes from the presentations (Section 4.3).

4.1 Analysis of responses about the managerial decision

To test the study hypothesis, we compare participants' answers regarding the managerial decision as part of their suggested solutions as provided before and after the task. If there is no significant difference between the users of SM and AM, then the research hypothesis is supported. We also compare the two simulation conditions with NM to establish whether there is a difference between using the model at all, as a means of checking that the case study and model work. If the simulation users demonstrate a shift in their attitudes towards the solution as opposed to NM, then we can support that the case study and model work. We use Pearson's Chi-square and Fisher's Exact Test for the comparisons as the variables are considered ordinal (Bryman and Cramer, 2011).

The participants are asked both in the pre- and post-questionnaire to answer whether and why they agree, disagree, or are not sure about the proposed managerial decision to increase the percentage of patients that are redirected to the CAT team for re-evaluation. The answers to this question before and after using the models are compared to establish whether a change in participants' perception of the solution occurs. . Out of the 39 valid answers, 24 agreed and 15 were not sure, while none disagreed with the intended increase in the managerial decision in the pre-questionnaire. In the post-questionnaire 21 agreed and 10 disagreed, while 8 only were not sure about the managerial decision.

We identify three possible options occurring as a result of the treatment condition i.e. after the participants have used the models:

- Positive change of attitude, when the participants' thinking improves indicating that they would move to a better decision against choosing the 50% alternative or at least expressing uncertainty about the increase of this percentage. Our premise is that if students alter their initial views from "Agree" or "Not Sure" to "Disagree", or from "Agree" to "Not Sure" then they would have acquired a change in attitude through this process. This change in the participants' attitude as corroborated by their answers, would suggest a change in their beliefs, and subsequently lead to the occurrence of learning about this problem (Argyris and Schön, 1996; Schacter et al., 2014).
- No change of attitude, when the participants thinking does not change, indicating that it does not lead to learning. This considers the cases when a participant continues to agree, disagree, be not sure both before and after the experiment.
- Unlearning, in the case that participants change their attitude to the solution for the worse, representing the case of moving from disagree to agree or not sure or from not sure to agree.

The participants' change of attitude for each group is summarised in table 4 below. The column "Change of attitude" presents the number and proportion of participants that experienced a positive change, which suggests that learning occurred after the treatment, while the column "No change of attitude" presents the number and proportion of participants that did not experience of change and the column "Negative change" the number and proportion of participants that experienced a negative change of attitude towards the solution. The proportions represent the percentage calculated out of the total 39 participants.

	Group		Change of Attitude		No change of Attitude		Negative change	Total Participants for each group
1	No model groups	0	0%	7	18%	3	8%	10
2	Simple model groups	7	18%	7	18%	2	5%	16

3	Adequate model groups	8	21%	3	8%	2	5%	13
4	Total	15	38%	17	44%	7	18%	39

Table 4 *Change of participants' attitude based on comparison of pre- and post-treatment answers.*

A shift in the participants' answers is noted. Out of the 39 participants, 15 (38.5%) changed their initial views towards the solution. More specifically, 8 participants were from the AM groups, 7 from the SM groups, and no one (as expected) from the NM groups. This means that 15 out of 29 (51.7%) simulation group participants were able to find the appropriate solution to the problem, with regards to the alternative managerial decision (i.e. keep the percentage at 30%). No change of attitude was experienced by 17 participants, the majority of which were in the NM and SM groups. Unlearning occurred for 7 cases, out of which, 3 belonged to the control (NM) group, 2 to SM, and 2 to AM group. This occurred only for the case of moving from "Not Sure" to "Agree" to the 50% alternative, which is interesting to note that there was not someone who categorically disagreed with the 50% alternative before the experiment. Considering the distribution of answers in table 4, it can be noted that there is a difference in the change of attitude between the simulation model groups compared to the NM (control) group. Furthermore, it appears that a higher proportion of the adequate model group has experienced a positive change of attitude compared to the NM (control) and SM groups.

