*Title:*

Are "wrong" models useful? A qualitative study of Discrete Event Simulation modeller stories

*Authors:*

*Dr. Naoum Tsioptsias*, Loughborough University, Epinal Way, Loughborough LE11 3TU, n.tsioptsias@lboro.ac.uk

*Dr. Antuela Tako*, Loughborough University, Epinal Way, Loughborough LE11 3TU, a.takou@lboro.ac.uk

*Professor Stewart Robinson*, Loughborough University, Epinal Way, Loughborough LE11 3TU, s.l.robinson@lboro.co.uk

*Supplementary material*

*Appendix A: Information on research method selection*

The following information provides additional explanations as to the selection and implementation of the research method followed in this paper.

 Qualitative research methods that were considered included: ethnography, case studies and interviews. Although the first two may offer further in-depth understanding for their contents, they both are slow techniques that also require luck in the sense of encountering "wrong" models. Instead, factors of wrongness and uses of "wrong" models can be found much faster in a story and for a much larger number of cases when using interviews. As long as the interviewees can recall the events and the interactions and dialogs with their clients, possible reasons that lead to a "wrong" model and possible uses of such a model can be derived. Additionally, studies exist that use interviews within OR and Simulation that can be considered for properly forming our processes. These include works from Tako (2015) where expert modellers are interviewed on how they develop DES models, and from Gogi (2016) where DES modellers are interviewed on their experience of insight generations from models. Our research here follows the latter, where semi-structured interviews are put together. This means that every interview is expected to have unique elements and dialogues, with a general set of questions used as a referral for gathering the material but the discussions are to remain flexible to adapt these questions based on what the interviewees mention.

 The sample does not have to be probabilistic as long as it serves the research objectives' purpose (Guest et al., 2006). Instead, it needs to be adequate to offer data saturation for the topic (Bryman, 2012). This means that a sample may be viewed as adequate when theory has started saturating which can indicatively happen at around 12 interviews (Guest et al., 2006). In our case a specific threshold cannot be allocated due to lack of previous investigations on the topic.

 In view of the data analysis, the coding was based on literature sources with guidelines for social sciences (Brinkmann and Kvale, 2015; Robson and MacCartan, 2016) but also derives by examining previous OR works that implement social studies' principles (e.g. Tako, 2015) and coding approaches (e.g. Gogi, 2016). A list of initial codes from the literature on the various topics discussed (on wrongness and uses) was prepared and populated with additions based on what was found from the analysis. Similar codes were collapsed together. Once the final codes were created, comparison with literature took place. The rest of the paper describes the findings and their analysis.

 Two last comments should be made here. First of all, all data derives from discussions with modellers. It thus depends on their memory and understanding of how the clients perceived a model and we incorporate all options under this premise. Secondly, as suggested in the main text, the term "wrong" is subjective and we use it for any model that is not considered good enough by clients and/or modellers for at least one of its components. We do not equate use of a "wrong" model to a modeller willingly developing or adapting it.

*Appendix B: Information on the interviewees of the study*

Appendix Table 1. Demographics of interviewees

| **Participant (M)** | **Company** | **Experience** | **Stories shared** |
| --- | --- | --- | --- |
| M01 | A | J/I | 5 |
| M02 | A | J/I | 3 |
| M03 | A | J/I | 3 |
| M04 | A | J/I | 2 |
| M05 | A | S | 5 |
| M06 | A | J/I | 3 |
| M07 | B | J/I | 2 |
| M08 | B | S | 0 |
| M09 | B | S | 5 |
| M10 | B | J/I | 2 |
| M11 | B | J/I | 1 |
| M12 | B | J/I | 3 |
| M13 | B | J/I | 3 |
| M14 | C | S | 2 |
| M15 | C | S | 2 |
| M16 | C | S | 2 |
| M17 | C | S | 1 |
| M18 | C | J/I | 2 |
| M19 | C | S | 2 |
| M20 | D | S | 2 |
| M21 | E | S | 2 |
| M22 | F | S | 2 |
| *Experience is deduced by the researcher based on participants' background. Senior (S) is considered a participant with at least 10 years of experience, regardless of position as opposed to Junior (J) or Intermediate (I).* |

