The Influence of Software Design Representation on the Design Communication of Teams with Diverse Personalities

Rodi Jolak
rodi.jolak@cse.gu.se
Chalmers | University of Gothenburg
Sweden

Maxime Savary-Leblanc
maxime.savary-leblanc@univ-lille.fr
University of Lille
France

Manuela Dalibor
dalibor@se-rwth.de
RWTH Aachen University
Germany

Juraj Vincur
xvincurj@stuba.sk
Slovak University of Technology
Slovakia

Regina Hebig
regina.hebig@cse.gu.se
Chalmers | University of Gothenburg
Sweden

Xavier Le Pallec
xavier.le-pallec@univ-lille.fr
University of Lille
France

Michel Chaudron
m.r.v.chaudron@tue.nl
Eindhoven University of Technology
Netherlands

Sébastien Gérard
sebastien.gerard@cea.fr
CEA LIST
France

Ivan Polasek
ivan.polasek@fmph.uniba.sk
Comenius University Bratislava
Slovakia

Andreas Wortmann
wortmann@isw.uni-stuttgart.de
University of Stuttgart, ISW
Germany

ABSTRACT

Software is the main driver of added-value in many of the systems that surround us. While its complexity is increasing, so is the diversity of systems driven by software. To meet the challenges emerging from this combination, it is necessary to mobilize increasingly large and heterogeneous multidisciplinary teams, comprising software experts, as well as experts from various domains related to the systems driven by software. Hence, the quality of communication about software between stakeholders of different domains and with different personalities is becoming a key issue for successfully engineering software-intensive systems. The goal of this study, thus, is to investigate the effect of the representation of software design models on the communication of design decisions between stakeholders with diverse personality traits. As a result, this study finds that graphical representations of software design models are better than textual representations in enhancing the communication and increasing the productivity of stakeholders with diverse personalities.

CCS CONCEPTS

• Software and its engineering → System description languages; **Designing software**; **• Social and professional topics**;

MODELS '22, October 23–28, 2022, Montreal, QC, Canada

ACM 978-1-4503-9466-6/22/10…$15.00
https://doi.org/10.1145/3550355.3552398

KEYWORDS

Software Engineering, Software Design, Human Aspects, Personality Traits, Communication

ACM Reference Format:
Rodi Jolak, Maxime Savary-Leblanc, Manuela Dalibor, Juraj Vincur, Regina Hebig, Xavier Le Pallec, Michel Chaudron, Sébastien Gérard, Ivan Polasek, and Andreas Wortmann. 2022. The Influence of Software Design Representation on the Design Communication of Teams with Diverse Personalities. In *Proceedings of MODELS '22: Proceedings of the 25th International Conference on Model Driven Engineering Languages and Systems (MODELS) (MODELS 2022).* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3550355.3552398

1 INTRODUCTION

Communicating knowledge about software development and design decisions is one of the core activities in software and systems engineering [14]. This knowledge, representing pieces of a software design model, is often either encoded and presented *textually* as formatted text paragraphs or *graphically* as software diagrams. On one hand, the human mind processes these two forms of knowledge representation differently [23]. On the other hand, personality traits influence human communication and can, thus, influence team productivity and software quality [1, 3]. Hence, selecting the right choice of representation can make software design communication between stakeholders of different personalities more efficient. To better address the challenges arising from the collaboration between different personalities in software engineering, this study aims at understanding how stakeholders communicate design decisions depending on their personality and how this can be applied to support software engineering teams. This study reproduces a face-to-face, co-located communication routine between two stakeholders for a knowledge transfer purpose. One stakeholder (i.e., the knowledge owner) explains a software design of a system to another stakeholder (i.e., the knowledge receiver). This task echoes several common everyday routines in software engineering companies, such as

- informing a newcomer in a software development team about the undergoing project,
- transferring knowledge from one company to another about a software product, or
- discussing a feature or design between two team members.

More often than not, these routines involve teams made up of varied profiles, expressing potentially diverse personality traits. Hence, these traits may affect software design communication. In particular, we address the following research question:

RQ.1 *How does the representation (textual vs. graphical) of a software design model influence the design communication between developers with diverse personalities?*
To investigate this, we assume the following two definitions [32, 33]:

Definition 1. *Textual Representation of Software Design Models*: A software design representation that spatially arranges and reflects the organization of design decisions by using textual notations, e.g., natural languages.

Definition 2. *Graphical Representation of Software Design Models*: A software design representation that provides a visuospatial arrangement of design decisions, and can recraft these decisions into a multitude of forms by using graphical notations, e.g., nodes and links.

In model-driven development (MDD), often graphical and textual syntax for the same language exist. For instance, both UML and SysML are often considered graphical modeling languages, but there are textual variants of UML [5, 27] and the SysML v2 is even being developed as a textual language first [30]. Hence, the results of this study can guide researchers and practitioners in using the most appropriate representations for communicating software knowledge manifested in textual or graphical representations.

To investigate design communication, we consider three interpersonal communication dimensions defined by [13], as often exhibited during collaborative knowledge sharing [29]:

- *Active Discussion*: questioning and informing others.
- *Creative Conflict*: arguing about others' discussions.
- *Conversation Management*: acknowledging and maintaining communicated knowledge.

The quality of interpersonal communication is determined by the amount of active discussions and creative conflicts, and the effort spent on conversation management. In particular, active discussion is considered a core activity of knowledge sharing and effective teams communication [29]. Moreover, arguing and reasoning in creative conflicts yield divergence, which contributes in supporting knowledge building and sharing [22, 29]. In contrast, spending considerable effort on conversation management hinders knowledge building that can be otherwise achieved via active discussion and creative conflict and, thus, reduces the quality of the communication [29].

To investigate how personality predispositions effect into the previous interpersonal communication dimensions, we use the Big Five model of personality [4]. This model is empirically validated and distinguishes five fundamental dimensions of personality that have been found in individuals from different cultures [20]. The premise of this model is that each factor predisposes a person to behave in a certain way [24]:

- *Extraversion*, the extent of being social and talkative.
- *Agreeableness*, the extent of being gentle and cooperative.
- *Conscientiousness*, the extent of being well-organized.
- *Neuroticism*, the extent of being calm and poised.
- *Openness*, the extent of being imaginative and curious.

In this study, we collect personality traits data and analyze the software design communication of 196 software engineering students from four European universities. In summary, we find that:

(1) Graphical design representations are better than textual representations in promoting *active discussion* between stakeholders, given that there is no high difference in *openness* level between the stakeholders.
(2) Graphical design representations are better than textual representations in reducing the *conversation management* effort between stakeholders, given that there is no low difference in *conscientiousness* level between the stakeholders.
(3) By using textual design representations, we observe that as the difference in *extraversion* level between the stakeholders increases, the frequency of *creative conflict* discussions decreases and the effort of *conversation management* increases. The correlation is statistically significant.