To test the above observations, we perform a Chi-square test to compare the proportion of participants that change their attitude about the managerial decision between the three different groups. Due to the small sample size, we also report Fisher's exact test (Table 4) (Bi and Kuesten, 2015). This confirms that there is a difference in the change of attitude that occurs within the NM groups and those that used a simulation model (Pearson's Chi-square p-value = 0.004, Fisher's exact test p = 0.006), meaning that simulation affected the answers provided. It however does not reveal a statistically significant difference between the AM and SM user groups (Pearson's Chi-square p-value = 0.340, Fisher's Exact Test p-value = 0.462) supporting the research hypothesis. Table 5 summarises these results.

Group comparison	**Pearson's Chi-Square p-value**	**Statistically significant difference?**	**Fisher's Exact Test**	**Statistically significant difference?**
NM vs SM and AM	0.004	YES	0.006	YES
NM vs SM	0.014	YES	0.023	YES
NM vs AM	0.002	YES	0.003	YES
AM vs SM	0.340	NO	0.462	NO

Legend: NM = No Model, SM = Simple Model, AM = Adequate Model

Table 5 *Statistical analysis at 5% for the managerial decision*

The explanations provided by the participants about the proposed managerial decision to increase the percentage of patients that are redirected to the CAT for re-evaluation in the pre- and post-questionnaire, helped the researchers to explain the differences between the groups. On one hand, those agreeing with the managerial decision for increasing the percentage explained that it would help in creating a faster, better, and cheaper system, as explained in the case study. It should be noted, however, that certain participants were "not sure" of the managerial decision in the pre-test questionnaire as they doubted the description provided by commenting as follows: "change may not have an impact" or "not enough information is provided". On the other hand, participants who changed their attitude towards the managerial decision and disagreed or stated that they were not sure about the increase, commented mainly about the numeric outcomes of their experiments, for example: "we can reach all targets without increasing …", or about the model's simplifications, a need for further data, or that the model may already work without needing any changes. These answers advocate that the research hypothesis is supported suggesting that a change of attitude had occurred as a result of using the models, either simple or adequate.

4.2 Analysis of Likert-style questions

The post-test questionnaire included a number of 5-point Likert scale questions where participants were asked to rate their level of confidence in the model as well as their opinion about their understanding of the model, model representativeness and trust in results. A 5-point Likert scale was used. The analysis involves a comparison of the differences in the answers to establish whether there is a statistical difference in the opinions of participants between the two simulation groups. The Mann-Whitney test is considered appropriate to use since we consider the variables to be ordinal (Bryman and Cramer 2011). The results are presented in Table 6 below.

Opinions variables about the model	Mann-Whitney test p-value	Statistically Significant difference?
Confidence	0.931	NO
Understanding	0.098	NO
Representativeness	0.055	NO
Trust in model results	0.105	NO

Table 6 *Statistical analysis comparing Likert-style variables for the model treatment groups (simple and adequate) at 5% confidence level using the Mann-Whitney test.*

The results show that there is no statistically significant difference between the two treatment groups, those using the adequate and the simple model, in terms of confidence in the models used (p-value= 0.931). Similarly, we compared the participants' opinions about their understanding, model representativeness and their trust in model results for the two simulation model treatment groups using the Mann-Whitney test. Similarly, we found no statistical difference with regards to understanding, suggesting that there was not a better or worse predisposition towards either model (p-value = 0.098). Considering participants' trust in model results, again no difference is present between the two groups (p-value = 0.105). Participants' opinions about model representativeness show are not significantly different, although this is borderline suggesting that the participants' perception about the model may have been slightly different between the two treatment conditions (p-value = 0.055), where the Man-Whitney mean rank for the simple model was higher than that for the adequate model (17.53 compared to 11.88). This suggest that the simple model was found to be slightly more representative by the participants compared to the adequate model. Further considerations about this are made also in the qualitative analysis above (see Section 4.1) and the group presentations where participants from the simple model groups suggested they needed a better model, while 2 out of 3 groups said the models are not very representative but they could work with them (see more in Section 4.3).

In addition, we consider the participants' confidence in their answers before and after completing the task. A Wilcoxon signed ranks test is used to compare the results of participants of the two simulation groups. We find no statistical significance for those using the simple model (p-value = 0.129) and neither for those using the adequate one (p-value = 0.470). Surprisingly, when comparing the before and after answers of the NM users, we find a statistical significance (p-value = 0.037) which could be attributed to a level of overconfidence from the participants due to the fact that they had not used any simulation model, but only created a conceptual model of the described case study. Table 7 summarises the findings of the analysis.

