

Introducing Normal Forms to Students: A Comparison Between Theory-First and Project-First Educational Approaches

DAKOTA C. COOKENMASTER* and JACOB A. BAHN*, Southern Adventist University, USA
GERMÁN H. ALFÉREZ, Southern Adventist University, USA

Educating the future generation of computer scientists and engineers often proves to be challenging, and how the content is introduced plays a large role in how well students will learn. One of the primary challenges that instructors face is regarding the introduction of important theory to students, both to show its essential nature to the field as well as its practicality. This paper analyzes two pedagogical methods for the instruction of normal forms in database management systems, a mandatory topic in any database course. The first of these methods is a theory-based approach that relies on written works and practices (i.e., theory) to introduce the concept. The second of these focuses on a project-based approach (i.e., practice) which aligns with the normal form as students implement a database schema. Through a small study, it was determined that most students have a strong predisposition to theory-first education, though students seemed to prefer the practice-based approach more than the theory-first approach. This paper compares the two methodologies, given this insight, and advises the use of an appropriate method for future educators.

CCS Concepts: • **Information systems** → **Database design and models**; **Deduplication**; • **Theory of computation** → **Incomplete, inconsistent, and uncertain databases**.

Additional Key Words and Phrases: pedagogy, databases, normal forms, theory-first education, practice-first education, experiential learning

ACM Reference Format:

Dakota C. Cookenmaster, Jacob A. Bahn, and Germán H. Alférez. 2022. Introducing Normal Forms to Students: A Comparison Between Theory-First and Project-First Educational Approaches. In . ACM, New York, NY, USA, 8 pages.

1 INTRODUCTION

Approaches to education have been pondered, adopted, reconsidered, and reapplied with substantial vigor over the last millennium. The academic community has proven itself to be dedicated to providing students with educational methods and approaches that encourage true learning and problem-solving over rote memorization as shown in [3]. In this way, it is expected that these students will be true pioneers in their respective fields and continue the work we fight so hard to produce today. Given this strong and ardent dedication to true education, it is required that the community consistently reassess their instructional methods with the intention of auditing their practices, as emphasised in [4]. The hopeful result is a cohesive and comprehensive learning approach that adapts and changes as necessary to best fit the needs of the students.

One of the most challenging topics for students to learn in Computer Science (regarded in [8] as a notoriously challenging subject area) – specifically database systems – is the concept of a “Normal Form.” For the reader unfamiliar with database architectural patterns, database schema architects frequently follow a tiered system of “Normal Forms,” or rather certain cascading rules designed to protect data integrity and prevent data duplication in database schema design [7][6]. The theoretical terminology introduced as part of these forms is often convoluted at worst and confusing at best. Examples of this can be seen in [5].

*Both authors contributed equally to this research.

53 Students seeking to learn these normal form definitions, with the explicit goal of using them for real-world design
54 projects, are often stumped by the theoretical verbiage and thus hesitate to apply the knowledge in their designs. A
55 question thus arises: “Is the theory the problem, the instruction of the theory, or the ordering of when theory and
56 practice are encountered?” This paper hopes to provide insight into how instructors should consider the education of
57 normalization in databases, with specific emphasis on student comfort and retention.
58

59 Section 2 of this paper seeks to provide some historical background on practice-based pedagogy in computer science
60 and elsewhere. Section 3 describes the concepts behind the practice-based approach we encourage educators to use.
61 Section 3 emphasizes the learning method proposed and our hypothesis for it. Section 4 explains the steps taken to
62 develop and execute a small study on the hypothesis. Section 5 details the result of the study. Section 6 discusses the
63 results found in the previous section. Lastly, section 7 presents conclusions and future work.
64
65
66

67 2 LITERATURE REVIEW

68 Advances in computer science tied with those in education have resulted in a nearly ubiquitous positive change for
69 students. In this way, in [8] researchers recognized that practical student involvement in computer science education
70 was fundamentally essential for student growth. This study highlighted the importance of implementing the principles
71 of experiential learning in order to obtain significant progress. Even between disciplines, the authors in [3] recognize
72 the helpful nature of experience-based pedagogy. In their research, the authors found that this philosophy, as opposed
73 to solely theory, assisted student farmers in Slovak towns and was, by and large, a success. Its methods are gradually
74 being adopted by the general public and the broader academic community.
75
76