Accordingly, we encourage using software design diagrams (graphical representations of the software design model) in the communication routine of software engineering teams, especially in teams with diverse personality traits, as this has the potential to enhance the quality of communication of the teams and increase their productivity.

In the following, Section 2 discusses related work and Section 3 details the empirical study. Afterward, Section 4 presents the results, Section 5 summarizes implications, and Section 6 discusses threats. Finally, Section 7 provides concluding remarks.

2 RELATED WORK

The usefulness of languages depending on the application area is stressed by a literature review and online questionnaires of educators on various modeling languages for designing systems including Flowcharts, Entity-Relationship Diagrams, and UML [2]. Practitioners perceived graphical representations as easy to understand, but less suitable to formulate complicated relations. This leads to the question if there are tasks for which specific representations are appropriately suited [18].

The communication of design decisions is vital for software development. When several software engineers work together on a solution or when new employees are trained, it is essential to communicate and explain these decisions understandably. Since only little research covers the influence of representations on communication of software, our study can help to recommend representations for specific tasks.

A person's prior knowledge influences whether a graphical or textual representation is suitable. Experts are more efficient and effective in performing tasks if the external task representation matches the experts' internal representation of prior knowledge [26]. This means that experts work better if they get a graphical model for a task that is considered graphically affine, and textually vice versa. This effect is less noticeable but still present for beginners.

However, personality traits can also directly impact a person's communication skills and behavior [21]. Thus, further research on which representations suit which personality traits to supports software developers while describing their design decisions is also vital to improve team collaboration.

A literature review of 90 papers covering different aspects of teamwork in software engineering, such as pair programming, team effectiveness, and education reveals conflicting results regarding the influence of personality [3]. Particularly regarding the composition of pairs during pair programming, the study points out that there is no consensus on whether heterogeneous or homogeneous pairs are more productive. The same holds for education, where studies contradict each other concerning the influence of personality traits on academic success. All studies agree that team composition influences team effectiveness, but they contradict each other whether diversity or homogeneity is preferable when building teams.

Scrum is one of the main standards for agile software development and highly relies on frequent team meetings and open communication. By interviewing eleven members of Scrum teams from seven different companies [1], it was found that *agreeableness* is a crucial character trait for making Scrum meetings successful. In another study, data from 47 software engineers working at ten different software development organizations were collected [8]. These data uncovered that there is a significant influence of personality on team performance. In particular, it was found that *extroverted* software engineers often work more efficiently when working according to a schedule, and that *openness* is related to taking responsibility for a project. Other studies suggest that: (1) process-oriented communication correlates highly with *extraversion*, *openness* to experience, and *conscientiousness* [17]; (2) the success of teamwork activities might correlate with the participants' *extraversion* [25]; and (3) successful teamwork and performance are also dependent on the team's belief in the capabilities of the group and its cohesion [19].

In summary, the research results are diverse, but all studies confirm the influence of personality on communication. To give recommendations concerning team composition and analyze the impact of personality traits on communication, further research is needed, especially in software engineering, where agile methods rely on the regular successful transfer of knowledge between team members.

Jolak et al. [16] conducted an experiment to study the effect of software design representation (graphical vs. textual) on design communication. They found that using graphical design representations during design discussion is better than using textual design representations in fostering active discussions and enhancing the memory recall of design details.

Data on interpersonal communication used in our work are taken from the study of Jolak et al. [16]. In our work, we analyze data on personality traits, where the Big Five factors of personality: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness are collected. We analyze the distribution of the five personality traits across two groups: textual vs. graphical design representations. Moreover, we study the interaction of two independent variables:

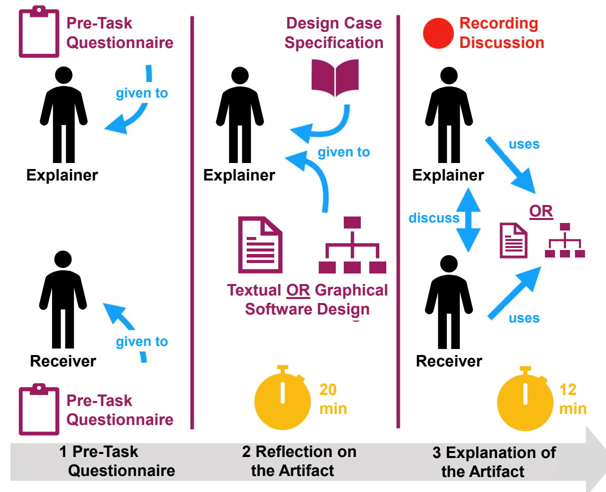

Figure 1: The three main tasks of the Whispers experiment.

design representation and *difference in personality traits*, and investigate how the interaction of these two variables influence the interpersonal communication. Finally, we analyze the correlation between the difference in personality traits and design communication per each studied group.

3 EXPERIMENTAL DESIGN

The goal of this study is to understand the effect of using graphical vs. textual software design representations on the communication of software engineering teams with diverse personalities. To achieve this goal, we designed and analyze data from an experiment conducted by Jolak et al. [16] with a goal to understand the effect of using a graphical vs. textual software design representation on the behavior and communication of software developers. The design of this experiment is inspired by the *Experience Engine* of the Ericsson company [12]. The *Experience Engine* is a practice of knowledge transfer between two individuals where one explainer (i.e., knowledge communicator) is in charge of explaining knowledge (e.g., artifacts of a specific systems) to a receiver who does not own such knowledge. The experimental design is detailed in the following subsections.

3.1 Participants

A mix of 196 B.Sc. and M.Sc. Software Engineering students are involved in the experiment. The participants were gathered following a *convenience sampling* approach. According to [31] and [11], this sampling approach enables the results to be generalized to junior and senior developers. Since the conductors of the experiment are academics, bachelor and master students who have followed a software design course in which UML is taught are selected. In order to limit pressure and stress bias, participants are informed that participation in the experiment would not influence their academic results. The experiment is replicated in four universities in four European countries to limit any cultural bias on communication. Finally, the major ethical issues according to [28] are considered: informed consent, beneficence– do not harm, and respect for anonymity and confidentiality.

3.2 Tasks

Figure 1 shows the three main tasks of the experiment. First, the Big Five factors of personality: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness are measured by asking the participants to fill a pre-task questionnaire containing a 20-item Mini-IPIP personality test [6]. Moreover, data about design experience and communication skills of the participants were collected. The questions are available in the experimental material online [15]. The participants are asked to anonymously fill the questionnaire using an ID number before the experiment session.

Second, the participants were randomly divided into two groups: the graphical group and the textual group. The participants of each group are furthermore randomly assigned one role, explainer or receiver. The explainers receive a design case specification and a graphical or textual software design of the specified case [15]. The design case describes a mobile application of a fitness center in natural language. This case is selected since it uses established concepts (e.g., sports).

