Correlation between students' background and the knowledge on Conceptual Database Modelling

Lili Nemec Zlatolas[1][0000-0002-2086-9825], Aida Kamišalić[1][0000-0002-8574-8506] and Muhamed Turkanović[1][0000-0002-5079-5468]

[1] University of Maribor, Faculty of Electrical Engineering and Computer Science, Koroška cesta 46, 2000 Maribor, Slovenia
{lili.nemeczlatolas, aida.kamisalic, muhamed.turkanovic}@um.si

Abstract. Students taking a Databases Course have different pre-knowledge of Conceptual Database Modelling. At the beginning of the course, they were asked to self-evaluate their knowledge of database design concepts. We have compared the results of self-evaluation of students with Technical and General High School education, as well as the results they achieved at the final exam. All together 132 students have collaborated in both parts of the survey. The results show that the students with a technical background have a better pre-knowledge of database design constructs, but, on the other hand, received lower results in the exam where the knowledge on Database Modelling was tested.

Keywords: Database Modelling, comparison, database design learning.

1 Introduction

Conceptual Database Modelling is an essential step in creating a good database. Teaching the students how to create a good conceptual model is also a challenge for all teachers, since there is no straight line for the perfect solution of any specific scenario [1, 2]. Studies show that differences reside between basic and expert database (model) designers, whereby many conceptual mistakes are produced by the former, while modelling a database [3-5]. Becoming a good database designer is a cumbersome process, requiring a lot of studying and practicing.

Students with different secondary education level backgrounds are joining Universities to study Computer Science or similar ICT related studies. The Database course usually serves as a fundamental course in such studies, due to the importance of understanding database fundamentals, including modelling, as soon as possible. However, many students come to the University with some pre-knowledge of databases and tend to overestimate their actual knowledge. The studies have shown that pre-knowledge can have an effect on students, so that they believe already having enough knowledge on the topic and not needing to learn any more [6, 7]. This might lead to the effect that such students do not achieve a level, which is required or desired by University standards.

In our previous studies, we examined the effectiveness of learning database fundamentals, depending on the notation used for conceptual design [8]. A multi-level experimental study was set up. Considering the influencing factors, students' achievements were examined throughout the learning process. Additionally, the influence of notation used for conceptual design on student knowledge perception [7] as well as the influence of students' educational background on their learning outcomes [9] were examined.

In this work, we present another aspect of the aforementioned study, where we analysed if students with different secondary education level backgrounds, thus different pre-knowledge of Database Modelling and their self-perception of this knowledge, influences their results. The analysis was performed with active students, after they had taken the course lessons and completed the practical laboratory work.

The structure of the paper is as follows. Subsection 1.1 presents the research questions set up for this study. In Section 2, information on research methods is provided. The main contribution of the paper is presented in Section 3 where results and discussion are detailed. Finally, the conclusions are presented in Section 4.

1.1 Research questions

We have set up the following two research questions:

RQ1: How does previous education affect the knowledge on conceptual design concepts`, between students with technical or general secondary education backgrounds?
RQ2: How does the pre-knowledge of conceptual design concepts affect the students' final grade gained for Database Modelling?

2 Research methods

For the purposes of the study we used a paper format questionnaire that was designed to answer the research questions. Survey questions were designed based on existing literature and discussion with fellow University members.

2.1 Data collection and participants

At the beginning of the semester, students attending the Databases Course were asked to fill in a self-evaluation questionnaire on the topic of their knowledge on conceptual database design concepts. After the course was finished and they had attended lectures on Database Modelling within the course, we tested their knowledge on a conceptual database design. The sample of demographics is presented in Table 1. All together 132 students collaborated in both parts of the survey – the self-evaluation questionnaire and the test after the course was finished. Since students who are taking the Databases Course are signed into the Study Programmes at the Faculty of Electrical Engineering and Computer Science, there are more male than female students taking this course.

Most of the students are in their first or second year of studies, so the average age concurs with this. The previous secondary education (High School) of students was either general or technical.

Table 1. Sample of demographics (n=132).

Variable	Sample results
Gender	Male 84.10% Female 15.90%
Age mean	M 19.86 SD 1.23
Study Programme	Computer Science and Information Technologies 62.1% Informatics and Technologies of Communication 37.9%
Previous education	General 55.30% Technical 44.70%

2.2 Measures

Measurement items were tested with a 5-point Likert-scale ranging from 1 to 5. The measurement items are presented in Table 2. There were altogether 3 demographical questions, 1 question on previous education and 5 questions regarding the Database Modelling concepts. The grade for Database Modelling was awarded by a teacher in the course after the students had already taken the lectures. To connect the pre-knowledge of each student and the grade for Database Modelling, we used a unique ID number that a student had to enter when filling in the self-evaluation questionnaire on his/her knowledge and on the exam at the end of the course.

Table 2. Measurement of variables.