Before and after difference in confidence per group	**Mann-Whitney test p-value**	**Statistically significant difference?**
AM	0.470	NO
SM	0.129	NO
NM (control)	0.037	NO

AM: Adequate Model, SM: Simple Model; NM: No model

Table 7 *Statistical analysis at 5% comparing the before/after confidence of individuals in each treatment group.*

4.3 Analysis of observations made based on the group presentations

At the end of the session, participants were asked to present their findings in their groups, reporting the analysis carried out, the scenarios they experimented with, their recommendations to the management team regarding the most appropriate decision chosen and to explain how they reached their conclusions. They were also asked to report on the targets reached while experimenting with the models. The NM group, who did not have access to a model, instead of discussing the targets, they were asked to present and explain the conceptual model developed. A prize incentive was offered for the best two presentations. The outcomes and main points discussed in the presentations by the three treatment groups are summarised in Table 8 below. The main points considered relevant are:

- Correct solution (row 1) – recommendation about the managerial decision, i.e. decision to keep the initial 30% routing to CAT intervention (as discussed in section 3.4 above).
- Learning Process (row 2) is used to assess the process the groups followed in achieving the solution. More specifically, for the adequate and simple model groups we check as to whether the group considered the importance of certain targets (such as the 8-min response target for urgent calls) over others and whether they focus on specific targets. In the case of NM, this criterion is instead applied to their solution to conceptual model instead of the answers to the problem. As a side note, it was not expected that the participants would randomly guess that the 30% alternative is the solution as there was no indication in the case study that this is the

solution. Instead, the 30% alternative would only be expected as an answer only if one experiments with the model and learns from the outcomes and the process, regardless of whether all the targets are met.

- Model evaluation (row 3) identifies whether the SM groups inherently comment about the simplicity or aggregated results present in their model.

Criteria/condition	No Model	Adequate Model	Simple Model
Correct solution	0 out of 3	1 out of 3	2 out of 3
Learning Process	2 out of 3 developed a representative conceptual model	0 out of 3 achieved all targets	0 out of 3 achieved all targets
Model evaluation	N/A	N/A	2 out of 3 (needed a better model)

Table 8 *Summary of the presentations*

The groups' performance was rated by two of the authors based on answers' insightfulness and general format of each presentation. Overall, the presentations showed a good understanding of the problem and the task.

All three NM (control) groups worked on creating a conceptual model that would represent the case study, with 2 out of 3 creating a very representative one (judged against the adequate simulation model).

The AM groups presented controversial results. Although the use of simulation offered more details to aid their understanding about the system, their final answers were not based on solving all targets. 2 out of 3 groups suggested that the managerial decision to increase the proportion of patients directed to CAT intervention should be 40%. This means that instead of trying to solve the initial problem through minor changes of parameters they redirected their attention to the alternative options of increasing the proportion of patients directed to CAT intervention to either 40% or 50%. Only one group chose the correct final solution and adequately justified their choice.

Similarly, the SM groups demonstrated a good understanding of how the system works. Still, no group met all the targets set, but 2 out of 3 groups (namely, groups 1 and 5) recommended that management should not change the current percentage for the managerial decision. This suggests that using a simpler version of the model the participants got the right answer with more ease. Another interesting outcome from these groups' presentations was that 2 out of the 3 groups commented that their model was not representative, which they considered as the reason for not being able to meet the targets. The members of group 5, specifically, reported that they felt a lot of information was missing from the model they were given. We note that they were not aware that they used a simplified version of the model. This suggests that they considered the simplified model given less credible, we however note that it was still adequate to find a solution to the problem, given that 2 out of the 3 groups were able to do so.

5. DISCUSSION OF RESULTS AND LIMITATIONS

This section discusses the outcomes of the study and its contributions to the existing simulation literature on the impact of model level of detail on learning. We then consider the limitations of the study and potential for future work.

5.1 Study outcomes

Our study set out to identify whether a simulation model's level of detail affects the learning achieved by users. Our study hypothesis tests out whether a simplified simulation model can offer the same level of learning as a more complex (adequate) one. Through the analysis undertaken we make a few observations regarding the use of simplified versus more complex models.