77 Relating to the concept of engineering, practice-based learning (PBL) can successfully be applied in engineering
78 programs, as mentioned in [5]. Also, the focus should be placed on application and integration of knowledge rather than
79 on knowledge acquisition. In [4], the authors state that teachers who underwent practice-based teacher professional
80 development training were able to more readily adapt new long-lasting positive pedagogical changes, aiding in the
81 theory that experiential learning is effective. The authors of [2], found that practice-based pedagogy was able to greatly
82 enhance teaching models. With the concept of web design, in [1], a controlled study found that experiential learning
83 was suited for complicated topics and instructors were able to focus more on problem-solving activities. Recent research
84 thus corroborates the idea that integrating a strong emphasis on experiential and practice-based learning is largely
85 positive for students and can be utilized to assist in the education of challenging topics, particularly those in computer
86 science.
87
88
89

90 In the context of databases, normal forms are a challenging collection of topics related to database schema design; it
91 is a vital component for ensuring data integrity and limiting data redundancy. The implementation challenge arises
92 because database designers have to make sure that non prime attributes do not depend on a pure subset of the database
93 table’s candidate key and also do not depend on another non-prime attribute. The only way to check that the database
94 meets these requirements is by manually parsing through the database. Thus meaning that a database schema needs to
95 be constantly checked and normalized if new tables are added to the schema [7].
96
97

98 Although there are positive results when instructors and institutions engaged in practice-based pedagogy, we
99 discovered that there is an absence of sufficient research in the area of database learning methods, and particularly
100 Normal Forms. In fact, although [7] and [6] discuss methods for normalization, there is no insight regarding the way
101 normalization should be taught. Furthermore, [8] and [5] argue for the use of experiential- and problem-based learning
102 in computer science and engineering, but the study of databases is not explicitly mentioned.
103
104

3 THEORETICAL FRAMEWORK

Figure 1 explains the general concept behind this research, namely that practice-based pedagogy should diverge from theory-based pedagogy through the explicit inclusion of practice at the forefront of the learning experience. Theory-based pedagogy often relies firstly on theory and may potentially exclude real-world examples entirely. An anecdotal example of this might be a simple math problem for children where an individual walks to the supermarket to purchase some 200 watermelons. Theory-only education will never suffice in preparing students for work in the real world, especially considering the fact that theory-based education often acknowledges little to no practical limitations. For database systems, this is wholly unhelpful, as all database systems have physical constraints, policy constraints, and real-world data requirements. This paper does not seek to argue the merit of theory, which is most certainly a necessity in the educational process. Rather, this paper seeks to encourage real-world practical examples in education first, which come to eventually rest squarely on solid, theoretical foundations.

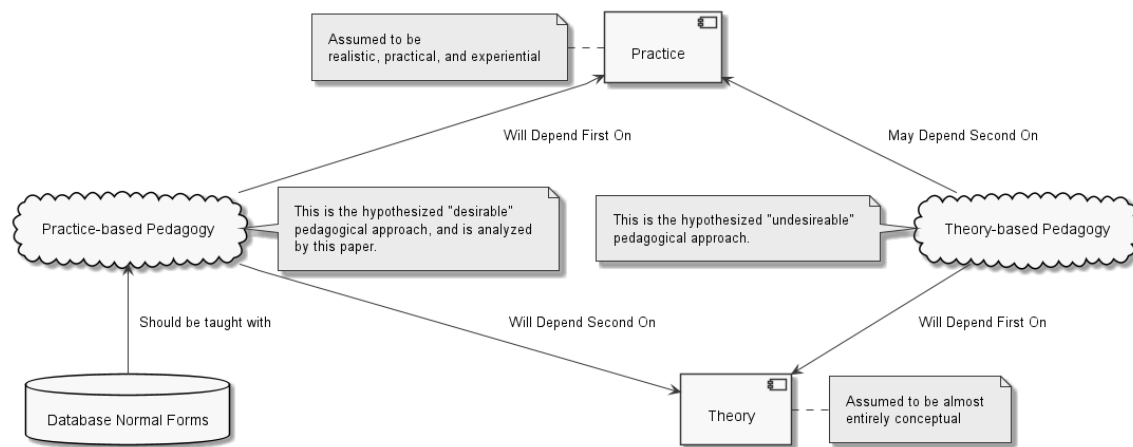


Fig. 1. Conceptual map describing different pedagogical approaches and some related examples.