The graphical and textual design representations provide the same information and describe one structural design of the fitness mobile application. Moreover, to ensure a systematic grouping in both representations, the well-known Model-View-Controller (MVC) pattern is employed. The graphical design is represented by a UML class diagram which features 17 attributes and 23 operations distributed across 28 classes (21 model entities, 3 controllers, and 4 views) linked by 30 relationships. Each element of the graphical design representation is systematically used to create exactly one corresponding element (e.g., one paragraph or sentence) in the textual design representation, thus to maintain a one-to-one correspondence between the two representations.

The explainers were asked to understand the received software design case and representation as good as possible. To do so, they are taken to a different room, separated from the receivers, where they are given the textual or graphical representation after a short explanation of the task to be done.

Third, by using either the textual or graphical software design representation, the explainers have to explain the design to the receivers. Definitely, the receivers could ask questions and discuss the design with the explainers. Moreover, during this phase the participants are asked to record the discussion using their laptop or smartphone.

We conducted pilot studies with a set of students matching the expected profile that we selected randomly. The pilot study participants had to perform each task of the experiment as described previously, in real conditions. By observing their progress and reporting their questions, we mainly adjusted the time allowed for each task, updated the general instructions for the experiment so no ambiguous sentence was provided, and prepared different solutions for students to record their discussions (software or mobile phone). This pilot study enabled us to set the time to 20 minutes for the design reading and mastering phase, and 12 minutes for the discussion.

3.3 Data Analysis

The communication between explainers and receivers is transcribed. The transcriptions are then analyzed and coded using the collaborative communication taxonomy of McManus and Aiken [22] (See Figure 2). First, coding's reliability is ensured by conducting two-way mixed Intraclass Correlation Coefficient (I-C-C) tests with 95% confidence interval on 9% of the data. In particular, the I-C-C value is 0.97 for the graphical group and 0.96 for the textual groups. Afterwards, the raters collaboratively continued to code the rest (i.e., 91%) of the data. As a result, data on *Active Discussions*, *Creative Conflicts*, and *Conversation Management are collected*. For instance, the following transcribed sentence: "*Can you explain why this component is needed?*" is a *Request* for *Justification* (See Figure 2) which contributes to *Active Discussion*.

3.4 Variables and Hypotheses

The independent variables are the *design representation* and *difference in personality traits*. The design representation is a nominal variable and corresponds to two treatments: textual design representation group and graphical design representation group. The difference in personality traits is an interval variable based on the subjective Mini-IPIP personality test questionnaire. Its values range between 0 and 4 and is measured by calculating the difference between the personality trait score of the explainers and receivers. Moreover, three dependent variables are considered (see Table 1). These variables correspond to the three considered interpersonal communication aspects: *active discussion*, *creative conflicts*, and *conversation management*. Hence, for this study, the following *null* $\mathbf{H}_0$ and *alternative* $\mathbf{H}_a$ hypotheses are investigated:

- $\mathbf{H}_0$: Software design representation and difference in personality traits *have no impact* on design communication.
- $\mathbf{H}_a$: Software design representation and difference in personality traits have impact on design communication.

4 RESULTS

In this section, we first analyze the distribution of the five personality traits across the two groups: textual vs. graphical design representations. Second, we study the effect of the two independent variables (i.e., design representation and difference in personality traits) and the interaction thereof on design communication. Finally, we analyze the correlation between the difference in personality traits and design communication per each group.

Dep. Var.	**Type**	**Source**	**Measurement**	
			Instrument	**Scale**
AD*	Ratio	Objective	Counting Occurrences in Recorded Conversations	AD/(AD+CC+CM) Values from 0 to 1
CC*	Ratio	Objective	Counting Occurrences in Recorded Conversations	CC/(AD+CC+CM) Values from 0 to 1
CM*	Ratio	Objective	Counting Occurrences in Recorded Conversations	CM/(AD+CC+CM) Values from 0 to 1

*__AD__: Active Discussion; **CC**: Creative Conflict; **CM**: Conversation Management

Table 1: Dependent variables (Dep. Var.) and measurement.

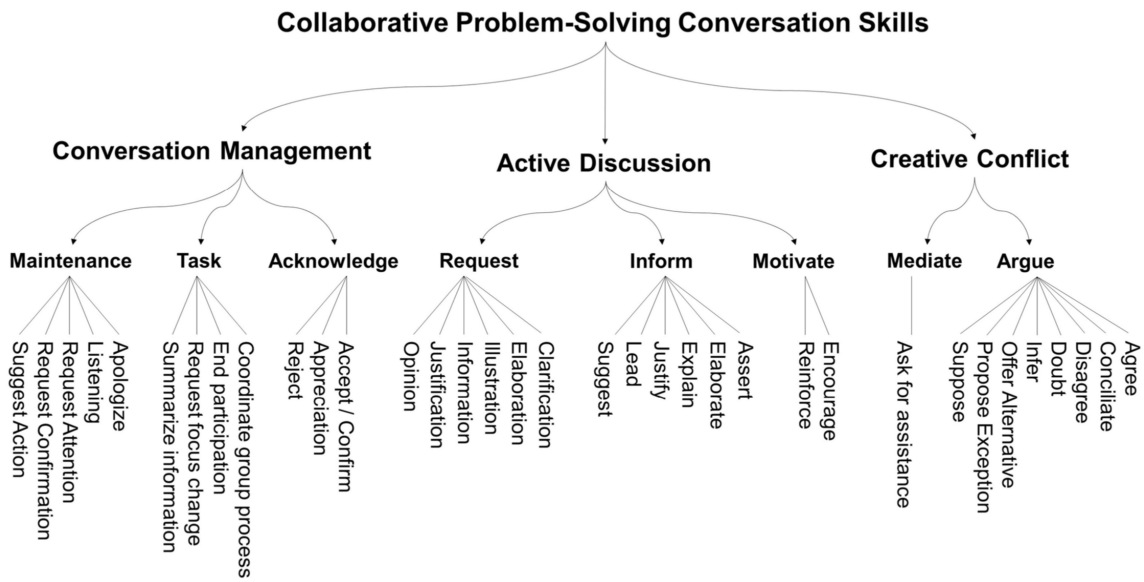

Figure 2: Collaborative interpersonal problem-solving communication skills [22].

4.1 Personality Traits Distribution

We evaluate whether the participants working in each group (textual vs. graphical design representations) are comparable concerning the personality traits.

In Table 2, we present the descriptive statistics by reporting; means, medians, and standard deviations of the five personality traits across the two groups. These descriptive statistics help to analyze central tendencies and dispersion of the five personality traits. We observe that the mean and distribution of the five personality traits across the two groups are similar. To statistically investigate this tendency, we conduct Mann-Whitney's independent-samples test. The results are presented in the last column (i.e., Mann-Whitney Test) of Table 2. The p-value is the probability of obtaining the observed results of a test, assuming that the null hypothesis is correct (i.e., there is a difference). We set the probability of type I error (i.e., α, probability of finding a significance where there is none) to 0.05. As a result, we find that the distribution of the five personality traits is the same across the two groups (i.e., there is no statistically significant difference). The effect size is medium for Neuroticism and small for the other personality traits.

