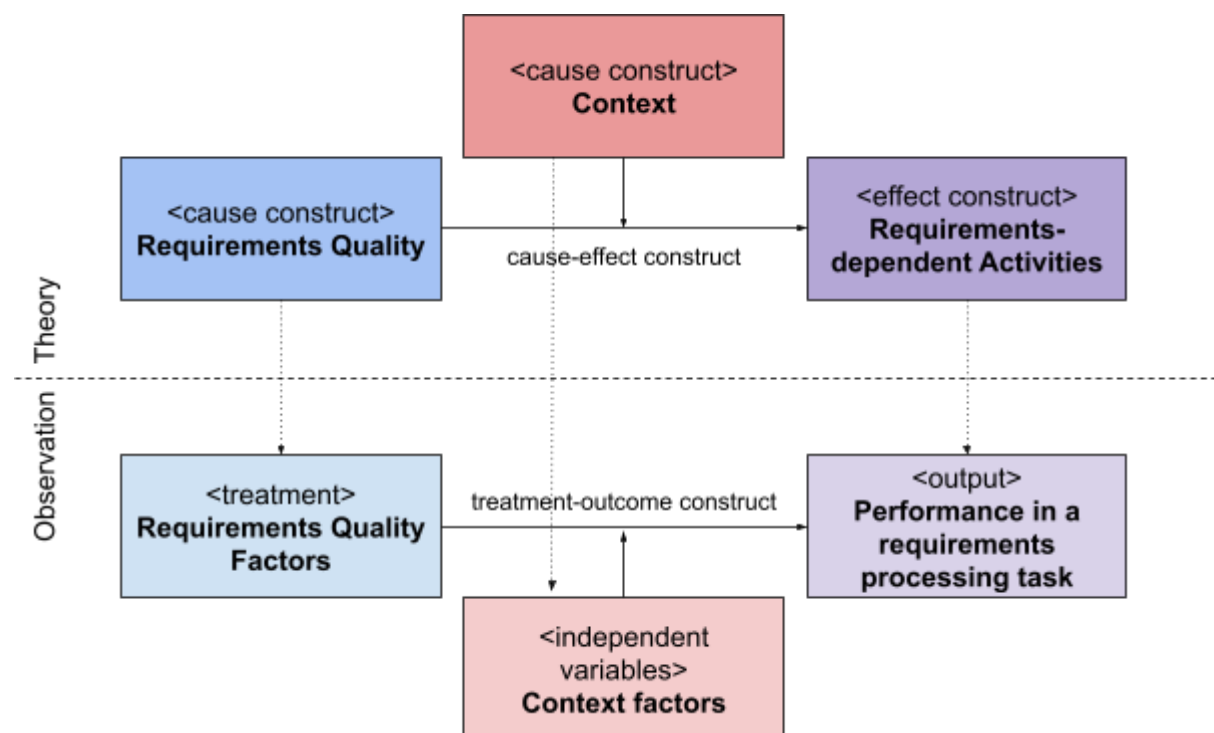# Experiment Design

Design of an experiment according to Wohlin et al. [1].

Fundamentals

Hypothesis

Our underlying hypothesis is that the **quality of natural language requirements** impacts **activities in which these requirements are used,** given the context in which those activities are performed.

Adapted experiment principles from Trochim et al. [2].

Experimental setup

During the experiment, we present subjects with NL requirements and expect them to translate the requirement into a domain model. For example, the requirement: "The system shall be capable of returning the search results after the user has entered the search criteria." is expected to be modeled as follows:
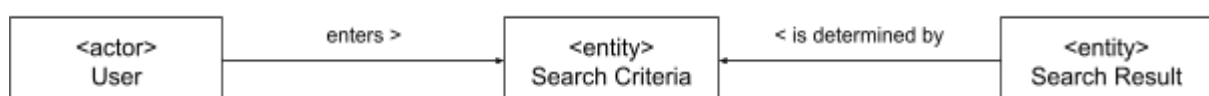
Components of the Experiment

Variables

The following list of variables will be considered in the experiment. Their type (column T) is classified as either dependent (dep), independent (ind), or context (con). The independent variable will be manipulated, while the context variables will only be sampled.

Variable	T	Data type	Values
Requirements Quality Defect	**ind**	categorical	[none, passive voice, ambigous pronoun, passive voice + ambiguous pronoun
~~Sentence length~~	~~con~~	~~Count~~	~~Number of characters~~
~~Number of tokens~~	~~con~~	~~Count~~	~~Number of tokens (words)~~
~~Count of nouns & verbs~~	~~con~~	~~Count~~	~~Number of nouns/verbs~~
~~Conjunctive Complexity~~	~~con~~	~~Count~~	~~Number of conjunctions (and/or) [3]~~
Experience in SE	con	Count	Years of working in SE
Experience in RE	con	Count	Years of working in RE
Experience in roles	con	Count matrix	List of roles, each annotated with a number of years
Level of education	con	Orginal	High-school, Bachelors, Masters, PhD
Experience with modeling	con	Ordinal	[never, rarely, occasionally, often]
Formal modeling training	con	boolean	true/false
Domain experience	con	Ordinal matrix	List of relevant domains, each annotated with a ordinal scale (1-5) of familiarity
~~Level of English proficiency~~	~~con~~	~~Ordinal~~	~~CEFR levels (A1-C2)~~
Missing actors	dep	Count	
Missing entities	dep	Count	
Missing associations	dep	Count	
Superfluous actors	dep	Count	
Superfluous entities	dep	Count	
Superfluous associations	dep	Count	

	Wrong associations			
	Time	dep	Count	Number of minutes/seconds to complete the task
	~~Consideration time~~	~~dep~~	~~Count~~	~~Minutes until the first entity is drawn~~
	~~Output/minute~~	~~dep~~	~~Ratio~~	~~Model elements drawn per minute~~
	~~Revisions…~~	~~dep~~	~~Count~~	~~Deleted elements per model~~

The independent variable leaves us with 4 treatment configurations. All **quality factors** and **context factors** are hypothesized to influence the **activity attributes**. In addition to that, we assume the following interrelations between context factors:

Variable	Impacts	Hypothesis
Native English speaker	Level of English proficiency	Being a native English speaker exposes a person to the language early and increases the likelihood of a high level of proficiency
Experience in RE	Experience in task	Domain modeling may be assumed to be common in RE.
Experience in roles (RE-related)	Experience in RE	The longer one has spent time in RE-related roles, the higher their RE knowledge

Objects

The objects of the experiment are the requirements containing the seeded defects.

	active voice	passive voice
No ambiguous pronoun	If <actor> does <action> with <system>, then the system <reaction>.	If <system> <**passive action**>, then the system <reaction>.
Ambiguous pronoun	If <actor> does <action> with <system>, **it** <reaction>.	If <system> <**passive action**>, **it** <reaction>.

References

[1] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media.

[2] Trochim, W. M., & Donnelly, J. P. (2001). *Research methods knowledge base* (Vol. 2). Macmillan Publishing Company, New York: Atomic Dog Pub..

[3] Antinyan, V., Staron, M., Sandberg, A., & Hansson, J. (2016, October). A complexity measure for textual requirements. In *2016 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement (IWSM-MENSURA)* (pp. 148-158). IEEE.