Appendix Table 2. Categorisations of interview stories

| **Category** | **Number of stories** |
| --- | --- |
| **Sector of client** | Healthcare | 9 (16.6%) |
| Pharmaceutical | 1 (1.8%) |
| Services | 7 (13%) |
| Manufacturing/Production | 30 (55.6%) |
| Retail | 3 (5.6%) |
| Public sector | 4 (7.4%) |
| **Problem type** | Resource utilisation | 32 (59.3%) |
| Additions | 17 (31.5%) |
| Support sales or bidding | 4 (7.4%) |
| Digitilise material | 1 (1.8%) |
| *Stories are matched to only 1 Sector and Problem type. Percentages are out of 54 stories.* |

*Appendix C: Additional results*

The additional analysis here regards reasons for using a "wrong" model as well as how it was implemented.

 Regarding the former, and despite other possible categorisations that could be considered, we have selected to split those 25 models under 4 different sets of reasons for using "wrong" models. The next table summarises these categories:

Appendix Table 3. Results on interviews regarding reasons for using models

|  |  |
| --- | --- |
| **Reasons** | **Number of models** |
| Further investigation | 8 (32%) |
| No alternative | 7 (28%) |
| Client happy | 7 (28%) |
| Third party | 3 (12%) |
| **Total** | **25** |
| *Each model is attributed to one main category of reason for being used* |

Three of the four categories have a similar number of stories attributed to them. The first one refers to *further investigation* taking place. This means that because of exploring options, a "wrong" model was decided to be used regardless of wrongness. This investigation may have shown evidence that a "wrong" model is useful because of specific details that came to light related to the project, either proving to the clients that the model is useful or revealing why the model is not performing as expected through further experimentation/exploration. As such, the model, regardless of denoted wrongness, was found possible to use. The second reason regards lack of alternatives, where a "wrong" model may be used due to having no other options except to implement this model or because changes made have no effect on its wrongness. The next reason addresses the fact that clients may actually be happy with their models considering them credible for the purpose at hand regardless of wrongness attributed. The last reason is the involvement of third parties which includes personnel (e.g. higher management or external collaborators from the clients’ side) not involved in the development of a model insisting on use of that model.

 Next, we briefly refer to how a "wrong" model was implemented. The distinction here does not address uses but rather some understanding on the way that a "wrong" model – regardless of its actual use – may be put to use: either exactly as it was provided by the modeller or with small changes taking place. The next table summarises the stories per way of implementation:

Appendix Table 4. Results of interviews regarding implementation of uses

|  |  |
| --- | --- |
| **Implementation** | **Total** |
| Used as is | 14 (56%) |
| Used after changes | 11 (44%) |
| *Percentages out of 25 stories* |

The first category refers to models implemented as presented to the clients because they are considered useful by the clients from start and they can be used without actual changes. The second category refers to models implemented after minor amendments taking place that do not though change the model into a "good" one. We notice that there is an almost equal split between the way that "wrong" models that are used are implemented, with more of them being used as presented to the clients.

 To summarise and comment upon the additional results, we found four main reasons that "wrong" models are used for – three of which were mentioned at a similar frequency: further investigation taking place that changes the clients' opinion for a model, no alternatives to using a "wrong" model are available, the client is happy with using a model regardless of its adequacy or wrongness, and third parties decide on use of a "wrong" model. This suggests that clients may actually choose to use a "wrong" model if they are shown evidence of model usefulness or if they are not concerned about its wrongness. It is also noted that third parties may affect that decision. More than half of the "wrong" models that were used, were implemented in the state presented to the clients (14/25). This happened either because clients found a model useful or because it was proven useful. The rest were implemented after minor changes although those models were still not considered as adequate. This shows that even if changes apply to a model, it may not necessarily become good enough which suggests that some models are not "rectifiable". This aligns with Bankes' (1998) suggestions that we need to learn how to use "wrong" models. Indeed we noticed that for some models, changes did not need to take place in order to take advantage of them.