Finding 1
The distribution of the personality traits of the participants is the same across the two groups: textual vs. graphical design representation (i.e., there is no statistically significant difference).

4.2 Effect of the Independent Variables and Their Interactions on Design Communication

We investigate the effect of design representation and difference in personality traits on design communication. Also, in order to understand whether these two independent variables are interacting with each other and, thus, creating a specific combined effect on the communication, we study the interaction effect. An interaction effect exists when the effect of an independent variable (i.e., the design representation) on a dependent variable (i.e., the communication) changes, depending on the value of one or more other independent variables (i.e., difference in personality traits).

We create interaction plots to study the single effect of the independent variables and interaction effect. Figure 3 shows the interaction plots. In particular, for each difference in personality traits between explainers and receivers, the mean ratio of the communication dimension by design representation (graphical vs. textual) is shown. The x-axis of each plot in Figure 3 shows the difference in personality traits. The values range between "0.00" and "4.00" and are measured by calculating the difference in the personality trait score between explainers and receivers. Since the maximum difference in personality traits in the collected data does not exceed "3.75", we consider a difference in personality traits as:

- *low*, when the difference is equal or higher than "0.00" and less than "1.25",
- *medium*, when the difference is equal or higher than "1.25" and less than "2.50", and
- *high*, when the difference is between "2.50" and "3.75".

The y-axis of each plot in Figure 3 shows the mean ratio of the communication dimension: *active discussion*, *creative conflicts*, and *conversation management*. The values range between "0.00" and "1.00" and are measured by calculating the ratio between the frequency of one specific communication dimension and the total frequency off all the communication dimensions in a design discussion. The graphical design representation is shown by a set of blue lines, whereas the textual design representation is shown by a set of red lines. These sets of lines represent how much communication happened when using a graphical (in blue) or textual (in red) design representation

Table 2: Personality traits difference between the two groups: textual and graphical design representation

	Textual Design Representation			Graphical Design Representation			Mann-Whitney Test	
Personality Trait	*mean*	*median*	*std. dev.*	*mean*	*median*	*std. dev.*	*p-value*	*effect size*
Extraversion	2.935	3.000	0.900	2.825	2.875	0.872	0.285	0.072
Agreeableness	3.726	3.750	0.905	3.805	3.750	0.600	0.804	0.046
Conscientiousness	3.432	3.500	0.738	3.318	3.250	0.717	0.253	0.293
Neuroticism	2.699	2.750	0.949	2.825	3.000	0.840	0.337	0.488
Openness	3.851	3.750	0.679	3.761	3.750	0.681	0.504	0.132

in a discussion between two individuals having certain differences in personality traits.

For instance, by looking at plot (1) (c) of Figure 3 we observe the following: by using the *textual design* representation (represented in red-lines/squares), pairs who have high difference in *extraversion* levels did more *conversation management* discussion during their communication of the software design. In contrast, the same plot indicates that by using the *graphical design* representation (represented in blue-lines/circles), pairs who have high difference in *extraversion* levels did less *conversation management* discussion during their communication of the design.

The plots of Figure 3 graphically transcribe the effect observed on mean values for all pairs, and help to visualize potential correlations before running the proper statistical tests.

By looking at Figure 3, we can preliminary investigate single and interaction effects as follows.

- A main effect of the design representation on the communication tends to occur when one of the two sets of lines (blue or red) is graphically above the other.
- A main effect of the difference in personality traits on the communication tends to occur when the mean value between each red-square and blue-dot is changing too much.
- An interaction effect tends to occur when the graphs of the two sets of lines (i.e., blue and red) are not parallel.

In order to inspect these effects statistically, we conduct a two-way Anova test. The results of this test are presented in Table 3. The statistical power is the probability that a test will reject a null hypothesis when it is in fact false. As the power increases, the probability of making a type II error (β-value) decreases. A power value of 0.80 is considered as a standard for adequacy [7]. β-value is used to estimate the probability of accepting the null hypothesis when it is false. By referring to Figure 3 and Table 3, we present the effects on each communication dimension in the following subsections.

4.2.1 Active Discussion. *Active discussion* accounts for actively questioning and informing others, and positively influences the quality of the communication [29]. By looking to the first row of plots in Figure 3 (1a-5a), we observe a possible effect of design representation on active discussions. Indeed, if we do not consider the high difference in *openness* trait, the set of blue lines is always above the set of red lines. This indicates a main effect of the design representation on the active discussion between explainers and receivers with (i) low, medium, or high difference in *extraversion*, *agreeableness*, *conscientiousness*, and *neuroticism* traits, and (ii) low and medium difference in *openness* trait. Table 3 shows that there is a statistically significant effect of design representation on active discussion (P-value < 0.05) when there is no high difference in *openness* trait between the explainer and receiver.

Regarding the main effect of the difference in personality traits on active discussion, Figure 3 indicates a change in the mean value between each red-square and blue-dot at each difference in *conscientiousness* trait. Table 3 shows a statistically significant effect of the difference in *conscientiousness* trait on active discussions (P-value < 0.05).

Regarding the interaction effect, the blue and red set of lines of the plots related to *active discussion* in Figure 3 are not parallel. This indicates a possible interaction effect. However, this effect is not statistically significant at the 0.05 level.

Pairs that used the graphical design representation outperformed the pairs that used the textual design representation in the amount of active discussions (mean = 0.509±0.911 vs. mean = 0.436±0.960). Therefore, the results show that using graphical representations of software designs tends to promote a higher communication quality than using textual representations by increasing the amount of *active discussions*. However, Figure 3 show that this tendency is affected by a high difference in *openness* trait. According to [17], *openness* to experience relates to the breadth, depth, originality, and complexity of an individual's mental and experiential life. In other words, individuals with high *openness* to experience are likely to be open to adventures, artistic, and creative. In contrast, individual with low *openness* to experience tend to be practical, traditional, and rational [34]. Previous work on *openness* [35] demonstrated that high heterogeneity in *openness* to experience in a team fosters knowledge sharing, which in turn generates and integrates new ideas and knowledge. In our study, Figure 3 indicates that high heterogeneity in *openness* triggers more *active discussion* for both design representations. This tendency, even though not statistically significant, is inline with the expected behaviour from [35], and might explain why high difference in *openness* trait mitigates the advantage of using the graphical design representation over the textual.

Finding 2
Graphical design representations are better than textual representations in promoting Active Discussion between stakeholders.