In support of our study hypothesis, the users of the simple and the adequate model demonstrate similar levels of change of attitude towards the solution. The results show that a significant number of participants from both treatment conditions provided correct solutions (8/13 adequate model users and 7/16 simple model users). This suggests that both models were useful in helping participants understand the problem. The statistical analysis supports our hypothesis that there is no significant difference in the users' confidence in their answers, their confidence in the results and their understanding of the problem.

From our qualitative analysis of the group presentations, we found that one adequate model group and two simple model groups came up with the correct answer, while two of the simple model groups mentioned they found their model was lacking details but could still work with it. A relevant observation is that in our experiment the simple model seemed easier for participants to handle especially due to the limited timeframe required to interact with the model. This outcome justifies our initial expectations as well as corroborates the opinions found in the literature. For example, Robinson (2015) posits that a simple model can help model understanding as well as support decision making and implementation. Model simplification is not a widely researched field as identified in a recently published review (Van der Zee, 2019) and our understanding of and guidance in developing simple models is limited.

Furthermore, model users found the models used not different in terms of representativeness as a marginally statistically insignificant result was obtained (p-value 0.055%) suggesting that, despite of the model used (simple or adequate) they were considered equally representative. The qualitative analysis corroborates this finding that the simplified model was not considered credible, but that it was still helpful in finding the correct solution. Another finding from this experiment is that simulation helped model users to demonstrate a shift in their attitudes towards the solutions compared to the control groups. Despite confidence and understanding were not found to be significantly different for either the no model or the simulation user groups before and after the treatment, it can be still concluded that simulation assisted the groups' understanding during the task.

The findings of this paper suggest that users' understanding, and behaviour can change as a result of using a simulation model, which also confirms the findings of a previous study by Gogi et al. (2015). Furthermore, the results of our study provide evidence to support that using simplified simulation models can result in similar levels of change in attitude, subsequently leading to learning, compared to using a more complex model. This is an interesting finding that provides some initial empirical evidence about the benefits of using simplified models. It is however noted that from the modellers' point of view, approximations need to be carefully chosen to ensure that using a simplified model does not affect the users' credibility into the model and its results. As shown in this study, the participants in the different model (treatment) groups did not find the models used significantly different.

Our paper is the first study to consider the effect of using simulation models at different levels of detail on users' learning and understanding about a problem situation. We find that so long as the model provides an adequate representation of the problem and it includes the key outputs relevant to the decision to be made, it can help its users to form opinions about the best courses of action. Obviously, we appreciate that simplified models provide approximate results, which can be interpreted as indicative as opposed to accurate results. Depending on the context or aim of the study, simple models can help to gain an overall understanding of the different courses of action, without requiring a huge amount of effort. Simplified models are often and typically used in facilitated simulation workshops, where the emphasis is on learning about the indicative behaviour of complex systems in a simple way rather than focusing on accurate results (Robinson et al 2014, Robinson 2015, Tako and Kotiadis 2015, Kunc et al 2018).

The findings of our study have implications to existing simulation theory and practice and more specifically to facilitated simulation, where models are developed and analysed in workshops with stakeholders, who often do not have technical simulation knowledge. From a practical perspective, our study shows that similar benefits can be gained from using simple models compared to more adequate ones, which calls for the simulation community to consider developing the simplest possible models. Indeed, Green and Armstrong (2015) found that complexity does not necessarily improve models' forecasting accuracy. Similarly, Kunc et al. (2019) posit that simple models can clarify stakeholder conflict and support the need for further work to understand how users behaviour is impacted by interacting with and using models. At the same time, our study opens avenues for further academic research, to answer questions such as: "what level of model detail is appropriate to build a sufficiently simple model for the situation at hand?", "what are the requirements that specify adequate details for producing simple models?", etc. We hope that more researchers will find an interest in this topic and carry out more research to provide answers to such interesting questions.

5.2 Limitations

There are certain limitations related to the study that may have affected our findings, which need to be considered.