Questions	Answers
Previous education	General education / Technical education
Perception of knowledge: • Entity (E) • Relationship (R) • Attribute (A) • Primary Key (PK) • Foreign Key (FK)	Possible answers: • I am not familiar with the term (1) • I am familiar with the term, but not with the meaning (2) • Undefined (3) • I am familiar with the meaning but I do not know how to use it (4) • I am familiar with the meaning and I know how to use it (5)
Grade for database modelling	0-100 %

3 Data analysis and results

A data analysis was performed with the use of IBM SPSS Statistics 24.0. First, we did a comparison of means for the pre-knowledge between general and technical students,

which is presented in Table 3. To test the significance of the results, we used the Mann-Whitney U test presented in Table 4 [10]. The students who finished a general secondary education self-evaluated their knowledge of database design concepts lower than the students with a technical secondary education level.

Table 3. Comparing means between groups with different high school education on Database Modelling knowledge

	Previous education	N	Mean Rank	Sum of Ranks
Entity	General	73	61.94	4521.50
	Technical	59	72.14	4256.50
Relationship	General	73	54.06	3946.50
	Technical	59	81.89	4831.50
Attribute	General	73	54.82	4002.00
	Technical	59	80.95	4776.00
Primary key	General	73	60.38	4407.50
	Technical	59	74.08	4370.50
Foreign key	General	73	51.36	3749.00
	Technical	59	85.24	5029.00
Mean for all database properties	General	73	53.84	3930.00
	Technical	59	82.17	4848.00

The Mann-Whitney U test in Table 4 shows that there are significant differences between the two groups with different higher education for almost all the database design concepts, except for the Entity.

Table 4. Mann-Whitney U test using grouping variable of previous education

	E	R	A	PK	FK	Mean of database properties
Mann-Whitney U	1820.5	1245.5	1301.0	1706.5	1048.0	1229.0
Asymp. Sig. (2-tailed)	.11	.00	.00	.03	.00	.00

Next, we divided the respondents into groups based on their self-evaluated pre-knowledge on Database Modelling. We calculated the mean for each participant for perception of knowledge of Entity, Relationship, Attribute, Foreign and Primary Key. The mean was between 1 and 5. Then, we divided the participants into 4 similarly sized groups. The first, with no pre-knowledge, rated their pre-knowledge with 1-2.4. The second group, with some pre-knowledge, rated it with 2.6-3.4. The third group, with good pre-knowledge, rated it with 3.6-4.8. The last group rated their knowledge high with a mean of 5. As can be observed in Figure 1, the students with general pre-education received higher results in the exam at the end of the lectures of the Databases

Course. The lowest grade that the students with high pre-knowledge of Database Modelling received was 28%, which is actually the lowest grade among all students. One possible explanation of this outcome could be that some students might think that they already know a lot and do not pay much attention to the Course itself.

In Table 5, the mean results of the grade for Database Modelling is presented for two groups of students with different secondary education background and different self-evaluation categories on their knowledge. In the majority of cases, students with general secondary education background wrote the exam for Database Modelling better than the ones with a technical background. Some of the students with general background and very high pre-knowledge on database design concepts achieved a bit lover results at the exam than the students with a technical background. If we combine all students together, then the results show that the more pre-knowledge the students had, the higher the grade they received for Database Modelling, but, further analysis shows that the students with general background and high pre-knowledge were probably too confident with their knowledge, since they received worse results than the ones with good pre-knowledge. On the other hand, the students with a technical background and some or good pre-knowledge, were also a bit too confident in their knowledge of database design, and received worse exam results for Database Modelling than the rest of the students with no or high pre-knowledge.

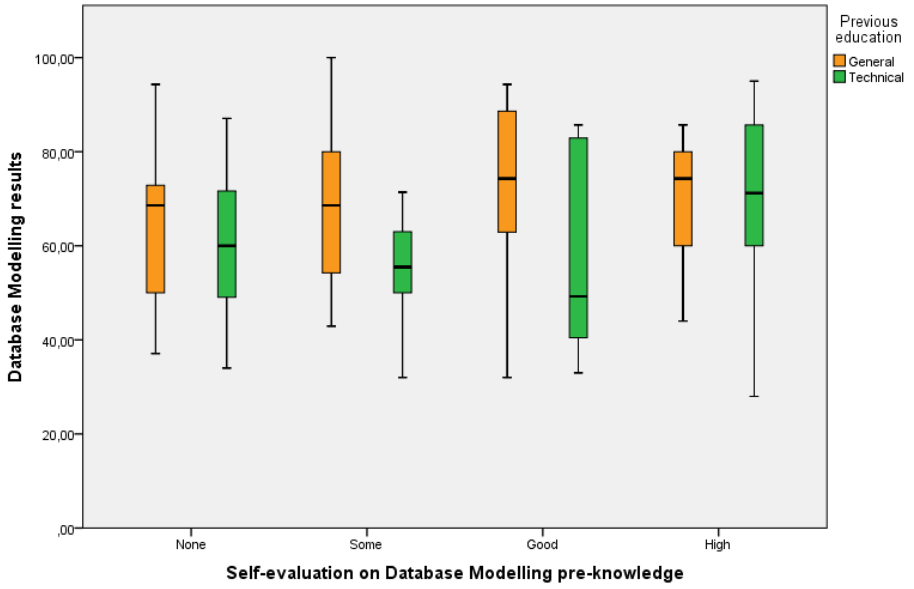

Fig. 1. Comparison of test results of Database Modelling for groups of students with Technical or General High School education and different self-evaluation categories.

