Empirical Study on Product Configuration and Traceability in UML-based Product-Lines

Thais Nepomuceno[1], Edson OliveiraJr[1], Raqueline Penteado[1], Marco A. Graciotto Silva[2], and Avelino F. Zorzo[3]

[1] State University of Maringá, Maringá-PR, Brazil
thais.nepomuceno1@gmail.com, {edson,raque}@din.uem.br
[2] Federal University of Technology - Paraná, Campo Mourão-PR, Brazil
magsilva@utfpr.edu.br
[3] Pontifical Catholic University of Rio Grande do Sul, Porto Alegre-RS, Brazil
avelino.zorzo@pucrs.br

Abstract. Software Product Line (SPL) can be defined as a set of systems that share common and variable parts. The elements that vary are those that allow the differentiation among products of this family, thus managing the variabilities is an important part of the SPL engineering. In the literature there is a lack of experiments that evaluate and compare approaches for SPL variability management. In this work, two of the existing approaches based on UML, SMarty and PLUS, are compared to verify: (i) the effectiveness of configuring products from use case diagrams and class diagrams; (ii) the influence the knowledge level of participants has when using each approach; (iii) the amount of consultations in the material needed for those using SMarty and PLUS; and (iv) the effectiveness in tracing variable elements from different diagrams. For this, we chose an experimental 1x2 factorial design. Although the overall results do not provide evidence of significant difference between the approaches, assumptions and improvements are discussed for future experiments.

Keywords: Empirical Study · Product Configuration · Software Product Lines · Traceability · UML.

1 Introduction

Software Product Line (SPL) is a steadily increasing approach as it brings several advantages to companies [9]. In SPL, software development targets a family of products and not individual products. To do so, there is an analysis of which characteristics of the software family are common and which are variable [3].

Elements which differentiate SPL products have associated variabilities, which is the ability to change or customize a system [3]. The importance of variability for an SPL makes its management an essential activity as its incorrect application can compromise the degree of SPL reuse [3]. There are several approaches based on feature models [2] and UML [8,19,16]: the former yields problem space, whereas the latter focuses on solution space. Morever, UML-based approaches can be used with current UML tools and builds upon known modeling concepts. Thus, in this study we address approaches on the solution space based on UML.

Despite the existence of several variability management approaches in the literature, they have not been properly evaluated using rigorous scientific methods [4,7]. In a tertiary study conducted by Raatikainen et al. [17], results showed that much of the empirical assessments of SPLs have not been adequately planned or reported.

Motivated by this scenario, we evaluated the management of variability, considering configurability and traceability, of two approaches: Product Line UML-based Software Engineering (PLUS) method [8], and SMarty approach [16].

PLUS is a widely referenced method in the literature and is an important example of a method for managing variabilities based on UML (use cases and classes), using stereotypes to represent these variabilities. Similarly, SMarty defines an UML profile and uses stereotypes and tagged values, supporting also rastreability. This paper answers the following research question: "Can we have more effectiveness and efficiency, at configuring SPL products, and provide traceability using SMarty compared to the PLUS method for use case and class diagrams?".

This study is part of a larger project of experimental evaluation of the SMarty approach throughout Continuous Experimentation, a process of systematic experimentation that has been studied in several aspects such as variability representation capability, product configuration, and traceability. Therefore, the results obtained by Marcolino et al.'s [10,11,12,14,13] and Nepomuceno and OliveiraJr [15] contributed significantly to the experimentation/learning process, proposed by Wohlin et al. [18], by allowing the overall evolution of SMarty. The main idea is that SMarty can be evaluated continuously, based on all SMarty supporting paired diagrams (use case, class, component, sequence, and activity), thus evolving from the results of the experiments. In this study, we considered use case and class diagrams for Mobile Media SPL [6] in the academic setting. Further studies are being prepared to address other UML diagrams and more complex SPL, resembling industrial context.