One limitation of the study relates to the choice of participants. We use a convenience sample, consisting of undergraduate students that the authors had access to. We acknowledge that students' decision making and performance in the task may differ compared to managers. The reason for choosing students was to ensure that the groups had all similar prior knowledge about the resource allocation problem for the ambulance service. This is furthermore supported by the experimental study of Bakken, Gould and Kim (1992), who found that students performed better than managers on their simulation based experimental task because the students did not possess prior knowledge and experience and as a result did not freely experiment with the model. Obtaining a sample of managers would have been very difficult and most importantly, it would have been difficult to identify a sample with a homogenous level of experience about the problem. Furthermore, as the participants were final year students, it should be noted that they all had one year of work experience, albeit not relevant to health care setting, hence had some understanding of management type decisions.

Another concern related to the participants is the relatively small sample size in each group (NM=10, SM=16 and AM =13). Therefore, some caution should be placed in the interpretation of the inferential statistics used to compare the participants opinions about the models in the Likert-type questions. Due to the relatively small sample size we only tested one model at two complexity levels. If we had access to a larger group, we would have been able to test different levels of model complexity.

Not all participants provided elaborate answers to the open-ended questions, leading to a small number of qualitative answers to support the analysis. We also observed a variation in the level of participation between the different groups, with some groups leaving the preparation of their presentation to the very end. This partly explains the lack of unanimous responses with regards to the recommended solution. A future addition to the experiment could be to include managers as participants, to understand how the interaction between modellers and decision makers into organisations can lead to changes in their behaviour (Kunc et al. 2018).

Furthermore, group composition, dynamics, cohesion and interdependence may have affected the outcomes of group results and the quality of participants' answers (Forsyth, 2018). Bearing in mind these limitations, we next plan to run another set of experiments with different participants on an individual basis in order to limit the impact of group-related factors. A further extension to this study would be to identify the minimum requirements in terms of simplifications introduced to a model to be considered useful for learning purposes. Further work is underway to understand more widely the validity and credibility of models in practice, by administering interviews with simulation experts to identify how and whether simplified and more widely wrong models are used in practice.

6. CONCLUSIONS

This paper explores the learning achieved from simulation models focusing on the level of model detail. A laboratory experiment was set up to identify whether the level of model detail can affect the learning achieved by users. We compare two types of models: an adequate and a simplified one, where the latter was developed by introducing simplifications and approximation to the adequate model. The adequate model was an adapted version of a model developed and validated in real life facilitated workshops with stakeholders from an ambulance service. For this reason, the adequate model is considered requisite. The results suggest that differences in learning from using the two models were not significant. We however found that simple model users had a better understanding of the problem, albeit they were able to comment that they needed a more detailed model to be able to solve the task set. These results are encouraging, providing some initial empirical evidence that simple models, so long as they are not inaccurately presented, can be useful in supporting clients to understand their problems and take decisions. This work can be particularly useful to inform the current facilitated modelling practice with regards to the level of detail (simplifications and complexity) included in the models used in facilitated modelling workshops with clients.

REFERENCES

Akpan, I. J., & Brooks, R. J. (2014). Experimental evaluation of user performance on two-dimensional and three-dimensional perspective displays in discrete-event simulation. *Decision Support Systems*, *64*, 14–30. http://doi.org/10.1016/j.dss.2014.04.002.

Argyris, C., and Schön, D. A. (1996). *Organizational Learning II - Theory, Method, and Practice*. Addison-Wesley.

Bahtiyar, M. (2005). *Simulation modeling of a shop floor activities for SMEs in virutal enterprises*. Middle East Technical Unirvesity. Retrieved from http://etd.lib.metu.edu.tr/upload/12606184/index.pdf

Baker, A. (2006). Occam's Razor in science: A case study from biogeography. *Biology and Philosophy*, *22*(2), 193–215. http://doi.org/10.1007/s10539-006-9027-9.

Bakken, B., Gould, J., & Kim, D. (1992). Experimentation in learning organizations: A management flight simulator approach. *European Journal of Operational Research*, *59*(1), 167–182. http://doi.org/10.1016/0377-2217(92)90013-Y

Balci, O. (1994). Validation, verification, and testing techniques throughout the life cycle of a simulation study. *Annals of Operations Research*, *53*, 121–173. http://doi.org/10.1109/WSC.1994.717129

Bankes, S. (1998). Policy Analysis for Complex and Uncertain Systems Through Computational Experiments. In *IEEE Aerospace Conference* (pp. 9–21 vol.1).