4.2.2 Creative Conflict. *Creative conflict* accounts for arguing and reasoning about others' discussion, which contributes in supporting knowledge building and sharing [22, 29]. By looking to the second row of plots in Figure 3 (1b-5b), we neither observe a possible effect of design representation nor difference in personality

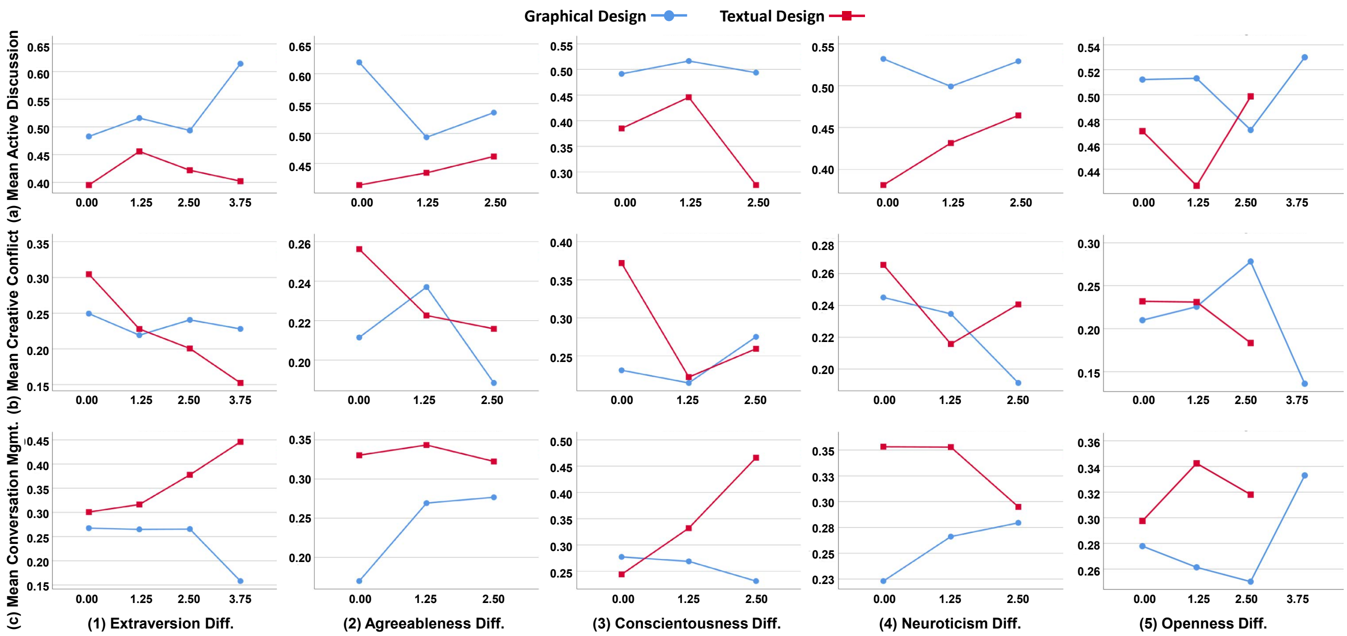

Figure 3: Interaction among the independent variables: design representation and difference in personality traits

traits on creative conflict. Yet, Table 3 shows that the effects of the design representation and difference in personality traits on creative conflict are not statistically significant (P-value > 0.05).

Regarding the interaction effect, the blue and red set of lines of the plots related to *creative conflict* in Figure 3 are not parallel. This suggests a possible interaction effect. However, this effect is not statistically significant at the 0.05 level.

Creative conflicts relate to arguing and reasoning, and positively influence the communication quality [22, 29]. Explainer-receiver pairs who used the graphical design representation generated an amount of creative conflicts (mean = 0.228±0.884) equivalent to what is generated by the pairs who used the textual design representation (mean = 0.228±0.931). Indeed, it is shown that neither the representation nor difference in personality traits has a statistically significant impact. In other words, as differences in personality traits appear, no general trend (statistically significant or not) can be deduced on the effectiveness of one design representation over the other.

4.2.3 Conversation Management.

Conversation management accounts for acknowledging and maintaining communicated information which hinders knowledge building that can be otherwise achieved via active discussion and creative conflict, and thus negatively influences the quality of the communication [29]. By looking to the third row of plots in Figure 3 (1c-5c), we observe a possible effect of design representation on conversation management. Indeed, if we do not consider the low difference in *conscientiousness* trait, the red set of lines is graphically above the blue set of lines in each plot. Table 3 shows that there is a statistically significant effect of design representation on conversation management (P-value < 0.05), when there is no low difference in *conscientiousness* trait. Regarding the main effect of the difference in personality traits on conversation management, we do not observe a change in the mean value between each red-square and blue-dot at each difference in personality traits. Statistically, Table 3 shows that the effect of the difference in personality traits on conversation management is not statistically significant (P-value > 0.05).

Regarding the interaction effect, the blue and red set of lines of the plots related to *conversation management* in Figure 3 are not parallel. This suggests a possible interaction effect. However, this effect is not statistically significant at the 0.05 level.

Pairs that used the graphical design representation performed less *conversation management* than those that used the textual design representation (mean = 0.263±0.103 vs. mean = 0.336±0.126). Therefore, the results show that using graphical representations of software designs tends to promote a higher communication quality than using textual representations by reducing the effort of *conversation management*. This tendency is affected by a low difference in *conscientiousness* trait, as shown by Figure 3. However, the lack of significant effect suggests that a difference in pairs' *conscientiousness* interacts with the nature of the design representation in a complex way so that no impact of the independent variables or their interaction can be observed on *conversation management*.

A potential explanation of this observation is that *conscientiousness* relates to the organizational skills of the participants, which is at the core of *conversation management*. A difference in *conscientiousness* trait might thus represent a difference in the problem-solving approach of the participants, especially in reading as well as comprehending the design task and steering the design communication. Such non-trivial interaction between the problem-solving approach and the pedagogical medium is the root of the Software Engineering education research field. Moreover, observing no effect of one

Table 3: Two-way Anova test for independent variables and their interaction effect on design communication