2 Background and Related Work

This section presents fundamentals on the PLUS method [8] and the SMarty approach [16], as well as related work.

2.1 Product Line UML-based Software Engineering (PLUS)

The PLUS method [8] was conceived for developing SPLs based on UML artifacts. PLUS encompasses the following phases: Inception, which determines whether an SPL is feasible based on its context, functionalities, degree of similarity and variability; Elaboration, in which use case model and features are reviewed and elaborated in greater detail, identifying their variation points and the SPL architecture is expanded including optional and variant components;

Construction, in which components are developed and tested; and Transition, in which components are integrated and made available for users to test.

PLUS allows the identification of variant components through the use of UML stereotypes to represent variability: «kernel» used in mandatory use case and class diagram elements; «optional» applied to use case and class diagram elements that might be present in a specific SPL product; and «alternative» used only for use case mutually exclusive elements. The PLUS method does not provide any kind of support in traceability of elements among different diagrams.

2.2 Stereotype-based Management of Variability (SMarty)

SMarty [16] is a variability management approach composed of an UML profile, the SMartyProfile, and a process for guiding users on how to identify and represent variabilities, the SMartyProcess.

The SMartyProfile encompasses a set of stereotypes and tagged-values for representing variability, variation points and variants in use case, class, component, sequence, and activity diagrams [16]. Such stereotypes are: «variability» represents the concept of variability; «variationPoint» represents a variation point; «mandatory» applied to variants present in every product;«optional» represents a variant which might be present in a product; «alternative_OR» applied to variants of an inclusive group; «alternative_XOR» applied to variants of an exclusive group. The SMartyProcess [16] guides users on identifying, representing, and tracing variabilities by means of a systematic process.

In the SMarty approach, traceability is performed based on the tagged-values of the «variability» stereotype: `realizes+` and `realizes-`. An example of traceability is depicted in Figure 1.

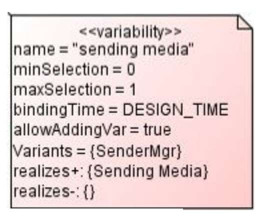

Fig. 1. Tracing elements in the SMarty «variability» tagged-values.

For the identification of traceable elements in a higher level of abstraction, we use the "+" symbol. For lower level of abstraction, we use the "-" symbol. In Figure 1 `realizes+` is a collection of variability names of a high-level element, in this case a variability named "Sending Media" in the use case diagram. Therefore, traceability of elements is facilitated, thus providing round trip course.

2.3 Related Work

Based on non-systematic searches and on the works of Ahnassay et al. [1], Chen and Babar [4] and Galster et al. [7], no work in the literature is directly related to the comparative experimental evaluation between SMarty and PLUS for generating product configurations and traceability based on use case diagrams and class diagrams. However, our research group has been developing several experiments that aim to show the effectiveness of SMarty. Several studies were conducted for identification and resolution of variabilities, and product configuration, while this is the first one aiming at analyzing traceability.

In 2013, Marcolino et al. [12] compared SMarty with PLUS in relation to identification and resolution of variabilities in use case diagrams. In such experiment, SMarty showed to be more effective than PLUS.

In 2014, Marcolino et al. [11] compared SMarty and PLUS to identify and solve variabilities in class diagrams. In such experiment, the PLUS method showed to be more effective. In 2017, two other experiments were conducted by Marcolino et al. [14] and Marcolino and OliveiraJr [13] comparing SMarty and PLUS in class diagrams. In the first experiment, there was no statistical difference between the effectiveness in relation to the ability to interpret and configure products. In the second experiment, PLUS evidenced to be more effective for identification and resolution of variabilities.

In 2018, Nepomuceno and OliveiraJr [15] evaluated the SMarty and PLUS approaches to the effectiveness of configuring correct products using the SPL AGM in use case diagrams. The results showed no advantage from one approach to another.

3 The Empirical Study

This section presents an experimental evaluation to compare the effectiveness of the SMarty approach and the PLUS based on the Wohlin et al.'s template [18].