Bi, J., and Kuesten, C. (2015). Revisiting Fisher's "Lady Tasting Tea" from a perspective of sensory discrimination testing. *Food Quality and Preference*, *43*, 47–52. http://doi.org/10.1016/j.foodqual.2015.02.009

Blumer, A., Ehrenfeucht, A., Haussler, D., and Warmuth, M. K. (1987). Occam's razor. *Information Processing Letters*, *24*, 377–380.

Brooks, R. J., and Tobias, A. M. (1996). Choosing the best model: Level of detail, complexity, and model performance. *Mathematical and Computer Modelling*, *24*(4), 1–14. http://doi.org/10.1016/0895-7177(96)00103-3

Bruno, W. J., and Halpern, a L. (1999). Topological bias and inconsistency of maximum likelihood using wrong models. *Molecular Biology and Evolution*, *16*(4), 564–566. http://doi.org/10331281

Bryman, A. and Cramer, D. (2011). *Quantitative Data Analysis with IBM SPSS 17, 18 and 19*. 1st edition. Hoboken: Routledge.

Cárdenas Duarte, M. A., Rojas Cepeda, J. A., González-Neira, E. M., Barrera, D., Cortés, V. R., and Rey, G. Z. (2017). Performance evaluation of a GRASP-based approach for stochastic scheduling problems. *Uncertain Supply Chain Management*, *5*, 359–368. http://doi.org/10.5267/j.uscm.2017.4.002

Carvalho, A. N., Scavarda, L. F., and Lustosa, L. J. (2014). Implementing finite capacity production scheduling: Lessons from a practical case. *International Journal of Production Research*, *52*(4), 1215–1230. http://doi.org/10.1080/00207543.2013.848484

Chwif, L., Paul, R. J., & Barretto, M. R. P. (2006). Discrete event simulation model reduction: A causal approach. *Simulation Modelling Practice and Theory*, *14*(7), 930–944. http://doi.org/10.1016/j.simpat.2006.05.001.

Dickert, B. F., & Zanakis, S. H. (1979). Simulation of an industrial computer system-A case study. *Computers and Industrial Engineering*, *3*(3), 203–215. http://doi.org/10.1016/0360-8352(79)90014-7.

Elsawah, S., McLucas, A., and Mazanov, J. (2017). An empirical investigation into the learning effects of management flight simulators: A mental models approach. *European Journal of Operational Research*. http://doi.org/10.1016/j.ejor.2016.10.011

Forsyth, D. (2018). *Group dynamics*. 7th ed. Boston: Cengage.

Franco, L. A., and Montibeller, G. (2010). Facilitated modelling in operational research. *European Journal of Operational Research, 205*(3), 489–500. http://doi.org/10.1016/j.ejor.2009.09.030

Gass, S. I. (1993). Model accreditation: A rationale and process for determining a numerical rating. *European Journal of Operational Research*, *66*(2), 250–258. http://doi.org/10.1016/0377-2217(93)90316-F

Gell-Mann, M. (1995). What is complexity. In *John Wiley and Sons, Inc.: Complexity* (Vol. 1, pp. 16–19). http://doi.org/10.1002/bies.10192

Gell-Mann, M., and Lloyd, S. (2003). *Effective Complexity*. *Nonextensive Entropy - Interdisciplinary Applications*.

Gogi, A., Tako, A. A., and Robinson, S. (2016). An experimental investigation into the role of simulation models in generating insights. *European Journal of Operational Research, 249*, 931–944. http://doi.org/10.1016/j.ejor.2015.09.042

Goldberg, J., Dietrich, R., Chen, J. M., Mitwasi, M., Valenzuela, T., and Criss, E. (1990). A simulation model for evaluating a set of emergency vehicle base locations: Development, validation, and usage. *Socio-Economic Planning Sciences*, *24*(2), 125–141. http://doi.org/10.1016/0038-0121(90)90017-2

Green, K. C., and Armstrong, J. S. (2015). Simple versus complex forecasting: The evidence. *Journal of Business Research*, *68*(8), 1678–1685. http://doi.org/10.1016/j.jbusres.2015.03.026