Independent Variable or Interaction	Dep. Var.	P-value	Power	β-value
Design Representation (DR)	AD	***0.001***	0.910	0.090
Extraversion Difference (EDiff)	AD	0.294	0.326	0.674
DR*EDiff	AD	0.576	0.185	0.815
Design Representation (DR)	CC	0.675	0.070	0.930
Extraversion Difference (EDiff)	CC	0.202	0.401	0.599
DR*EDiff	CC	0.407	0.259	0.741
Design Representation (DR)	CM	***0.003***	0.866	0.134
Extraversion Difference (EDiff)	CM	0.722	0.136	0.864
DR*EDiff	CM	0.246	0.363	0.637
Design Representation (DR)	AD	***0.000***	0.995	0.005
Agreeableness Difference (ADiff)	AD	0.122	0.431	0.569
DR*ADiff	AD	0.057	0.563	0.437
Design Representation (DR)	CC	0.435	0.121	0.879
Agreeableness Difference (ADiff)	CC	0.494	0.167	0.833
DR*ADiff	CC	0.484	0.170	0.830
Design Representation (DR)	CM	***0.003***	0.852	0.148
Agreeableness Difference (ADiff)	CM	0.331	0.241	0.759
DR*ADiff	CM	0.413	0.200	0.800
Design Representation (DR)	AD	***0.002***	0.896	0.104
Conscientiousness Difference (CDiff)	AD	***0.028***	0.668	0.332
DR*CDiff	AD	0.134	0.414	0.586
Design Representation (DR)	CC	0.267	0.197	0.803
Conscientiousness Difference (CDiff)	CC	0.109	0.452	0.548
DR*CDiff	CC	0.361	0.225	0.775
Design Representation (DR)	CM	0.082	0.413	0.587
Conscientiousness Difference (CDiff)	CM	0.447	0.185	0.815
DR*CDiff	CM	0.120	0.434	0.566
Design Representation (DR)	AD	***0.000***	0.966	0.034
Neuroticism Difference (NDiff)	AD	0.333	0.241	0.759
DR*NDiff	AD	0.401	0.205	0.795
Design Representation (DR)	CC	0.481	0.108	0.892
Neuroticism Difference (NDiff)	CC	0.512	0.160	0.840
DR*NDiff	CC	0.290	0.267	0.733
Design Representation (DR)	CM	***0.012***	0.772	0.228
Neuroticism Difference (NDiff)	CM	0.667	0.114	0.886
DR*NDiff	CM	0.315	0.251	0.749
Design Representation (DR)	AD	0.273	0.194	0.806
Openness Difference (ODiff)	AD	0.860	0.096	0.904
DR*ODiff	AD	0.236	0.307	0.693
Design Representation (DR)	CC	0.446	0.118	0.882
Openness Difference (ODiff)	CC	0.728	0.134	0.866
DR*ODiff	CC	0.307	0.256	0.744
Design Representation (DR)	CM	0.137	0.318	0.682
Openness Difference (ODiff)	CM	0.900	0.085	0.915
DR*ODiff	CM	0.727	0.100	0.900

specific design representation indicates that there is no perfect representation for everyone, but maybe an ideal representation for one person performing a specific task.

Finding 3
The graphical design representation is better than the textual representation in reducing the conversation management effort between stakeholders.

4.3 Correlation between Difference in Personality Traits and Interpersonal Communication

Finally, we analyze the correlation between the difference in personality traits and interpersonal communication per each group: textual and graphical design representation. To this end, we conducted Spearman's Rank-Order correlation test. Tables 4 and 5 show the results of the test by reporting the correlation coefficient and significance.

Spearman's rho (83)		EDiff	ADiff	CDiff	NDiff	ODiff
AD	Correlation Coeff.	-0.059	0.117	-0.144	0.124	0.081
	Sig. (2-tailed)	0.711	0.459	0.362	0.435	0.608
CC	Correlation Coeff.	-0.409	-0.056	-0.152	0.146	0.109
	Sig. (2-tailed)	**0.007****	0.725	0.337	0.355	0.491
CM	Correlation Coeff.	0.333	0.101	0.221	-0.218	-0.105
	Sig. (2-tailed)	**0.031***	0.524	0.160	0.166	0.509

* *Significant at the 0.05 level (2-tailed)*

** *Significant at alpha level corrected by Bonferroni-Holms method*

Table 4: Textual Design Representation: Correlation between the difference in the pairs' personalities and communication

Spearman's rho (113)		EDiff	ADiff	CDiff	NDiff	ODiff
AD	Correlation Coeff.	0.047	0.062	0.167	-0.026	-0.157
	Sig. (2-tailed)	0.729	0.647	0.214	0.848	0.243
CC	Correlation Coeff.	-0.053	-0.220	0.152	-0.165	0.046
	Sig. (2-tailed)	0.698	0.100	0.260	0.220	0.735
CM	Correlation Coeff.	0.018	0.173	-0.204	0.186	0.100
	Sig. (2-tailed)	0.893	0.198	0.129	0.166	0.459

Table 5: Graphical Design Representation: Correlation between the difference in the pairs' personalities and communication

By using a textual software design representation for design communication, we observe that there is a significant relationship between the difference in pairs' *extraversion* trait and the amount of *creative conflicts*. The correlation coefficient is -0.409 and it is significant at the alpha level corrected by Bonferroni-Holm due to analyzing multiple correlations. The *corrected-α* is α/n, where n is the number of performed correlations. This relationship is *negative* which means that as the difference in *extraversion* trait between explainers and receivers increases, the frequency of *creative conflict* discussions decreases. Hence, explainer-receiver pairs having similar *extraversion* traits (e.g., the explainer and receiver are both extroverts, introverts, or in between) generate more *creative conflict* than pairs who are farther apart from the same level of *extraversion* trait.

Moreover, we observe a significant relationship between the difference in pairs' *extraversion* trait and the effort of *conversation management*. The correlation coefficient is 0.333, and it is significant at the alpha level of 0.05. This relationship is *positive* which means that as the difference in *extraversion* trait between explainers and receivers increases, the effort of *conversation management* increases. Hence, explainer-receiver pairs who are closer to the same *extraversion* trait (e.g., the explainer and receiver are both extroverts, introverts, or in between) spend less effort on *conversation management* than the pairs who are farther apart from the same level of *extraversion* trait.

Finding 4
In the textual design group, we observe that as the difference in Extraversion trait between the stakeholders increases, the frequency of Creative Conflict discussions decreases and the effort of Conversation Management increases. The correlation is statistically significant.

By using a graphical software design representation during design communication, we observe that there are no significant relationships between differences in personality traits and quality of interpersonal communication.

5 IMPLICATIONS

The results of our study show that while some personality traits have a negligible impact on the quality of communication, others interact either directly or in a complex way with the nature of the design representation. In the latter case, our observations do not allow us to draw any conclusions that could directly be applied by practitioners and researchers, but rather call for a deeper investigation of these phenomena.

Moreover, the results show that a homogeneous or heterogeneous distribution of the majority of the considered personality traits does not influence the general trend of having a better communication quality (i.e., more *active discussion* and less *conversation management*) when using graphical design representations compared to using textual design representations. An exception is made in the case of *conscientiousness* for *conversation management*, where the results are not conclusive, as in the case of *openness* for *active discussion*.

Furthermore, the results show that using textual representations of software design is one main factor that plays a role in reducing design reasoning and argumentation, increasing the effort of *conversation management*, and hence adversely influencing communication quality between individuals at different levels of the *extraversion* trait. In contrast, adopting graphical representations of software design as a means for supporting communication does not impact the communication quality of stakeholders, especially when significant *extraversion* disparity appears.

The previous observations lead to the following Implication 1.

Implication 1
Introducing graphical representations of software design in the communication routine of software development teams (especially in teams with diverse personalities) has the potential to enhance the communication quality of the team and increase its productivity. This effect might however be mitigated by specific combinations of openness and conscientiousness disparity among team members.

According to the results of this study, using graphical representations of software design would be of benefit for team confidence and cohesion by enabling introverted collaborators to communicate more easily with colleagues at a stronger extroverted personality.