Table 5. The results of the grade for the Database Modelling results for different groups.

	All students			General background			Technical background		
	Mean	**N**	**SD**	**Mean**	**N**	**SD**	**Mean**	**N**	**SD**
No pre-knowledge	63,51	30	15,54	64,47	23	14,84	60,36	7	18,55
Some pre-knowledge	63,81	34	17,74	67,80	24	18,13	54,23	10	13,05
Good pre-knowledge	66,36	33	20,17	75,00	17	16,02	57,18	16	20,48
High pre-knowledge	71,15	35	16,20	69,01	9	14,63	71,88	26	16,92
Total	66,32	132	17,60	68,58	73	16,38	63,54	59	18,77

3.1 Answering the research questions

The students with technical backgrounds reported higher pre-knowledge of conceptual database design concepts in comparison to students with a general high education background. The students with technical background reported to have higher knowledge of what Relationship, Attribute, Primary Key and Foreign Key mean in conceptual database design. There are no statistically significant differences for Entity as a concept of conceptual database design among the students with technical or general education backgrounds.

The students with technical backgrounds self-reported that they have better knowledge of the meaning or use of Entity, Relationship, Attribute, Primary and Foreign Keys, but they received lower exam results where conceptual database design was tested than the students with less pre-knowledge of conceptual database design.

4 Conclusion

At the Faculty of Electrical Engineering and Computer Science at the University of Maribor, students are taking a Databases Course in the first or second year of their studies. We asked the students at the beginning of the Course to fill in a self-evaluation questionnaire about their knowledge of Database Modelling. After they had attended lectures, they took an exam to check their Database Modelling knowledge. We then compared their self-evaluated pre-knowledge with their exam results obtained at the end of the Databases Course. We found out that the students who visited High Schools with general education self-evaluated their pre-knowledge much lower than the students who visited Technical High Schools before entering to University. This should not be a surprise, because the students from Technical High Schools have already attended some lectures on databases within their secondary education. On the other hand, we would also expect that the students with technical backgrounds would get higher grades for Database Modelling after attending the lessons of the Databases Course at

the University. However, we found out that the students with general backgrounds actually achieved higher grades in the Database Modelling part of the Course than the ones with technical backgrounds. There could be a number of reasons for this. Our limitation was the small sample, which cannot be generalized to all the population. Another point is that the students with technical backgrounds could have had a different pre-education on conceptual model language, and might mix some basic concepts of Conceptual Modelling.

Acknowledgements

The authors acknowledge the financial support from the Slovenian Research Agency (Research Core funding No. P2-0057) and from the European Union's Horizon 2020 Research and Innovation program under the Cybersecurity CONCORDIA project (GA No. 830927). We would also like to thank the participants of this research project.

References

1. Embley, D.W., Thalheim, B.: Handbook of conceptual modeling: theory, practice, and research challenges. Springer (2012).
2. Mylopoulos, J.: Conceptual modelling and Telos. Conceptual modelling, databases, and CASE: An integrated view of information system development 49-68 (1992).
3. Batra, D., Davis, J.G.: Conceptual data modelling in database design: similarities and differences between expert and novice designers. International Journal of Man-Machine Studies 37, 83-101 (1992).
4. Batra, D., Antony, S.R.: Novice errors in conceptual database design. European Journal of Information Systems 3, 57-69 (1994).
5. Engels, G., Gogolla, M., Hohenstein, U., Hülsmann, K., Löhr-Richter, P., Saake, G., Ehrich, H.-D.: Conceptual modelling of database applications using an extended ER model. Data & Knowledge Engineering 9, 157-204 (1992).
6. Sewell, A.: Constructivism and student misconceptions: Why every teacher needs to know about them. Australian Science Teachers Journal 48, 24 (2002).
7. Kamišalić, A., Turkanović, M., Heričko, M., Welzer, T.: Knowledge Perception influenced by Notation Used for Conceptual Database Design. In: Heričko, M. (ed.) 21st International Multiconference Information Society - IS 2018, pp. 19-22. Institut "Jožef Stefan", Ljubljana, Slovenia (2018).
8. Kamišalić, A., Heričko, M., Welzer, T., Turkanović, M.: Experimental Study on the Effectiveness of a Teaching Approach Using Barker or Bachman Notation for Conceptual Database Design. Computer Science & Information Systems 15, (2018).
9. Kamišalić, A., Turkanović, M., Welzer, T., Heričko, M.: Influence of Educational Background on the Effectiveness of Learning Database Fundamentals. In: Heričko, M. (ed.) The 29th International Conference on Information Modelling and Knowledge Bases, EJC 2019, Lappeenranta, Finland (2019).
10. Ruxton, G.D.: The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behavioral Ecology 17, 688-690 (2006).