Experiment with competency questions and explanation transformations

Elena Romanenko

December 2024

1 Questionnaire description

As ExpO is currently in the prototyping stage, and to avoid issues with the new interface while testing the explanation approach, we initiated a questionnaire as a preliminary step of the evaluation. The primary focus of the research was to ascertain whether the proposed transformation operations indeed contribute to an improved understanding of the given model.

The questionnaire consisted of three parts. The first one asked some questions about the interviewee's experience in conceptual modeling. The second one presented one or several models and asked questions about them. In the final part, we asked about satisfaction from the given models when considering them as explanations. These questions were adapted from the *System Causability Scale*.

The purpose of the System Causability Scale is to quickly determine whether an explanation process or an explanation itself is suitable for the intended purpose [4]. Built upon the widely adopted *System Usability Scale*, it consists of 10 questions. However, given that the explanation transformations were already completed and the models were fixed, some questions were not applicable in our context. Hence, we asked the following questions:

1. I found that the model(s) included all relevant information with sufficient precision and granularity.
2. I understood the model(s) within the context of the question.
3. I did not need support to understand the model(s).
4. I was able to use the model(s) with my basic knowledge.
5. I think that most people would learn to understand the model(s) very quickly.

As for the main part of the questionnaire, we selected two models from the OntoUML/UFO catalog, that were published together with the corresponding *competency questions* (CQs): 'aguiar2019ooc' [1] and 'jacobs2022sdpontology' [5].

According to the SABiO methodology [2], one of the first activities in ontology development is identifying CQs, i.e., requirements in the form of questions that the ontology must be able to answer [3]. Modelers who develop ODCMs, usually adopt the same methodology, but those questions are rarely published.

Table 1: Competency questions asked in the questionnaire.

Rule		Competency question	Request (concept)
aguiar2019ooc	CQ1	What makes up an Object-Oriented source code?	Focus (Object-Oriented Source Code) with hop=2
	CQ2	What is the visibility of an element present in an OO source code?	Focus (Element Visibility) with hop=2
	CQ3	How are classes logically organized in an Object-Oriented source code?	Focus (Class) with hop=2; Abstract aspects; Delete (Primitive type, Method, Variable, Attribute, Member, Type)
	CQ4	What elements compose a class?	Focus (Class) with hop=2; Abstract aspects; Delete (Physical module, Logical module, Type, Return type)
	CQ8	What is the mutability of a variable?	Focus (Mutability) with hop=2
jacobs2022sdpontology	CQ3	What are the proficiency levels of an employee?	Focus (Employee) with hop=2
	CQ4	Which role(s) does a member of the team have?	Cluster (Membership)
	CQ8	Which development method is being used?	Focus (Software development project) with hop=1
	CQ9	For which type(s) of platform(s) is the software being built?	Focus (Platform) with hop=1
	CQ12	What is the level of the project requirements?	Focus (Non-functional requirements) with hop=1

Table 1 overviews the selected CQs and transformations that were applied to produce the reduced versions of ODCMs, where the numbering of the five CQs that we selected for each model is left as in the respective literature source. For the 'aguiar2019ooc' model we selected questions that required familiarity only with one module (out of three); for the 'jacobs2022sdpontology' we combined all sub-modules.

2 Results of the questionnaire analysis

In total, we received 25 answers. However, the responses from one interviewee were disqualified as they were incorrect, and we had the impression that unfortunately, the individual did not understand the task. In general, most of the interviewees were high-qualified professionals with more than five years of experience in modeling (see also Figure 1).

Depending on the model, the interviewees have spent different amounts of time on the questionnaire. In the case of the 'aguiar2019ooc' model, the average time was about 20 minutes. Instead, for 'jacobs2022sdpontology' it was 12 minutes, and almost the same time, namely 13 minutes, for the explained models.

We were glad to see that there were no questions that everyone answered incorrectly, moreover, the percentage of the correct answers was not lower than

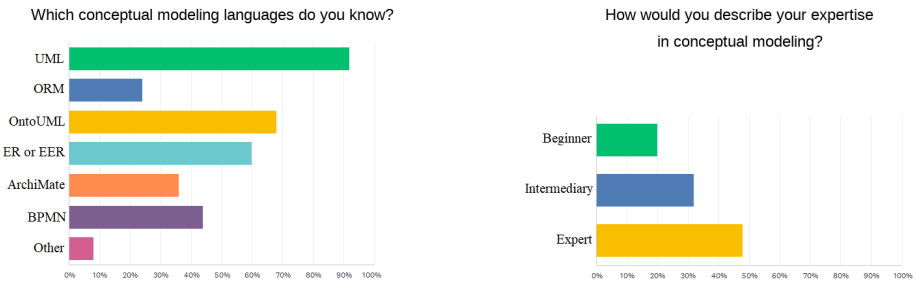

Figure 1: Characteristics of the interviewees.

