Requirements Quality Impact: Causal Model

This document contains a causal model of requirements quality impact, which we investigate in the scope of an experimental study.

Motivation

Requirements Quality research has proposed the fundamental hypothesis that the quality of requirements - i.e., the way a requirement is written - impacts its subsequent use [1,2]. This spawned the research stream of requirements quality factors [3], i.e., the proposal of metrics that evaluate natural language requirements and quantify their quality [4]. For example, using passive voice in a requirement is associated with bad quality [5].
However, the empirical evidence for this association is thin: previous research fails to provide actual, empirical evidence that would prove the impact of quality factors [2,3]: most claims are based on anecdotal evidence (e.g., authors assumed that passive voice harms the subsequent processing of requirements) which undermines their reliability [3].
Requirements quality research needs a method for assessing the impact of quality factors empirically. We propose using an experimental study to collect data for a Bayesian workflow for causality analysis. In this project, we demonstrate this method and show how it can be used to empirically evaluate the actual causal impact of requirements quality factors on subsequent activities.

Causal Model

Causal Assumptions

Our causal understanding of requirements quality includes two assumptions:

1. Activities performed on requirements depend on requirements quality
2. Activities performed on requirements depend on the context

The first assumption represents the hypothesis formulated in the introduction: the quality of written requirements impacts how they are used in later activities. The second assumption represents an alternative, postulating that it is not the quality of the requirements but rather the context (i.e., the experience of the involved engineers, the available domain knowledge, etc.) that determines how requirements are used. The two hypotheses are not mutually exclusive but represent different stances: one hinges requirements quality on the written requirement itself, and the other hinges on the people using the requirement.

Variables

We capture the two concepts in the following variables:

H	Variable	Type	Measurement
1	**Requirements Quality Defect**	Categorical	[none, passive voice, ambiguous pronouns, passive voice + ambiguous pronouns]
2	**Experience in SE**	Count	Years of working in SE
2	**Experience in RE**	Count	Years of working in RE
2	**Experience in roles**	Count matrix	List of roles, each annotated with the number of years spent in that role
2	**Level of education**	Orginal	High-school, Bachelors, Masters, Ph.D.
2	**Experience with modeling**	Ordinal	[never, rarely, occasionally, often]
2	**Formal modeling training**	boolean	true/false
2	**Domain experience**	Ordinal matrix	List of relevant domains, each annotated with an ordinal scale (1-5) of familiarity