Hahn, H. A. (2013). The conundrum of verification and validation of social science-based models. *Procedia Computer Science*, *16*, 878–887. http://doi.org/10.1016/j.procs.2013.01.092

Hämäläinen, R. P., Luoma, J., and Saarinen, E. (2013). On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems. *European Journal of Operational Research*, *228*(3), 623–634. http://doi.org/10.1016/j.ejor.2013.02.001

Hodges, J. S. (1991). Six (Or So) Things You Can Do with a Bad Model. *Operations Research, 39*(3), 355–365. http://doi.org/10.1287/opre.39.3.355

Homer, J. B., and Hirsch, G. B. (2006). System dynamics modeling for public health: Background and opportunities. *American Journal of Public Health*, *96*(3), 452–458. http://doi.org/10.2105/AJPH.2005.062059.

Innis, G., and Rexstad, E. (1983). Simulation model simplification techniques. *Simulation*, *41*, 7–15. http://doi.org/10.1177/003754978304100101

Kaplan, R. S. (1986). The role for empirical research in management accounting. *Accounting, Organizations and Society*, *11*(4–5), 429–452. http://doi.org/10.1016/0361-3682(86)90012-7

Katsikopoulos, K. V., Durbach, I. N., and Stewart, T. J. (2017). When should we use simple decision models? A synthesis of various research strands. *Omega*, *0*, 1–9. http://doi.org/10.1016/j.omega.2017.09.005.

Kotiadis K and Tako A.A. (2010) PartiSim User Guide to Facilitation:Version 1. (https://www.researchgate.net/publication/284030194_final_PARTISIM). DOI: 10.13140/RG.2.1.3659.1201.

Larroca, F., and Rodríguez, F. (2014). An Overview of WLAN Performance, Some Important Case-Scenarios and Their Associated Models. *Wireless Personal Communications*, *79*, 131–184. http://doi.org/10.1007/s11277-014-1846-4

Lee, Y. S., Yi, Y. K., and Malkawi, A. (2011). Simulating human behaviour and its impact on energy uses. *Proceedings of Building Simulation 2011*, 1049–1056.
Lödding, H., Yu, K. W., and Wiendahl, H. P. (2003). Decentralized WIP-oriented manufacturing control (DEWIP). *Production Planning and Control*, *14*(1), 42–54. http://doi.org/10.1080/0953728021000078701
Luoma, J. (2014). Model-based organizational decision making: A behavioral lens. *European Journal of Operational Research*, *249*(3), 816–826. http://doi.org/10.1016/j.ejor.2015.08.039
Monks, T., Robinson, S., and Kotiadis, K. (2014). Learning from discrete-event simulation: Exploring the high involvement hypothesis. *European Journal of Operational Research*, *235*(1), 195–205. http://doi.org/10.1016/j.ejor.2013.10.003
Monks, T., Robinson, S., and Kotiadis, K. (2016). Can involving clients in simulation studies help them solve their future problems? A transfer of learning experiment. *European Journal of Operational Research*, *249*(3), 919–930. http://doi.org/10.1016/j.ejor.2015.08.037.
Newell, B., and Siri, J. (2016). A role for low-order system dynamics models in urban health policy making. *Environment International*, *95*, 93–97. http://doi.org/10.1016/j.envint.2016.08.003.
Nwodo, T. C., and Okoro, O. I. (2015). Simulation Tools for Electrical Machines Modelling : Teaching and Research. *Nigerian Journal of Technology*, *24*(1), 33–38.
Pace, D. K. (2004). Modeling and simulation verification and validation challenges. *Johns Hopkins APL Technical Digest (Applied Physics Laboratory)*, *25*(2), 163–172. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-16244422023andpartnerID=40andmd5=f6e7e09b236f72a10d227c1ff03e0987
Posada, D. (2003). The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny. In *The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny*. http://doi.org/10.1186/1471-2148-7-S1-S5
Puntambekar, N. (2016). *Exploring the Efficiency of a UK Ambulance Service through Discrete-Event Simulation* (master's thesis). Loughborough University, UK.
Kunc, M., & Morecroft, J. (2007). System dynamics modelling for strategic development. Supporting Strategy: Frameworks, methods and models, 157-190.
Kunc, M., Harper, P., & Katsikopoulos, K. (2018). A review of implementation of behavioural aspects in the application of OR in healthcare. Journal of the Operational Research Society, 1-18.
Robinson, S. (2014). *Simulation - The Practice of Model Development and Use*, 2nd edition. Macmillan education, Palgrave.
Robinson, S. (2015). A tutorial on conceptual modeling for simulation. In *Proceedings of the 2015 Winter Simulation Conference* (pp. 1820–1834).
Saleh, M., Oliva, R., Kampmann, C. E., and Davidsen, P. I. (2010). A comprehensive analytical approach for policy analysis of system dynamics models. *European Journal of Operational Research*, *203*(3), 673–683. http://doi.org/10.1016/j.ejor.2009.09.016.
Schacter, D., Gilbert, D., Wegner, D. and Hood, B. (2014). *Psychology*. 3rd ed. Worth Publishers.
Schulze, J., Müller, B., Groeneveld, J., and Grimm, V. (2017). Agent-Based Modelling of Social-Ecological Systems: Achievements, Challenges, and a Way Forward. *Journal of Artificial Societies and Social Simulation*, *20*(2). http://doi.org/10.18564/jasss.3423
Shalizi, C. R. (2003). Methods and Techniques of Complex Systems Science: An Overview, 1–96. http://doi.org/10.1007/978-0-387-33532-2_2
Silver, E. A. (2004). An overview of heuristic solution methods. *Journal of the Operational Research Society*, *55*(9), 936–956. http://doi.org/10.1057/palgrave.jors.2601758.