3.1 Goal and Research Questions

The goal of this experiment is to **compare PLUS and SMarty**, **with the purpose of** identifying which one is more effective, **with respect to** configuration of specific products and tracing elements in use case diagrams and class diagrams, **from the point of view of** researchers in the role of SPL architects, **in the context of** undergraduate students from the State University of Maringá and the Federal University of Technology - Paraná, Campo Mourão campus.

To do so, we defined the following research questions:

RQ.1: Which approach is more effective at deriving SPL product configurations from use case diagrams and class diagrams?

RQ.2: Which approach requires less consultations to the instructional materials for configuring SPL products from use case diagrams and class diagrams?

RQ.3: Is there any influence of the knowledge level of participants on UML and SPL/variability for generating SPL products from use case and class diagrams?

RQ.4: Which approach is more effective in tracing elements in use case and class diagrams?

3.2 Planning

Context Selection This experiment is characterized as off-line. We carried out it in two days: one day at the State University of Maringá (UEM) and one day at the Federal University of Technology - Paraná (UFTPR-CM)

Hypotheses Formulation We formulated the following hypotheses (replace "eff" with "effectiveness" or "number of consultations", "act" with "product configuration" or "traceability", and "diag" with "use case diagram" or "class diagram"):

- Null Hypothesis ($H0_{eff\_act\_diag}$): there is no significant difference in the effectiveness or number of consultations of PLUS and SMarty in configuring products or traceability for use case diagrams or class diagram.
  $\mathbf{H0}_{eff\_act\_diag}$**:** μ**(SMarty)** $=$ μ**(PLUS)**

- ($H1_{eff\_act\_diag}$): there is a significant difference in the effectiveness of PLUS and SMarty in configuring products or traceability for use case diagrams or class diagram.
  $\mathbf{H1}_{eff\_act\_diag}$**:** μ**(SMarty)** $\neq$ μ**(PLUS)**

Variables and Metrics Selection We selected the following variables:

- **Independent variables:** approach being analyzed, which is a factor with two treatments: the PLUS method and the SMarty approach; pre-fixed variable for SPL, the Mobile Media, which was chosen since it has several types of variabilities and possibilities for product configuration, yet with diagrams simple enough to be used for an study without tool support.
- **Dependent variables:** effectiveness of configuring products, number of consultations to instructional material, influence of participant's knowledge on UML and SPL/variability, and effectiveness at traceability.

To calculate effectiveness we used the following equation:

$$\textbf{Effectiveness (z,d)} = \textbf{nCorrElem/TotalElem}$$

where:

- z = PLUS or SMarty;
- d = use case diagram or class diagram;

– nCorrElem = number of correct resolved/traced variability elements of a given diagram d using the z approach; and
– TotalElem = total number of variability elements of a given diagram d using the z approach.

Note that effectiveness will always be a value between 0.0 and 1.0. The number of consultation in the material is a discrete positive number. The influence of the participants knowledge is represented by a Likert scale with five options: none, only reading, basic, moderated, advanced. Such influence will be calculated as a correlation between the Likert scale options and the effectiveness at configuring products.

Selection of Subjects Participants were selected in a non-probabilistic manner. All participants from the two universities are undergraduate students enrolled in Computer Science programs.

Choice of Design Type We chose a 1x2 independent factorial experimental design type as treatment, and control has neither interactions nor relationships.

Instrumentation The following materials compose our instrumentation: Informed Consent Term (ICT): contains main information on the experiment to be applied, such as, confidentiality, procedures and benefits; characterization questionnaire, applied to participants to analyze the level of knowledge and experience on UML and SPL/variability; instructional material, with three parts, the first with main concepts of SPL, the second with general description of SPL Mobile Media, and the third with specific concepts for PLUS and SMarty; two use case diagrams of the SPL Mobile Media, one modeled using SMarty and another according to PLUS; two class diagrams of the SPL Mobile Media, one modeled according to SMarty and another according to PLUS; two documents with questions on product configuration and traceability for PLUS and SMarty.