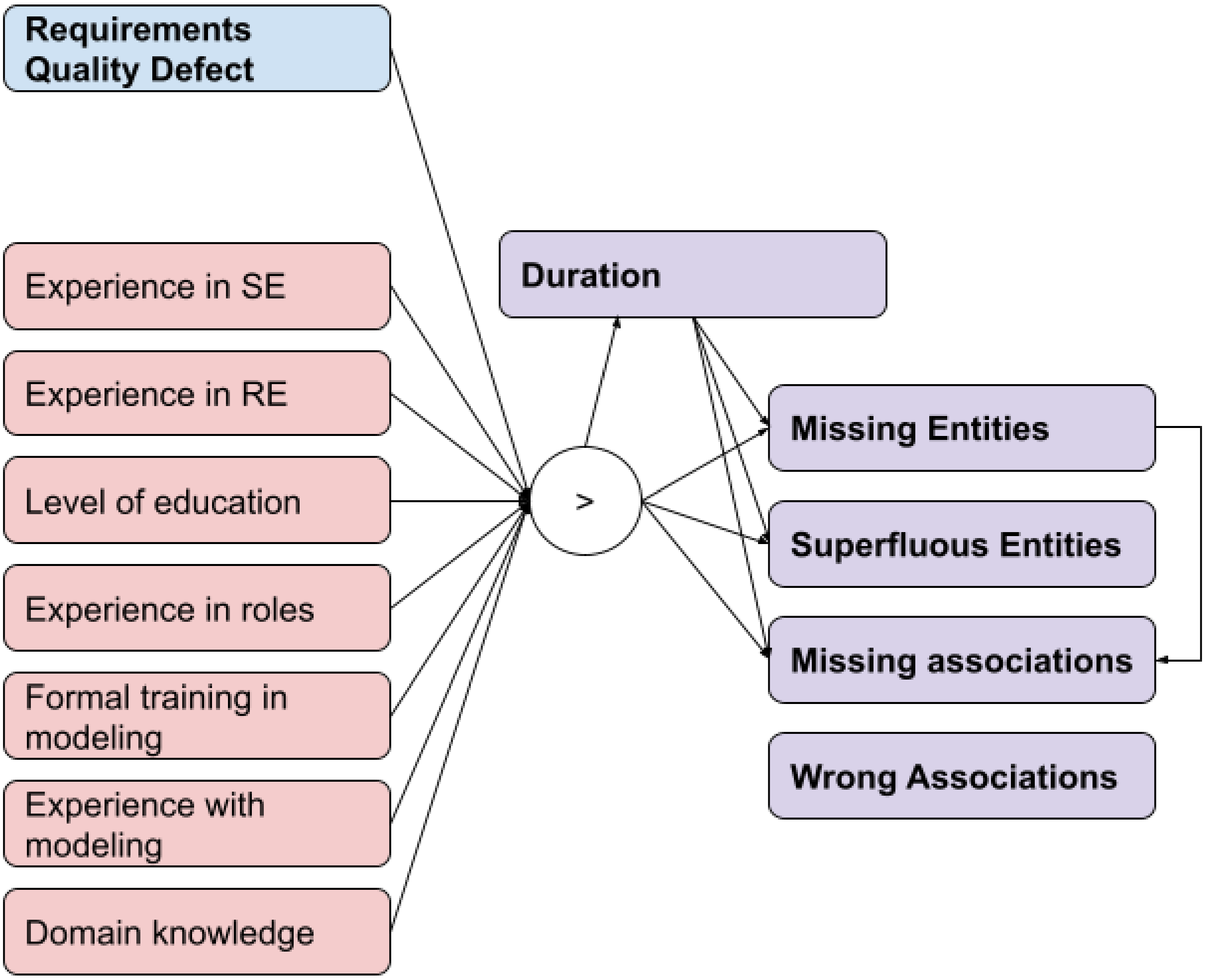
Passive voice and **ambiguous pronoun** are two popular requirements quality factors [3] we chose to represent the latter. Other requirements quality factors could have been chosen here, and a holistic model would need to include all of them. We selected these two for the demonstration purpose of this study as they are well-researched and easy to demonstrate. In the experimental study, these two variables will serve as the factors that we manipulate.

Experience in software engineering, **experience in requirements engineering**, **experience in roles**, and **level of education** approximates the skill of an engineer. This represents one part of the context on which hypothesis 2 hinges.

Experience with modeling and **formal modeling training** are task-specific context factors, as they also benefit processing a requirement and potentially mitigate the effect of the quality factors. Similarly, **domain experience** is likely to compensate for missing information in the requirement, which represents the impact of the context on the requirements processing according to hypothesis 2.

Directed Acyclic Graph

The following directed acyclic graph (DAG) visualizes our causal assumptions in the form of variables and their causal relationships.



We are employing a domain modeling task in our experiment, i.e., participants are presented with a requirement (where each of the four requirements has a different configuration of the **passive voice** and **ambiguous pronoun** factors) and have to translate it to a domain model. As response variables, we investigate the impact of **missing entities**, **missing associations**, **missing associations,** and **wrong associations** (i.e., associations where either the source or target of the edge is wrong). These represent the "performance" of a requirements-dependent task.

The goal of our study will be to see whether the different treatments (use of passive voice and/or ambiguous pronouns) impact the response variables given a representative sample of the other independent variables (context factors representing hypothesis 2). This would provide empirical evidence for or against hypotheses 1 and 2, as it allows us to determine which of these variables have the strongest influence on the response variables and, therefore, the processing of requirements.

References

[1] Femmer, H., & Vogelsang, A. (2018). Requirements quality is quality in use. *IEEE Software*, *36*(3), 83-91.

[2] Montgomery, L., Fucci, D., Bouraffa, A., Scholz, L., & Maalej, W. (2022). Empirical research on requirements quality: a systematic mapping study. *Requirements Engineering*, *27*(2), 183-209.

[3] Frattini, J., Montgomery, L., Fischbach, J., Unterkalmsteiner, M., Mendez, D., & Fucci, D. (2022, August). A Live Extensible Ontology of Quality Factors for Textual Requirements. In *2022 IEEE 30th International Requirements Engineering Conference (RE)* (pp. 274-280). IEEE.

[4] Femmer, H., Fernández, D. M., Wagner, S., & Eder, S. (2017). Rapid quality assurance with requirements smells. *Journal of Systems and Software*, *123*, 190-213.

[5] Femmer, H., Kučera, J., & Vetrò, A. (2014, September). On the impact of passive voice requirements on domain modelling. In *Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement* (pp. 1-4).