4 METHODOLOGY

Given the pressure by the academic community to move towards a practice-based approach, and more specifically in computer science education, it would stand to reason that other challenging computing topics would also benefit from experiential education. In particular, this paper analyzes the pedagogy behind database schema design and development using normal forms. While research conducted by authors in [6] has been undertaken to assist in making normal form theory easier to understand, this paper hypothesizes that – given an alternative approach to standard theory-based education – students will be even more readily able to understand why normal forms and their implications and impacts are essential, despite the proven complexity in making such database schema determinations outright, as expressed in [7].

This section covers the steps taken to test the aforementioned hypothesis. These steps included preparing the experiment, creating a survey, conducting the experiment, and gathering the results.

157 4.1 Preparing Experiment

158 The preparation of this experiment consisted of a strong consideration of the audience, the development of optimal
159 lecture methodologies, and focus on a relatively challenging topic in the database sciences that was primarily unknown
160 to the audience. The experiment was conducted in the undergraduate Database Management Systems class at Southern
161 Adventist University in Collegedale, Tennessee, USA.
162

- 163 (1) Audience: In this experiment, the audience is comprised entirely of university students enrolled in an introduc-
164 tory database course. These students have general knowledge about database design and modelling (concepts
165 including, but not limited to: tables, tuples, attributes, primary keys, candidate keys, etc.), but they had not yet
166 been introduced to database normal forms nor concepts related to deduplication, consistency, or isolation.
167
- 168 (2) Determining Topic: The research topic focuses on educating students in the normalization of data in a relational
169 database system, with specific emphasis on 3NF.
170
- 171 (3) Lecture Methodology: The lecture methodology utilized in this research reflects a comparison of theory-
172 based learning and practice-based learning. In the theory-based pedagogy conducted, theoretical concepts
173 and terminology were introduced first with generic (and potentially unrelated) examples provided to explain
174 differences between the different normal forms, and how to identify them. Little to no emphasis was placed on
175 walking students through a real-world example. Rather, this pedagogical approach focused on key terms and
176 rote memorization over application. The practice-based learning presented in this paper, on the other hand,
177 began with a plausible real-world example. Students were encouraged to build new functionality into a system
178 given a series of requirements, and eventually the students normalized the data in order to prevent issues that
179 arose. Theory is not completely avoided in this approach as terminology must still be introduced. However, the
180 emphasis focused mostly on developing a solution to the problem.
181
182
183
184

185 4.2 Creating Survey

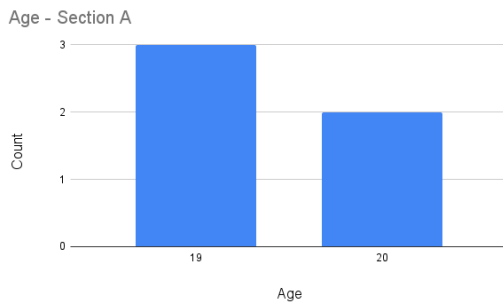
186 One of the important things taken into consideration during the creation of the post-lecture survey was the intentional
187 exclusion of leading questions (i.e., questions that illicit a specific answer). The survey initially inquires about how
188 comfortable the participant was with 3NF before the experiment, followed by their level of comfort post-experiment.
189 Also asked were questions relating to the student's comfort with the instructional methodology. All questions utilized
190 fall on an integral scale from 1 - 5.
191
192
193

194 4.3 Conducting Experiment

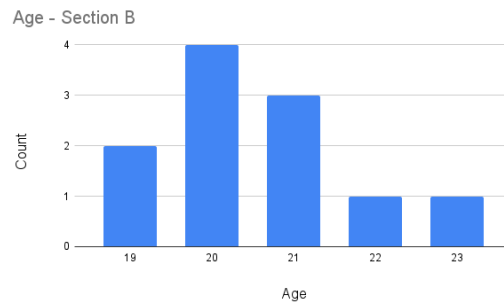
195 A group of 16 students between the ages of 18 and 23 were broken into two smaller groups, which this study labels
196 Section A and Section B, respectively. Section A was comprised of 5 students, and Section B was comprised of 11
197 students; for a detailed breakdown, see Figure 2. Each group was given a lecture, approximately 20 minutes long,
198 regarding the concept of Database Normal Forms. After the lectures were complete, the students were asked to complete
199 a survey regarding their level of comfort with normal forms both before and after the lecture, their perceptions on the
200 instructional methodology provided, and their opinions on when theory should be introduced in the classroom.
201