Soo, V. K., Compston, P., and Doolan, M. (2019). Life Cycle Modelling of End-of-Life Products: Challenges and Opportunities towards a Circular Economy. *Procedia CIRP*, *80*, 607–612. http://doi.org/10.1016/j.procir.2019.01.092.

Tako, A. A., and Robinson, S. (2009). Comparing discrete-event simulation and system dynamics: users' perceptions. *Journal of the Operational Research Society*, *60*(3), 296–312. http://doi.org/10.1057/palgrave.jors.2602566

Tako, A. A., and Kotiadis, K. (2015). PartiSim: A multi-methodology framework to support facilitated simulation modelling in healthcare. *European Journal of Operational Research*, *244*(2), 555–564. http://doi.org/10.1016/j.ejor.2015.01.046

Thompson, J. P., Howick, S., and Belton, V. (2016). Critical Learning Incidents in system dynamics modelling engagements. *European Journal of Operational Research*, *249*(3), 945–958. http://doi.org/10.1016/j.ejor.2015.09.048

van der Zee, D-J. (2019). Model Simplification in Manufacturing Simulation - Review and Framework, *Computers and Industrial Engineering*, *127, 1056-1067* doi: https://doi.org/10.1016/j.cie.2018.11.038

Van der Zee, D. J. (2017) "Approaches for Simulation Model Simplification". In *Proceedings of the 2017 Winter Simulation Conference*, edited by K. V. Chan et al., 4197-4208. Piscataway, New Jersey: IEEE.

Van Der Zee, D. J., Holkenborg, B., & Robinson, S. (2012). Conceptual modeling for simulation-based serious gaming. *Decision Support Systems*, *54*(1), 33–45. http://doi.org/10.1016/j.dss.2012.03.006

Van der Zee, D.J., Tako, A., Robinson. S., Fishwick, P. and O. Rose (2018) Panel: education on simulation model simplification - beyond rules of thumb. In Proceedings of the 2018 Winter Simulation Conference. Edited by Rabe M et al., 3987-4001. Piscataway, NJ, USA: IEEE.

Vanpaemel, W. (2009). Measuring model complexity with the prior predictive. *Advanced in Neural Information Processing Systems*, *22*, 1919–1927. Retrieved from https://papers.nips.cc/paper/3886-measuring-model-complexity-with-the-prior-predictive.pdf

Wahlström, B. (1994). Models, modelling and modellers: an application to risk analysis. *European Journal of Operational Research*, *75*(3), 477–487. http://doi.org/10.1016/0377-2217(94)90290-9

Ward, K. A. (2014). *A prospected Study For Development Of Berths Facility Services (Loading and Unloading) Of Gaza Seaport Using Simulation Techniques*.