3.3 Operation

The experiment period was one day for each university. For each period, there was a training on general concepts on SPL and, then, for each group, training specific for the given SPL technique and, finally, the experiment tasks.

Training We trained all participants in SPL and variability modeling, configuration of products and traceability. In addition, we trained half of participants using SMarty and the other half using PLUS. Training lasted 35 minutes at both universities. During training, participants received three documents: the ICT, the characterization questionnaire, and the instructional material. Training was based on excerpts of hypothetical SPL use case and class diagrams. Participants were allowed to ask questions at any time during training sessions.

Operation Procedures The following steps were taken in this experiment: experimenter divided participants into two groups balancing them according to their knowledge on UML and SPL/variability; experimenter informed participants they were allowed to use the instructional material during the experiment tasks; experimenter randomly distributed to participants an use case diagram and a class diagram according to PLUS or SMarty, and respective document with questions and tasks to be performed; experimenter asked participants to take notes on how many times they consult the instructional material during tasks; participants configured one specific instance (product) based on both use case and class diagrams from the Mobile Media SPL; participants answered questions on product configuration and traceability by analyzing whether the approach allows identifying variability from one level to another (use case to class diagram traceability) and vice-versa. Participants responded to a couple open questions on improvements and difficulties faced in the experiment; participants gave instruments to experimenter and finished their tasks in the experiment.

3.4 Analysis and Interpretation

Fourty-six (46) participants attended the experiment, 26 participants used the PLUS method and 26 the SMarty approach. All participants had at least basic knowledge of UML, and the median knowledge level was moderate. Regarding SPL/variability, the median knowledge level was low, with 31 with no knowledge on SPL and 15 that could read or had basic knowledge. The average experience was of 33.8 months (median of 36, minimum of 8 and maximum of 48 months of experience). There was no significant difference for knowledge level or experience between participants in control and intervention groups.

Data collected from products derivation are shown in Tables 1 and 2, which refer to PLUS and SMarty, respectively. In such tables there are information on: the number of correct elements (`Correct`), the number of total elements (`Total`) and the effectiveness of the approach (`Effect.`) for each product, the traceability effectiveness from use cases to classes and vice-versa, and the number of consultations on the material. We also inform means, standard deviation and median of such values.

In Table 2, for the SMarty approach, we discarded data from three participants as they either did not complete the tasks requested in the document or due to deviations in the performed experimental tasks. The interpretation of the results were performed to all other participants. Mean, standard deviation and median on Table 2 do not take into account data from excluded participants. All analysis was performed using the R environment.

Effectiveness on Configuring Products (RQ.1) This section presents analysis on the effectiveness of each approach at configuring products.

Normality Test We tested normality of our obtained effectiveness samples with Shapiro-Wilk test. All samples were considered non-normal as p-value < 0.05.

Table 1. Observed values for PLUS.

	Product #1 (use case)			Product #2 (class)			Traceability effectiveness from use cases to classes	Traceability effectiveness from classes to use cases	Number of material consultations
	Correct	Total	Effect.	Correct	Total	Effect.			
Mean	9.47	10	0.94	9.39	10	0.93	0.67	0.23	0.78
St. Dev.	0.99	0.00	0.09	0.59	0.00	0.19	0.46	0.36	2.02
Median	10	10	1.0	10	10	1.0	1.0	0.0	0

Table 2. Observed values for SMarty.