202 Section A was provided with a Theory-first approach to normal forms. Definitions on relational database terminology
203 were provided upfront, followed by specific, pointed, examples of designs that violated the normal form being presented.
204 Finally, the students thoroughly observed an example of a design which violated Third Normal Form and were shown
205 the proper way to correct the violation.
206
207
208

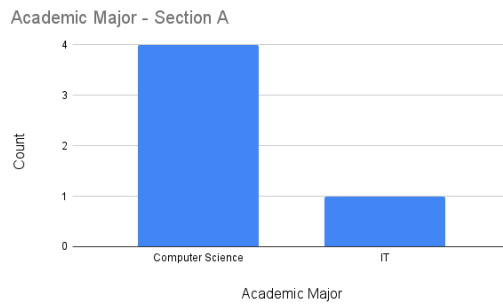
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260



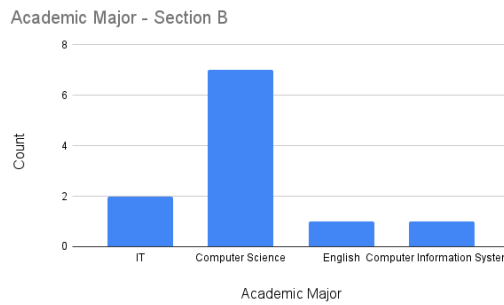
(a) Ages for Section A



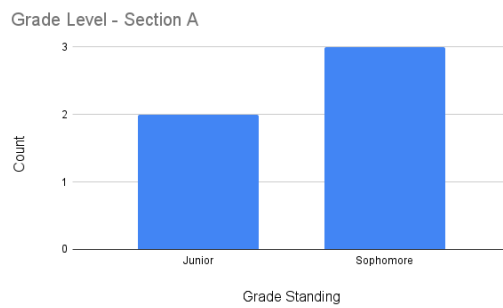
(b) Ages for Section B



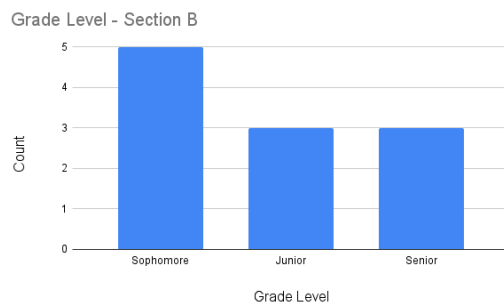
(c) Academic Majors for Section A



(d) Academic Majors for Section B



(e) Grade Levels for Section A



(f) Grade Levels for Section B

Fig. 2. Participant data shows varied ages, majors, and grade levels among survey groups.

Section B was provided with a Practice-first approach to normal forms. The lecture began with a real world example where the students brainstormed how to extend a database schema given a series of practical requirements. Following this, the students were provided examples of database designs that violated normal forms. Throughout these steps, they were asked targeted questions related to normal forms and pressed to update the real world example to be 3NF-compliant.

261 4.4 Gathering Results

262 The results-gathering took place via an online survey, conducted on the Google Forms platform. Students were asked
263 various questions about their level of comfort with normal forms before and after the lecture as well as their beliefs on
264 when theory should be introduced in the classroom. Other specific information gathered includes, name, age, grade
265 level, and academic major.
266
267

268 5 RESULTS

269
270 A series of questions were asked in a post-lecture survey, which were used to help determine whether or not students
271 were more or less comfortable with database normal forms after learning from a particular instructional methodology.
272
273

274 5.1 Level of Comfort Before the Lecture

275 Students were asked, on a scale from 1 - 5, how comfortable they were with the concept of database normal forms
276 before the lecture. A student selecting 1 would indicate low comfort, and a 5 would indicate high comfort. Students in
277 Section A were, on average, indifferent to the concept of database normal forms prior to the lecture. The calculated
278 average was $3.2 / 5$, which rounded to the nearest integer would be: **3 - Indifferent**. Students in Section B were, on
279 average, indifferent to the concept of database normal forms prior to the lecture. The calculated average was $2.7 / 5$,
280 which rounded to the nearest integer would be: **3 - Indifferent**.
281
282

