

Poster: The Effect of Noise on Requirements Comprehension

Simone Romano
University of Basilicata, Italy
simone.romano@unibas.it

Giuseppe Scanniello
University of Basilicata, Italy
giuseppe.scanniello@unibas.it

Davide Fucci
University of Hamburg, Germany
fucci@informatik.uni-hamburg.de

Natalia Juristo
Universidad Poli. de Madrid, Spain
natalia@fi.upm.es

Burak Turhan
Brunel University London, UK
burak.turhan@brunel.ac.uk

ABSTRACT

We conducted a controlled experiment with 55 final-year undergraduate students in Computer Science. We asked them to comprehend functional requirements exposing them or not to noise. We did not observe any effect of noise on requirements comprehension.

CCS CONCEPTS

• **Software and its engineering** → **Software creation and management**;

KEYWORDS

Comprehension, noise, experiment, functional requirement

ACM Reference Format:

Simone Romano, Giuseppe Scanniello, Davide Fucci, Natalia Juristo, and Burak Turhan. 2018. Poster: The Effect of Noise on Requirements Comprehension. In *ICSE '18 Companion: 40th International Conference on Software Engineering Companion*, May 27-June 3, 2018, Gothenburg, Sweden. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3183440.3194984>

1 INTRODUCTION

Noise, defined as an unwanted sound, is one of the commonest factors that could affect people's performance in their daily work activities. Nowadays, workspaces tend to have less privacy, with less dedicated space, which leads to noisy environment. The obvious reason for this trend is the cost. A "penny" saved on the workspace is a "penny" earned on the bottom line, or so the logic goes [2]. The savings of a cost-reduced workplace are attractive, but these savings need to be compared to the risk of performance reduction in daily working activities/tasks. Software companies that provide a noisy workplace are comforted by the belief that this factor does not matter [2], but noise exerts its specific influences on various forms of cognitive responses [5]. After all, software engineers are knowledge workers—they need to have their brain in gear to do their work—and thus their performance would be sensitive to a noisy workplace while accomplishing working activities/tasks.

We present a controlled experiment, whose overarching goal was to assess if noise influences the comprehension of functional requirements. A summary of this experiment is shown in Table 1.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICSE '18 Companion, May 27-June 3, 2018, Gothenburg, Sweden

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5663-3/18/05.

<https://doi.org/10.1145/3183440.3194984>

2 STUDY DESIGN

In our experiment, we defined and investigated the following research question: *Does noise worsen the comprehension of functional requirements?*

2.1 Experimental Units and Material

The participants were final-year undergraduate students in Computer Science at University of Basilicata in Italy. The participants attended the Software Engineering course in which we conducted the experiment as optional laboratory exercise. They had basic knowledge of requirements engineering. The participation was on voluntary basis (*i.e.*, in no case we obliged students to participate).

The systems used in our experiment were: **M-Shop** and **Theater**. M-Shop is a system for managing the sales in a music shop, while Theater is a system for the reservation of tickets in a theater. For each of these systems, one functional requirement with the corresponding models (*i.e.*, functional model, analysis object model, and dynamic model) was selected from its requirements analysis specification document. These experimental objects were previously used by Abrahão *et al.* [1] in their family of experiments.

2.2 Tasks

The participants had to perform the following tasks:

Comprehension task 1—We provided each participant with the models associated with the functional requirement "Search Album by Singer" of M-Shop. To evaluate the comprehension of such a requirement, we asked the participants to fill out a comprehension questionnaire consisting of 11 closed-ended questions. Each question admitted one or more right answers.

Comprehension task 2—We asked the participants to perform the same task as the previous one, but the experimental object was Theater. In particular, we gave each participant the models of the requirement "Buy Theater Ticket."

2.3 Hypotheses, Parameters, and Variables

The participants who had to perform the tasks in the presence of NOISE comprised the treatment group; while those who worked

Table 1: Summary of the experiment.

| | |
|----------------------|--|
| Schedule | 11:30 on 2016/12/12 |
| Kind of SE task | Comprehension of functional requirements |
| Task duration | 30 minutes |
| Experimental objects | M-Shop and Theater |
| Participants | 55 final-year undergraduate students |
| Group1/Group2 size | 28/27 |
| Experiment design | AB/BA crossover design |

on the tasks in NORMAL conditions (*i.e.*, they were not exposed to noise) comprised the control group.

To quantify the comprehension of functional requirements, we evaluated the answers of each participant to a comprehension questionnaire by using two strategies. The first one was information retrieval-based and consisted in computing precision (P) and recall (R) of the answers given. To get a single measure that represents a trade-off between correctness (P) and completeness (R), we compute the balanced F-measure as follows: $F_1 = \frac{2*P*R}{P+R}$. This metric assumes values in the interval $[0, 1]$, where 1 is the best value. The second strategy was inspired to that by Kamsties *et al.* [3]. In particular, we quantified comprehension by means of: $Avg = \frac{\sum_{i=1}^n count_i}{n}$. The $count_i$ assumes 1 as the value if the answer to the question i corresponds to the oracle for that question, while n is the number of questions. Avg assumes values in the interval $[0,1]$, where 1 is the best value.