	Product #1 (use case)			Product #2 (class)			Traceability effectiveness from use cases to classes	Traceability effectiveness from classes to use cases	Number of material consultations
	Correct	Total	Effect.	Correct	Total	Effect.			
Mean	8.9	10	0.89	7.35	10	0.73	0.57	0.40	**0.30**
St. Dev	3.05	0.00	0.30	3.80	0.00	0.38	0.43	0.47	0.80
Median	10	10	1.0	9	10	0.90	0.50	0.0	0

Hypothesis Test We chose Mann-Whitney-Wilcoxon as a non-parametric hypothesis test for configuring products as follows:

- Use Case Diagram:
  - the calculated p-value was $p = 0.1898$, which is greater than $\alpha = 0.05$;
  - the Cohen effect size test was applied to samples, thus we obtained $d = -0.4844$, which indicates a small difference in samples observed values;
  - the null hypothesis cannot be rejected, thus there is no significant difference between samples of PLUS and SMarty for use cases.
- Class Diagram:
  - the calculated p-value was $p = 0.02797$, which is less than $\alpha = 0.05$;
  - the Cohen effect size test was applied to samples, thus we obtained $d = 0.5168$, which indicates a medium difference in samples observed values;
  - the null hypothesis can be rejected, thus alternative hypothesis should be accepted. Therefore, PLUS obtained better results than SMarty for configuring products based on classes.

Number of Consultations to Materials (RQ.2) This section presents analysis of results on the number of times a participant consulted the provided instructional material. We intend to demonstrate how such material influences the effectiveness of configuring products for PLUS and SMarty.

Normality Test All samples were considered non-normal as p-value < 0.05.

Hypothesis Test We chose Mann-Whitney-Wilcoxon as a non-parametric hypothesis test for consulting materials on both use case and class diagrams as follows:

- the calculated p-value was $p = 0.6349$, which is greater than $\alpha = 0.05$;

– the Cohen effect size test was applied to samples, thus we obtained $d = 0.277$, which indicates a small difference in samples observed values;
– the null hypothesis cannot be rejected, thus we cannot claim the number of consultations on the material has influenced more one approach over another.

Participants Knowledge Influence (RQ.3) To analyze whether the participants knowledge has influenced more one approach over another, we performed correlations for each diagram. To do so, we converted each Likert scale option to a natural number from 1 to 5. We chose the Spearman correlation technique, as follows:

– Use Case Diagram:
  - for the PLUS method we obtained $\rho = 0.32$ which means a weak positive correlation; and
  - for the SMarty approach we obtained $\rho = 0.16$ which means a weak positive correlation.
– Class Diagram:
  - for the PLUS method we obtained $\rho = 0.24$ which means a weak positive correlation; and
  - for the SMarty approach we obtained $\rho = 0.41$ which means a weak positive correlation.

As we could observe, all correlations are positive and weak. It means: (i) there was no influence on the participants knowledge over the obtained effectiveness at configuring products in both use case and class diagrams; and (ii) PLUS and SMarty comprehensibility seems not dependent on knowledge, as undergraduate students could perform valid configurations on both PLUS and SMarty.

Effectiveness on Tracing Elements (RQ.4) This section presents analysis on the effectiveness on tracing elements from use cases to classes and vice-versa.

Normality Test All samples were considered non-normal as p-value < 0.05.

Hypothesis Test We chose Mann-Whitney-Wilcoxon as a non-parametric hypothesis test for the effectiveness on tracing elements as follows:

– From Use Cases to Classes:
  - the calculated p-value was 0.6442, which is greater than $\alpha = 0.05$;
  - the Cohen effect size test was applied to samples, thus we obtained $d = 0.078$, which indicates an insignificant difference in the observed values;
  - the null hypothesis cannot be rejected, thus there is no significant difference between samples from traceability from use case to class diagrams.
– From Classes to Use Cases:
  - the calculated p-value was 0.176, which is greater than $\alpha = 0.05$
  - the Cohen effect size test was applied to samples, thus we obtained $d = -0.488$, which indicates a small difference in samples observed values;
  - the null hypothesis cannot be rejected, thus there is no significant difference between samples from traceability from class do use case diagrams.

