



## ESPLAT2023 Conference submission form

### Title of Paper/Poster:

#### Author(s) (please indicate presenting author with \*)

\*Rainer Scheuchenpflug, Institut für Psychologie, Universität Würzburg,  
rainer.scheuchenpflug@uni-wuerzburg.de

Alexander Hörnlein, Rechenzentrum, Universität Würzburg, casetrain@uni-wuerzburg.de

#### Title:

How do they use it? Logfile analysis of learning material utilization during exam preparation

#### Abstract (100-250 words):

Students in my statistics courses can use 25 exams from previous semesters to prepare for their exam (which contains new questions but covers the same topics). Since the material is provided in an electronic format and user interactions are logged, it is possible to analyze how learners use this (subjectively very) relevant material which is available early in the semester for preparation. Logfile data from winter term 2021 (for Statistics 1) and summer term 2022 (Statistics 2) show that learners in general tend to use newer exams more often than older ones, even if older exams do provide similar training opportunities. Also, not surprisingly, learners tend to utilize material rather shortly before the exam, and a large percentage of users does not complete individual exams.

Interviews with students showed that presenting exam problems in the context of the original exam is hindering free interaction and self-selected learning. Therefore, a project starting in January 2023 will separate the existing exams into different subproblems and will add metadata like topic, difficulty, type of test, type of question etc., and integrate the collection of subtasks into the electronic system. The system will give students the opportunity to specifically select exam problems fitting to their current understanding and will provide individual feedback and inform about topics where they need further training. Hopefully, this change in the format learning material is presented will lead to a change of student learning and exercising behavior.

#### Extended Summary (max. 800 words including references; please adhere to APA-guidelines):

##### 1. Objectives or purposes.

This study examines how students of statistics in Psychology use course material in exam preparation by analyzing logfiles from an electronic system (CaseTrain) providing this material. A closer look on traces of students' use of online material can be used to identify maladjusted learning strategies and inform the design and future presentation of learning material.

##### 2. Perspective(s) or theoretical framework.

The analysis stands in the context of learning analytics, a field of research that has increased in recent years (Aldowah et al., 2019; Lang et al., 2022; Romero & Ventura, 2019).



### 3. Methods, techniques, or modes of inquiry.

Logfile data from two recent cohorts of students were used: Winter term 2021 for preparatory activities preceding the exam in Statistics 1, and data from summer term 2022 for Statistics 2. Statistics 1+2 are required courses for students of Psychology and Human-Computer-Systems, typically visited in the first and second semester.

### 4. Data sources, evidence, objects, or materials.

Students' interaction with CaseTrain are recorded in the system on a microlevel ("Person A selects answer 3 in question 10 of case Z on date Y"), but are presented to the instructor in different aggregated forms, which are used here to analyze resource utilization over time, search behavior, navigation behavior and use of feedback. Subject records are grouped according to performance in the final exam to test the possibility to identify beneficial or maleficial learning strategies.

### 5. Results and/or substantiated conclusions or warrants for arguments/point of view.

Students in both cohorts start to use the electronic learning material rather late in the semester. This can partially be expected, because as more and more content is covered in lecture, more exam problems become relevant and solvable for learners. But still, students start exam preparation activities rather shortly before the exam. One reason for this behavior could be the organization of the material: Exam questions are presented in the context of complete exams, not separately, and searching for problems which are solvable at a certain time in the course might be too difficult for students, or approaching a complete exam might be intimidating.

A second point is that exam material seems to age: Students use newer exams quite intensively, with an exponential drop in utilization with increasing age of the material. Since exam problems are not repeated, but every exam is created with new problems each semester, this behavior is not warranted by strategic decisions or by higher similarities of newer exams to the upcoming exam, but limits the material's sustainability. A simple reason for this behavior might be the order of presentation of past exams in the course management system, where the "newest" exams come first.

A third observation is that many (~50 %) exams are not completed by learners, especially shortly before the real exam. Interviews with students and teaching assistants support the assumption that students are searching for specific problems in different exams and close the file when they solve this specific problem (and process the automatic feedback and explanations).

### 6. Scientific or scholarly significance of the study or work.

While providing old exams as training opportunity for the upcoming exam is seen as relevant by students and this material is used rather intensively compared to additional available exercises created especially for training, the presentation of exam problems in the form of the original exams (as a collection of problems to different topics) seems to hinder individual learning processes. A new project starting in January 2023 will separate the existing exams into different subproblems and will enhance the material with metadata like problem difficulty in the original exam, and keywords like topic, type of test, type of question etc. The software system will be enhanced with a filtering interface, so that users can select problems according to topic, difficulty, or according to prior performance.

An obvious problem is the structure of metadata, which – in the minds of novice users - will probably differ from the structure used in textbooks or by experts (Chi, Feltovich & Glaser, 1981). The project will use questionnaires as well as logfile data in order to iteratively improve the system of terms used in metadata coding.

References:



# ESPLAT

**European Society of Psychology Learning and Teaching**

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13-49. <https://doi.org/https://doi.org/10.1016/j.tele.2019.01.007>

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, 5(2), 121-152.

[https://doi.org/https://doi.org/10.1207/s15516709cog0502\\_2](https://doi.org/https://doi.org/10.1207/s15516709cog0502_2)

Lang, C., Siemens, G., Wise, A.F., Gašević, D., & Merceron, E. (Eds.). (2022). *The Handbook of Learning Analytics* (2<sup>nd</sup> ed.). SoLAR. <https://doi.org/10.18608/hla22>

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1355.

<https://doi.org/https://doi.org/10.1002/widm.1355>