283 5.2 Level of Personal Comfort After the Lecture

284 Students were asked, on a scale from 1 - 5, how comfortable they were with the concept of database normal forms after
285 the lecture, specifically how comfortable they would feel using 3NF for their own personal schema designs. A student
286 selecting a 1 would indicate low comfort, and a 5 would indicate high comfort. Students in Section A were, on average,
287 indifferent to the concept of database normal forms after the lecture. The calculated average was $3.4 / 5$, which rounded
288 to the nearest integer would be: **3 - Indifferent**. While the overall result can be reduced to a 3, it is important to note
289 that the students felt, overall, more comfortable than they did before the lecture. Students in Section B were, on average,
290 indifferent to the concept of database normal forms after the lecture. The calculated average was $3.1 / 5$, which rounded
291 to the nearest integer would be: **3 - Indifferent**. While the overall result can be reduced to a 3, it is important to note
292 that the students felt, overall, more comfortable than they did before the lecture.
293
294
295
296
297

298 5.3 Level of Comfort Explaining 3NF to a Peer After the Lecture

299 Students were asked, on a scale from 1 - 5, how comfortable they would be explaining 3NF to a peer after the lecture. A
300 student selecting a 1 would indicate low comfort, and a 5 would indicate high comfort. Students in Section A were, on
301 average, comfortable with the concept of explaining 3NF to a peer after the lecture. The calculated average was $3.8 / 5$,
302 which rounded to the nearest integer would be: **4 - Comfortable**. Students in Section B were, on average, indifferent
303 to the concept of explaining 3NF to a peer after the lecture. The calculated average was $3.4 / 5$, which rounded to the
304 nearest integer would be **3 - Indifferent**.
305
306
307

308 5.4 Normal Form Complexity

309 Students were asked, on a scale from 1 - 5, what their impressions were on the complexity of database normal forms.
310 Specifically, a statement posits: "I find third normal form easy to understand." A student selecting 1 would indicate
311
312

313 strong disagreement, and a 5 would indicate strong agreement. Students in Section A were, on average, in agreement
314 with this statement. The calculated average was $3.6 / 5$, which rounded to the nearest integer would be: **4 - Agree**.
315 Students in Section B were, on average, in agreement with this statement. The calculated average was $3.6 / 5$, which
316 rounded to the nearest integer would be: **4 - Agree**.
317

318

319 **5.5 Test Readiness**

320 Students were asked, on a scale from 1 - 5, what their impressions were on their own personal readiness for a test
321 on 3NF. Specifically, a statement posits: "If a test were given today on Third Normal Form, I would ace it." A student
322 selecting a 1 would strongly disagree with this statement, and a 5 would indicate strong agreement. Students in Section
323 A were, on average, indifferent to this statement. The calculated average was $3 / 5$, which would be: **3 - Indifferent**.
324 Students in Section B were, on average, indifferent to this statement. The calculated average was $2.5 / 5$, which rounded
325 to the nearest integer would be: **3 - Indifferent**.
326
327

328

329 **5.6 Demonstration**

330 Students were asked, on a scale from 1 - 5, how intuitive they found the lecture. Specifically, a statement posits: "I found
331 the demonstration intuitive." A student selecting a 1 would indicate strong disagreement with this statement, and a
332 5 would indicate strong agreement. Students in Section A were, on average, in agreement with this statement. The
333 calculated average was $4 / 5$, which would be **4 - Agree**. Students in Section B were, on average, in agreement
334 with this statement. The calculated average was $4.2 / 5$, which rounded to the nearest integer would be **4 - Agree**.
335
336
337

338 **5.7 Educational Approach**

339 Students were asked, on a scale from 1 - 5, how much they enjoyed the educational approach. Specifically, a statement
340 posits: "I enjoyed the educational approach used in the demonstration." A student selecting a 1 would indicate strong
341 disagreement, and a 5 would indicate strong agreement. Students in Section A were, on average, indifferent to this
342 statement. The calculated average was $3.4 / 5$, which rounded to the nearest integer would be **3 - Indifferent**. Students
343 in Section B were, on average, in agreement with this statement. The calculated average was $4.2 / 5$, which rounded
344 to the nearest integer would be **4 - Agree**.
345
346
347

348 **5.8 Predisposition to Theory**

349 Students were asked, on a scale from 1 - 5, how much they believed the introduction of theory was important to
350 introduce first in education. Specifically, a statement posits: "It is important to learn theory before engaging in practice."
351 A student selecting a 1 would indicate strong disagreement with this statement, a 5 would indicate strong agreement.
352 Students in Section A were, on average, in agreement with this statement. The calculated average was $4 / 5$, which
353 would be **4 - Agree**. Students in Section B were, on average, in agreement with this statement. The calculated average
354 was $4.1 / 5$, which rounded to the nearest integer would be **4 - Agree**.
355
356
357