3.5 Validity Evaluation

We identified the following threats and tried to mitigate them. Regarding **internal validity**, all participants are undergraduate with no significant difference in the students skills, thus providing some kind of homogeneity in samples. We tried to mitigate it by inviting students from two different universities. We tried to mitigate lacking of knowledge in SPL/variability by leveling participants in training sessions. We provided them with at least same contents and materials. Therefore, we understand their answers are valid and significant. We avoided communication among participants with a human observer during the experiment tasks all the time.

External Validity: we understand the most external threat to this experiment was the instrumentation with regard to use case diagrams and class diagrams. As they do not come from a commercial and actual SPL, we tried to mitigate it using a well-known SPL in the academic environment, the Mobile Media. The knowledge level of participants might be a threat as they have different experiences in UML and SPL/variability from each university. We tried to mitigate it by performing training sessions and taking Falessi et al. [5] recommendations for granted in this experiment. Accordingly to such authors, populations should not be considered better than other.

Construct Validity: study instrumentation and its validity are major potential construct threats to this experiment. To mitigate it, we performed a pilot project with four master's students from UEM with grounded knowledge in SPL/variability aiming at evaluation of such instrumentation and potential improvements.

Conclusion Validity: the main conclusion validity of this experiment is related to: (i) environment, as only two universities were chosen; (ii) only 41 (46 minus 5 excluded) students took part of the experiment, lowering the behavioral statistical accuracy; and (iii) participant experiences should be more explored towards experts.

4 Discussion on Results

We discuss results in various perspectives in the next sections.

Configuration of Products (RQ.1) We have verified there is evidence that the SMarty approach, in its current version, does not stand out in relation to the PLUS method in the effectiveness of configuration of products based on use case diagrams of the Mobile Media SPL. Such lack of difference is corroborated by a very low effectiveness size value. As for the class diagram, we realized the participants who made use of the PLUS method had more effective product configurations than those who used the SMarty approach, corroborated by a high effectiveness size value.

Although using a more complex SPL compared to that used by Nepomuceno and OliveiraJr [15], our assumption is that the use case diagram of Mobile Media still has few use cases and is relatively simple to understand and configure. In this way, solving the variabilities and configuring specific products was not a complex task, even with the PLUS method that has no guidelines and a support process as the SMarty approach.

Another assumption that arises is the relative simplicity of configuration of products with PLUS, since their diagrams are modeled with only three stereotypes. This may have led to a greater effectiveness in configuring products by the participants based on the medians and standard deviations of each approach.

Usage of Instructional Material (RQ.2) As we observed, the number of consultations to the instructional material did not significantly influenced the effectiveness at configuring SPL products in both use case and class diagrams of one approach over another. Such claim is corroborated by a small effect size value.

With few consultations to materials, the following assumptions rise: training might provide enough basis on PLUS and SMarty for participants performing their tasks; the used diagrams from Mobile Media have few stereotypes representing variability aspects, thus there is no such difficult at syntactically and semantically comprehending them; as discussed in external validity, this might have occurred due to certain participants do not accurately inform the exact number of consultations.

Influence of Participants Knowledge (RQ.3) We verified there is no significant influence of participants knowledge at configuring products with PLUS and SMarty, corroborated with small values of correlations.

On the one hand, we noticed that for the configuration of a product using use case diagram, the knowledge of the participants who used the PLUS method had greater influence than SMarty. On the other hand, when configuring products from the class diagram, less qualified participants were more effective with PLUS than with SMarty.

In the case of configuration of products with use case diagrams, our assumption is that the UML note of the «variability» stereotype might have positively influenced participants using SMarty, which facilitates the resolution of variation points and respective variants.

For configuring products from class diagrams we identified three potential assumptions: (i) PLUS has only two stereotypes, which might justify its simplicity in comprehending such diagrams and, consequently, valid configurations; (ii) the simplicity of Mobile Media class diagrams might also have influenced valid configurations; and (iii) both assumptions (i) and (ii) might make sense.

Traceability (RQ.4) The initial evidence for a better result of SMarty over PLUS was with respect to tracing elements between UML diagrams, especially from the class diagram to the use case diagram.