According to our research question, we formulated the null hypothesis H_n as follows: *noise does not significantly affect the comprehension of functional requirements.*

2.4 Experiment Design and Procedure

The used design was an AB/BA crossover [6]. This kind of design has two treatments (*i.e.*, A and B). An AB/BA crossover design has two periods (*i.e.*, the times at which each treatment is applied). Participants are split into Group1 and Group2 and administered with every treatment only once. The groups represent the sequences, *i.e.*, the order in which treatments have administered to participants. In our experiment, A and B are NOISE and NORMAL, respectively. We randomly assigned the participants to Group1 (the sequence AB) and Group2 (sequence BA). We added a 30-minute wash-out period between the two periods. A wash-out period should leave sufficient time for the effect of a treatment to recede completely, and thus possibly neutralize carryover effect [6].

The participants exposed to noise accomplished the tasks with a noise exposure level¹ (L_{EX}) equal to 82 dB. We established this value on the basis of the Directive 2003/10/EC² of the European Parliament. A value of L_{EX} equal to 82 dB does not require the use of individual hearing protectors and it is not considered harmful to health. We used speech noise because it is the major type of practical distractive noise [5] and it is common in workplaces with open office plans [4]. As for the participants working in normal conditions, the L_{EX} value was equal to 42 dB, namely a common noise level for a quite office workplace [4].

2.5 Analysis Procedure

For each dependent variable, we perform a pre-test to check the presence of a carryover effect by comparing the within-participant sum (in the periods) as Wellek and Blettner [7] suggested.

If carryover is not statistically significant, we perform the following steps: (i) we compute descriptive statistics and (ii) we test whether the effect of noise is not statistically significant (*i.e.*, H_n) by comparing the within-participant differences (in the periods) for each dependent variable [7].

¹It is the time-weighted average of the noise exposure levels.

²eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02003L0010-20081211

Table 2: Descriptive statistics and H_n results.

| Variable | p-value | Statistic | Condition | |
|----------|---------|-----------|-----------|--------|
| | | | NORMAL | NOISE |
| F_1 | 0.7309 | Median | 0.6923 | 0.7143 |
| | | Mean | 0.6781 | 0.6833 |
| | | Std | 0.1268 | 0.1398 |
| Avg | 0.3119 | Median | 0.5455 | 0.5455 |
| | | Mean | 0.5025 | 0.5256 |
| | | Std | 0.1296 | 0.1552 |

To test both carryover effect and H_n , we run an unpaired two-sided t-test if data are normally distributed, a two-sided Mann-Whitney U test otherwise.

3 RESULTS

The pre-test results indicate that the carryover effect is not statistically significant for both F_1 (0.5179) and Avg (0.4312). We applied unpaired two-sided t-tests since data were normally distributed (*i.e.*, the Shapiro test returned p-values greater than 0.05).

In Table 2, we summarize descriptive statistics on the dependent variables. They indicate that there is not a huge difference between the participants administered with the NORMAL and NOISE treatments with respect to F_1 (0.6781 vs 0.6833 on average) and Avg (0.5025 vs 0.5256 on average).

To test H_n , we used an unpaired two-sided t-test for F_1 (data were normally distributed since the Shapiro test returned a p-value greater than 0.05), while the Mann-Whitney U test for Avg (data were not normally distributed since the Shapiro test returned a p-value equals to 0.0167 for Group2). The results of these tests (see p-values shown in Table 2) did not allow us to reject H_n for both the dependent variables.

4 REMARKS

The results suggest that noise does not significantly affect the comprehension of functional requirements. Thus, we cannot positively answer our research question. On the basis of the Szalma and Hancock [5] results, we speculate that the lack of a statistical significant difference might be due to the kind of Software Engineering task. That is, comprehending functional requirements might be little resource demanding, so neutralizing the possible negative effect of noise. Further empirical work is needed on this respect and our study poses the basis for future work.

REFERENCES

- [1] S. Abrahão, C. Gravino, E. Insfran, G. Scanniello, and G. Tortora. 2013. Assessing the Effectiveness of Sequence Diagrams in the Comprehension of Functional Requirements: Results from a Family of Five Experiments. *IEEE Trans. on Softw. Eng.* 39, 3 (2013), 327–342.
- [2] T. DeMarco and T. Lister. 2013. *Peopleware: Productive Projects and Teams* (3rd ed.). Addison-Wesley Professional.
- [3] E. Kamsties, A. von Knechten, and R. Reussner. 2003. A controlled experiment to evaluate how styles affect the understandability of requirements specifications. *Inf. Softw. Technol.* 45, 14 (2003), 955–965.
- [4] L. E. Maxwell. 2015. Noise in the office workplace. *Facility Planning and Management Notes* 1, 11 (2015).
- [5] J. L. Szalma and P. A. Hancock. 2011. Noise effects on human performance: a meta-analytic synthesis. *Psychological bulletin* 137, 4 (2011), 682–707.
- [6] S. Vegas, C. Apa, and N. Juristo. 2016. Crossover Designs in Software Engineering Experiments: Benefits and Perils. *IEEE Trans. on Softw. Eng.* 42, 2 (2016), 120–135.
- [7] S. Wellek and M. Blettner. 2012. On the Proper Use of the Crossover Design in Clinical Trials. *Dtsch Arztebl International* 109, 15 (2012), 276–281.