358 **5.9 Discussion**

359 The authors of this paper recognize that the sample size is small and acknowledge the limitations of these results.
360 Classroom size limitations, student availability, and time constraints resulted in a smaller-than-desirable set of students.
361 However, the methodology and results presented herein can be used to fuel another, larger study.
362
363
364

365 Based on the data collected, while it is currently indeterminate as to whether or not students were more or less
366 comfortable with the concept of normal forms after the experiment, what is certainly clear is that students across all
367 groups have a strong predisposition to theory-first education. Across both groups, the average was 4 or above (general
368 agreement) when asked how much they believed that the introduction of theory was important to introduce first in
369 education. This has lasting implications for educators, as students currently believe that theory is important to introduce
370 first, whether or not it actually behooves them.
371

372 That being said, it should be noted that while students had a strong predisposition to theory-first education, those in
373 the practice-first lecture were reportedly more comfortable on average with the educational approach afforded them, as
374 opposed to those in the theory-first lecture, who were on average indifferent.
375
376

377 6 CONCLUSION AND FUTURE WORK

378 This paper explored two pedagogical methods for the instruction of normal forms in database management systems.
379 The first approach was theory-based and the second was project-based. For testing these methods, we separated a
380 university class' students into two groups and presented normal forms with the theory-based approach to the first
381 group and the project-based approach to the second group. While the data collected here has definite limitations, what
382 is clear is the bias students have towards theory in education, whether or not it is truly beneficial to their learning
383 progress. Assisted by the preliminary results suggesting that students may well prefer a practice-based approach, we
384 propose that another, larger study be conducted to test this hypothesis with the hope that practice-based learning be
385 more heavily tested (and eventually adopted) in the classroom.
386
387
388
389

390 REFERENCES

- 391 [1] Maria Jakovljevic and Piet Ankiewicz. 2015. Project-based pedagogy for the facilitation of webpage design. *Int J Technol Des Educ* 7 (may 2015).
392 <https://doi.org/10.1007/s10798-015-9312-5>
- 393 [2] Limin Jao, Dawn Wiseman, Marta Kobiela, Allison Gonsalves, and Annie Savard. 2018. Practice-Based Pedagogy in Mathematics and Science teaching
394 Methods: Challenges and Adaptations in Context. *Ontario Institute for Educational Studies* 20 (jun 2018).
- 395 [3] Katarina Slobodová Nováková and Zuzana Giertlová. 2016. New models of theoretical and practical education in urban environment (On example of
396 experience-based pedagogy in Slovak Towns). *Procedia - Social and Behavioral Sciences* 228, 12 (June 2016), 305–310.
- 397 [4] Shannon Pella. 2015. Pedagogical Reasoning and Action: Affordances of Practice-Based Teacher Professional Development. *Teacher Education*
398 *Quarterly* 42 (2015), 81–101. <https://www.jstor.org/stable/10.2307/teaceducquar.42.3.81>
- 399 [5] J.C. Perrenet, P.A.J. Bouhuijs, and J.G.M.M. Smits. 2000. The Suitability of Problem-based Learning for Engineering Education; Theory and Practicet.
400 *Teaching in Higher Education* 5 (2000), 345–358. <https://doi.org/10.1080/713699144>
- 401 [6] Betty Salzberg. 1986. Third Normal Form Made Easy. *ACM SIGCSE Bulletin* 18, 1 (June 1986), 64–74. <https://doi.org/10.1145/15075.15092>
- 402 [7] Hyontai Sug. 2019. A Method for Normalization of Relation Schema Based on Data to Abide by the Third Normal Form. *WSEAS TRANSACTIONS on*
403 *MATHEMATICS* 19, 10 (December 2019). <https://doi.org/10.37394/23206.2020.19.20>
- 404 [8] Katerina Tzafilkou, Nicolaos Protogeros, and Adamandi Chouliara. 2019. Experiential learning in web development courses: Examining students'
405 performance, perception and acceptance. *Education and Information Technologies* 24 (December 2019). <https://doi.org/10.1007/s10639-020-10211-6>
406
407
408
409
410
411
412
413
414
415
416