The results might also apply to team/project managers who sometimes may be reticent to compose teams with contrasting personalities. This would result in creating heterogeneous teams (as personality is tied to social and cultural backgrounds), which already prove to be more creative, a key aspect in Software Engineering [10].

Implication 2
Knowing that graphical representations of software design are or will be exploited for communication in software development teams, would, in the light of the findings, enable managers to consider more comfortably disparate or diverse personality profiles to work together.

We consider these research findings relatively preliminary and encourage replication. While waiting for new experiments and studies to confront the results of this study, we strongly recommend the adoption of graphical representations of software design as a means for supporting design communication in software development teams, since they provide a potential benefit for improving software quality and team productivity.

6 THREAT TO VALIDITY

This section discusses the threats to validity and different tactics we adopted to mitigate them.

6.1 Internal Validity

To mitigate the effect of limited understanding and explanation times, multiple pilot studies were performed prior to the experiment. These pilot studies helped to figure out how much time is required by the explainer to understand the design task, and how much time is needed to communicate or explain the design to the receiver. Based on the feedback from the pilot study's participants, the suitable times are identified as follows: Explainers have 20 minutes to understand the software design that they get, and 12 minutes to communicate the design or explain it to the receivers. Moreover, the way in which the participants of our experiment are randomly assigned to the textual and the graphical design representation groups might lead to potential confounding factors in case we have specific personality types prevalent in one of the two groups. To mitigate such a threat to the internal validity, the distribution of the personality traits in the textual and graphical design representation groups are analyzed. It is found that the distribution is the same across the two groups (see Table 2).

The structure and representation of the design in the textual and graphical artifacts might have influenced the quality of communication. The textual artifacts are structured by indentation, indexing, and grouping information. The MVC entities in the graphical artifacts are highlighted by color. To mitigate such an effect, a thorough work is conducted to making these artifacts providing the same amount of information. Moreover, in practice, it is common to color the graphical artifacts (or models) and structure the textual artifacts (or documents). Accordingly, this cannot be considered as significant threat over using unstructured textual or uncolored graphical artifacts.

In the textual design representation group, it is observed that some receivers were drawing informal graphical diagrams of the design while the explainers were explaining it to them. This might have caused an interaction between the independent variables. However, the number of these cases is small (2.5%) and the effect of this interaction can be considered as negligible.

6.2 External Validity

The participants of the experiment are software engineering students. These participants might not represent the entire population of software engineering practitioners and professionals. This might influence the generalizability of the results to other practitioners with different expertise, experiences, and backgrounds. However, it was found that students and professional software developers often perform similarly in small tasks [11]. Moreover, using students is considered as a good alternative when it is impractical and expensive to obtain appropriate samples of professionals [9]. Regardless of that, we call for replication to extend and generalize the results of the study to other software engineering practitioners.

Having used specific types and structures of textual and graphical artifacts in the experiment might have also influenced the generalizability of the results. However, this is a threat independent of the specific choice of the representation. Moreover, our results are considered as preliminary, and it is planned to study the effect of different types of textual and graphical artifacts on the quality of communication.

The conditions of the experiment such as the size, topic, and complexity of the experimental object (i.e., design case), size of teams, size of documents (i.e., textual and graphical design representations) used for investigation, and understanding and explaining time duration might be different in real-word scenarios. These conditions might limit the generalizability of the results. To mitigate this effect, the industrial, real-word use case of onboarding of newcomers by experienced developers at Ericsson, the "*Experience Engine*" [12] is imitated. Moreover, replications can help to address this threat to the external validity of this study.

6.3 Conclusion Validity

In the experiment, hypotheses testing is performed to find out whether independent variables have an effect on dependent variables. The implications that are stated earlier in this manuscript might have been affected by type I error which leads to incorrectly reject the null hypothesis (false positive), or by type II error which leads to incorrectly accept the null hypothesis (false negative). To mitigate this effect, the risk of detecting a non-real effect is reduced by considering a significance at the alpha level of 0.05. Furthermore, the sensitivity is studied by exploring the effect size and statistical power of the performed statistical tests.

7 CONCLUSION AND FUTURE WORK

This paper describes an empirical study conducted to explore the influence of software design model representation (textual vs. graphical) on interpersonal communication between stakeholders with diverse personality traits. To this end, We collected personality traits data and analyzed the software design communication of 196 software engineering students from four European universities. Our analysis reveals that:

(1) Graphical design representations are better than textual representations in promoting *active discussion* between stakeholders, given that there is no high difference in *openness* level between the stakeholders.
(2) Graphical design representations are better than textual representations in reducing the *conversation management* effort between stakeholders, given that there is no low difference in *conscientiousness* level between the stakeholders.
(3) By using textual design representations, we observe that as the difference in *extraversion* level between the stakeholders increases, the frequency of *creative conflict* discussions decreases and the effort of *conversation management* increases. The correlation is statistically significant.

As an implication, we encourage using software design diagrams (graphical representations of the software design model) in the communication routine of software design and development teams. This has the potential to enhance the communication of the team and increase its productivity, especially in teams with diverse personality traits.

As a future work, we plan to replicate this experiment to address certain threats to its validity, including potential effects of country and culture, coloring of text graphical representation, or different textual representations. This can be done e.g., by involving professionals or using different graphical and textual representations of software design models. Furthermore, we plan to explore how these findings could be applied for creating new techniques or approaches that would enhance the effectiveness of software design and engineering processes.

REFERENCES

[1] Ruth Baumgart, Markus Hummel, and Roland Holten. 2015. Personality Traits of Scrum Roles in Agile Software Development Teams - A Qualitative Analysis. In *ECIS 2015 Completed Research Paper*.
[2] Marianne Bradford, Sandra B. Richtermeyer, and Douglas F. Roberts. 2007. System Diagramming Techniques: An Analysis of Methods Used in Accounting Education and Practice. *Journal of Information Systems* 21, 1 (03 2007), 173–212. https://doi.org/10.2308/jis.2007.21.1.173 arXiv:https://meridian.allenpress.com/jis/article-pdf/21/1/173/1741773/jis_-2007_21_1_173.pdf
[3] Shirley Cruz, Fabio Q.B. da Silva, and Luiz Fernando Capretz. 2015. Forty years of research on personality in software engineering: A mapping study. *Computers in Human Behavior* 46 (2015), 94 – 113. https://doi.org/10.1016/j.chb.2014.12.008
[4] Boele De Raad. 2000. *The Big Five Personality Factors: The psycholexical approach to personality.* Hogrefe & Huber Publishers.
[5] Gergely Dévai, Gábor Ferenc Kovács, and Ádám An. 2014. Textual, Executable, Translatable UML.. In *OCL@ MoDELS*. Citeseer, 3–12.
[6] M Brent Donnellan, Frederick L Oswald, Brendan M Baird, and Richard E Lucas. 2006. The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment* 18, 2 (2006), 192.
[7] Paul D Ellis. 2010. *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.
[8] Robert Feldt, Lefteris Angelis, Richard Torkar, and Maria Samuelsson. 2010. Links between the personalities, views and attitudes of software engineers. *Information and Software Technology* 52, 6 (2010), 611 – 624. https://doi.org/10.1016/j.infsof.2010.01.001
[9] Robert Feldt, Thomas Zimmermann, Gunnar R Bergersen, Davide Falessi, Andreas Jedlitschka, Natalia Juristo, Jürgen Münch, Markku Oivo, Per Runeson, Martin Shepperd, et al. 2018. Four commentaries on the use of students and professionals in empirical software engineering experiments. *Empirical Software Engineering* 23, 6 (2018), 3801–3820.