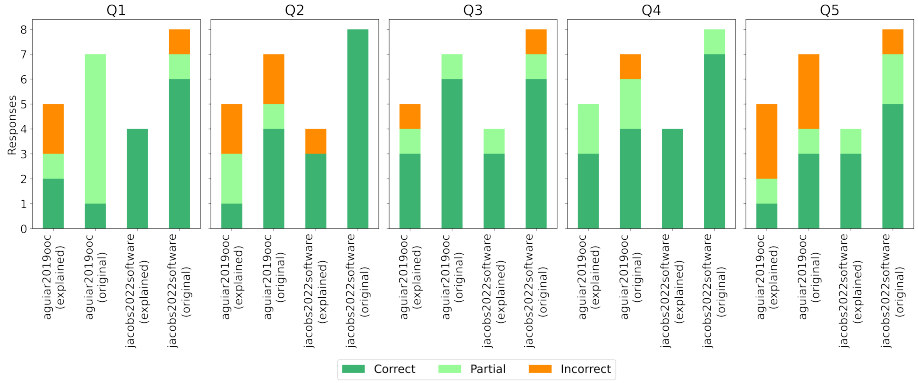

Figure 2: Results of the answers to the competency questions. Q1–Q5 refers to the 5 questions for each of the two models, in the order as listed in Table 1.

50%. For 'jacobs2022sdpontology' we received more accurate answers. For this model, on average, the percentage of incorrect answers for every question was less than 10%. Also, for all questions, there were correct answers on both — explained and original — models (see Figure 2). Although the order of the questions was randomized for every interviewee, here Q1–Q5 correspond to the questions in the order presented in Table 1, thus, e.g., Q1 corresponds to CQ1 or CQ3 depending on the model. For the 'jacobs2022sdpontology' model the answers, in general, were better with the explained models. For 'aguiar2019ooc', the results are more ambiguous.

As for the last part of the questionnaire, most interviewees did not complain about the quality of the explained models (see Figure 3). Nevertheless, it seems to be important not to shrink the original models too much, since people were expecting more complete models.

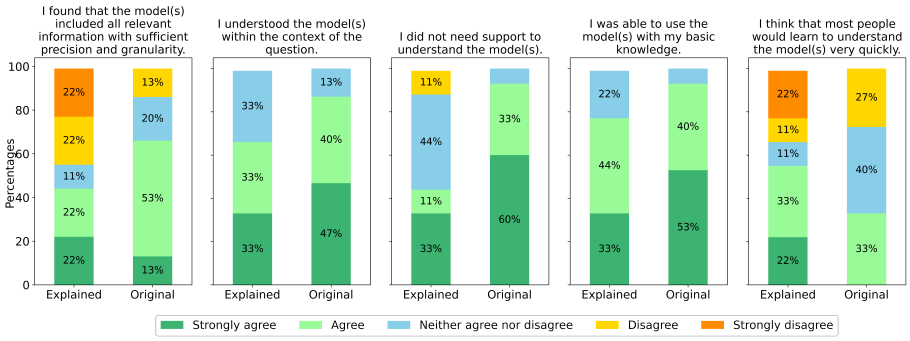

Figure 3: Results of the answers to the competency questions.

References

[1] de Aguiar, C.Z., de Almeida Falbo, R., Souza, V.E.S.: OOC-O: A reference ontology on object-oriented code. In: Laender, A.H.F., Pernici, B., Lim, E., de Oliveira, J.P.M. (eds.) Conceptual Modeling - 38th International Conference, ER 2019, Salvador, Brazil, November 4-7, 2019, Proceedings. Lecture Notes in Computer Science, vol. 11788, pp. 13–27. Springer (2019). https://doi.org/10.1007/978-3-030-33223-5_3

[2] de Almeida Falbo, R.: SABiO: Systematic approach for building ontologies **1301** (09 2014), `https://ceur-ws.org/Vol-1301/ontocomodise2014_-2.pdf`

[3] Gruninger, M., Fox, M.S.: Methodology for the design and evaluation of ontologies. In: International Joint Conference on Artificial Intelligence (1995)

[4] Holzinger, A., Carrington, A., Müller, H.: Measuring the quality of explanations: The system causability scale (scs): Comparing human and machine explanations. KI - Künstliche Intelligenz **34**(2), 193–198 (Jan 2020). https://doi.org/10.1007/s13218-020-00636-z, `http://dx.doi.org/10.1007/s13218-020-00636-z`

[5] Jacobs, M.: A software development project ontology. Master's thesis, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente (August 2022), `http://essay.utwente.nl/93228/`