The PLUS method, unlike the SMarty approach, does not provide any kind of support for tracing elements between UML diagrams. Then, our main assumption is the SMarty UML note for variability, with tagged values `realizes+` and `realizes`- might help participants to obtain better results at tracing elements from class diagrams to use case diagrams.

Results might get some influence due to the lack of an actual SPL and relatively low complexity diagrams, especially with few variation points and variants among use cases and classes levels. This currently led us to plan and conduct a replication of this experiment with a more complex SPL.

SMarty Prospective Improvements and Lessons Learned. With the results of this experiment and participants feedback on open questions, SMarty should be improved.

Guidelines will serve as guidance for providing users specific scenarios according to the type of variability of each diagram, as well as to ease the traceability among elements from different diagrams at different abstraction levels, including navigability of such diagrams.

Another improvement will be an optional feature to use icons rather than stereotypes to represent variability, which might graphically ease the comprehension of modeled variabilities. Icons are very common instruments used by different UML profiles, thus it will not impact the UML metamodel in any way.

We are also currently working on an automated tool to support UML-based SPLs as there is no specific tool for such. The advantage of having this tool is its support to every SPL modeled with Meta Object Facility (MOF) based[4] languages.

Configuring a valid product is not a trivial task, even using a simple SPL. Although we provided training for participants as well as documentation for consultation, they still faced difficulties to correctly set up products in both approaches. Thus, we understand that an automated tool will mitigate such issue. We are currently developing an automated environment to support the whole lifecycle of SMarty-based SPLs.

Although SMarty has support for traceability of variabilities, participants complained on the lack of a guide to do so. Therefore, we believe a set of guidelines should aid to execute this activity.

We tried to increase complexity of adopted SPL moving from the SEI's SPL to Mobile Media. To mitigate this issue, we understand more complex Mobile Media elements are needed for prospective experiments.

5 Conclusion

This paper presented an experiment conducted to compare PLUS and SMarty variability management approaches with relation to four research questions. All instrumentation and data of this experiment is available at `https://doi.org/10.5281/zenodo.3569422`.

With respect to the first research question, the results showed a slight advantage of the PLUS method in relation to the SMarty approach in the configuration

[4] `https://www.omg.org/mof/`

of products from class diagrams. For the use case diagram there was no evidence benefiting one over another.

The number of consultations to the instructional material of each approach did not influence the calculated value of respective effectiveness. This is a great result as we expected participants to understand the approaches comprehensibility in a short period of time.

There is evidence the participant's previous knowledge level in SPL/variability and UML may be related to better results on the use of the approaches for configuring products, especially with regard to the SMarty approach. Thus, less experienced participants were able to configure products with greater effectiveness with SMarty.

It is understood that the SMarty approach provides more subsidies for element tracing between the diagrams, however this aspect still needs to be further experienced in more complex SPLs.

We understand the results of this experiment together with those of the previous experiment for class and use case diagrams provided further subsidies for investing in an SPL approach that focuses on UML diagrams and consequently produces tooling support for UML-based SPL life cycle.

Our next planned experiments, using sequence, activity, and component diagrams, will consider assumptions discussed in this experiment as a way to evolve the SMarty approach, since the purpose is that it is more effective in the identification of variabilities, in the configuration of products and in the tracing of elements among diagrams in relation to the other existing approaches. We shall consider SPLs with more complexity at configuring products.