[10] Narasimhaiah Gorla and Yan Wah Lam. 2004. Who should work with whom? Building effective software project teams. *Commun. ACM* 47, 6 (2004), 79–82.

[11] Martin Höst, Björn Regnell, and Claes Wohlin. 2000. Using students as subjects—a comparative study of students and professionals in lead-time impact assessment. *Empirical Software Engineering* 5, 3 (2000), 201–214.

[12] Conny Johansson, Patrik Hall, and Michael Coquard. 1999. "Talk to paula and peter—They are experienced" the experience engine in a nutshell. In *International Conference on Software Engineering and Knowledge Engineering*. Springer, 171–185.

[13] David W Johnson and Roger T Johnson. 1987. *Learning together and alone: Cooperative, competitive, and individualistic learning*. Prentice-Hall, Inc.

[14] Rodi Jolak, Truong Ho-Quang, Michel RV Chaudron, and Ramon RH Schiffelers. 2018. Model-based software engineering: A multiple-case study on challenges and development efforts. In *Proceedings of the 21th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems*. 213–223.

[15] Rodi Jolak, Maxime Savary-Leblanc, Manuela Dalibor, Juraj Vincur, Regina Hebig, Xavier Le Pallec, Michel R.V. Chaudron, Sébastien Gérard, Ivan Polasek, and Andreas Wortmann. 2022. Experimental Material: The Influence of Software Design Representation on the Design Communication of Teams with Diverse Personalities. https://zenodo.org/record/4707552.

[16] Rodi Jolak, Maxime Savary-Leblanc, Manuela Dalibor, Andreas Wortmann, Regina Hebig, Juraj Vincur, Ivan Polasek, Xavier Le Pallec, Sébastien Gérard, and Michel R.V. Chaudron. 2020. Software engineering whispers: The effect of textual vs. graphical software design descriptions on software design communication. *Empirical Software Engineering* 25, 6 (2020), 4427–4471. https://doi.org/10.1007/s10664-020-09835-6

[17] Márta Juhász. 2010. Influence of personality on Teamwork behaviour and communication. *Periodica Polytechnica Social and Management Sciences* 18, 2 (2010), 61–74.

[18] Andrea Seaton Kelton, Robin R. Pennington, and Brad M. Tuttle. 2010. The Effects of Information Presentation Format on Judgment and Decision Making: A Review of the Information Systems Research. *Journal of Information Systems* 24, 2 (2010), 79–105. https://doi.org/10.2308/jis.2010.24.2.79

[19] Steve W.J. Kozlowski and Daniel R. Ilgen. 2006. Enhancing the Effectiveness of Work Groups and Teams. *Psychological Science in the Public Interest* 7, 3 (2006), 77–124. https://doi.org/10.1111/j.1529-1006.2006.00030.x

[20] Robert R McCrae and Oliver P John. 1992. An introduction to the five-factor model and its applications. *Journal of personality* 60, 2 (1992), 175–215.

[21] James McCroskey, Alan Heisel, and Virginia Richmond. 2001. Eysenck's BIG THREE and communication traits: three correlational studies. *Communication Monographs* 68, 4 (2001), 360–366. https://doi.org/10.1080/03637750128068

[22] Margaret M McManus and Robert M Aiken. 1995. Monitoring computer-based collaborative problem solving. *Journal of Interactive Learning Research* 6, 4 (1995), 307.

[23] Daniel L Moody. 2010. The" physics" of notations: a scientific approach to designing visual notations in software engineering. In *2010 ACM/IEEE 32nd International Conference on Software Engineering*, Vol. 2. IEEE, 485–486.

[24] Miranda AG Peeters, Harrie FJM Van Tuijl, Christel G Rutte, and Isabelle MMJ Reymen. 2006. Personality and team performance: a meta-analysis. *European Journal of Personality: Published for the European Association of Personality Psychology* 20, 5 (2006), 377–396.

[25] Jinny Rhee, David Parent, and Anuradha Basu. 2013. The influence of personality and ability on undergraduate teamwork and team performance. *SpringerPlus* 2, 1 (2013), 16. https://doi.org/10.1186/2193-1801-2-16

[26] Hamzah Ritchi, Mieke Jans, Jan Mendling, and Hajo Reijers. 2020. The influence of business process representation on performance of different task types. *Journal of Information Systems* 34, 1 (2 2020), 167–194. https://doi.org/10.2308/isys-52385

[27] Bernhard Rumpe. 2016. *Modeling with UML: Language, Concepts, Methods*. Springer International.

[28] Janice Singer and Norman G. Vinson. 2002. Ethical issues in empirical studies of software engineering. *IEEE Transactions on Software Engineering* 28, 12 (2002), 1171–1180.

[29] Amy Soller. 2001. Supporting social interaction in an intelligent collaborative learning system. *International Journal of Artificial Intelligence in Education (IJAIED)* 12 (2001), 40–62.

[30] SysMLv2. accessed April 2021. OMG System Modeling Language V2 Release. https://github.com/Systems-Modeling/SysML-v2-Release.

[31] Walter F Tichy. 2000. Hints for reviewing empirical work in software engineering. *Empirical Software Engineering* 5, 4 (2000), 309–312.

[32] Barbara Tversky. 2014. Visualizing thought. In *Handbook of human centric visualization*. Springer, 3–40.

[33] Barbara Tversky. 2018. Multiple Models. In the Mind and in the World. *Historical Social Research/Historische Sozialforschung. Supplement* 31 (2018), 59–65.

[34] Murat Yilmaz, Rory V. O'Connor, Ricardo Colomo-Palacios, and Paul Clarke. 2017. An examination of personality traits and how they impact on software development teams. *Information and Software Technology* 86 (2017), 101 – 122. https://doi.org/10.1016/j.infsof.2017.01.005

[35] Wei Zhang, Sunny Li Sun, Yuan Jiang, and Wenyao Zhang. 2019. Openness to Experience and Team Creativity: Effects of Knowledge Sharing and Transformational Leadership. *Creativity Research Journal* 31, 1 (2019), 62–73. https://doi.org/10.1080/10400419.2019.1577649 arXiv:https://doi.org/10.1080/10400419.2019.1577649