References

1. Ahnassay, A., Bagheri, D., D, G.: Empirical evaualtion in sofware product line engineering. Tech. rep., Ryerson University (2013), `http://ls3.rnet.ryerson.ca/wp-content/uploads/2013/08/TR-LS3130084R4T.pdf`, tR-LS3-130084R4T
2. Apel, S., Batory, D., Kästner, C., Saake, G.: Feature-Oriented Software Product Lines: Concepts and Implementation. Springer Publishing Company, Incorporated, 1st edn. (2016)
3. Capilla, R., Bosch, J., Kang, K.C.: Systems and Software Variability Management: Concepts, Tools and Experiences. Springer, USA (2013)
4. Chen, L., Ali Babar, M.: A systematic review of evaluation of variability management approaches in software product lines. Inf. Softw. Technol. **53**(4), 344–362 (Apr 2011)
5. Falessi, D., Juristo, N., Wohlin, C., Turhan, B., Münch, J., Jedlitschka, A., Oivo, M.: Empirical software engineering experts on the use of students and professionals in experiments. Empirical Software Engineering **1**(23), 452–489 (06 2017)
6. Figueiredo, E., Cacho, N., Sant'Anna, C., Monteiro, M., Kulesza, U., Garcia, A., Soares, S., Ferrari, F., Khan, S., Castor Filho, F., Dantas, F.: Evolving software product lines with aspects. In: 2008 ACM/IEEE 30th International Conference on Software Engineering. pp. 261–270 (May 2008). https://doi.org/10.1145/1368088.1368124

7. Galster, M., Weyns, D., Tofan, D., Michalik, B., Avgeriou, P.: Variability in software systems—a systematic literature review. IEEE Transactions on Software Engineering **40**(3), 282–306 (March 2014)
8. Gomaa, H.: Designing Software Product Lines with UML: From Use Cases to Pattern-Based Software Architectures. Addison Wesley Longman, Redwood City, CA, USA (2004)
9. Linden, F.J.v.d., Schmid, K., Rommes, E.: Software Product Lines in Action: The Best Industrial Practice in Product Line Engineering. Springer, Berlin (2007)
10. Marcolino, A., OliveiraJr, E., Gimenes, I.: Variability identification and representation in software product line uml sequence diagrams: Proposal and empirical study. In: SBES. pp. 141–150 (2014). https://doi.org/10.1109/SBES.2014.11
11. Marcolino, A., OliveiraJr, E., Gimenes, I., Barbosa, E.F.: Empirically based evolution of a variability management approach at uml class level. In: COMPSAC. pp. 354–363 (2014)
12. Marcolino, A., OliveiraJr, E., Gimenes, I.M.S., Maldonado, J.C.: Towards the effectiveness of a variability management approach at use case level. In: SEKE. pp. 214–219. KSI Institute (2013)
13. Marcolino, A.S., OliveiraJr, E.: Comparing smarty and plus for variability identification and representation at product-line uml class level: A controlled quasi-experiment. Journal of Computer Science **13**(11), 617–632 (2017). https://doi.org/10.3844/jcssp.2017.617.632
14. Marcolino, A.S., OliveiraJr, E., Gimenes, I.M., Barbosa, E.F.: Variability resolution and product configuration with smarty: An experimental study on uml class diagrams. Journal of Computer Science **13**, 307–319 (2017). https://doi.org/10.3844/jcssp.2017.307.319
15. Nepomuceno, T.S., OliveiraJr, E.: Configuring software product line specific products with smarty and plus: An experimental study on use case diagrams. In: SBQS. pp. 81–90. ACM (2018). https://doi.org/10.1145/3275245.3275254
16. OliveiraJr, E., Gimenes, I.M.S., Maldonado, J.C.: Systematic management of variability in uml-based software product lines. Journal of Universal Computer Science (JUCS) **16**(17), 2374–2393 (sep 2010)
17. Raatikainen, M., Tiihonen, J., Männistö, T.: Software product lines and variability modeling: A tertiary study. Journal of Systems and Software **149**, 485 – 510 (2019)
18. Wohlin, C., Runeson, P., Hst, M., Ohlsson, M.C., Regnell, B., Wessln, A.: Experimentation in Software Engineering. Springer, USA (2012)
19. Ziadi, T., Jezequel, J.M.: Software product line engineering with the uml: Deriving products. In: Käköla, T., Duenas, J.C. (eds.) Software Product Lines, pp. 557–588. Springer, Berlin (2006)