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.

2022. Manuscript submitted to ACM

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

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

2 LITERATURE REVIEW

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

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

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

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

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.

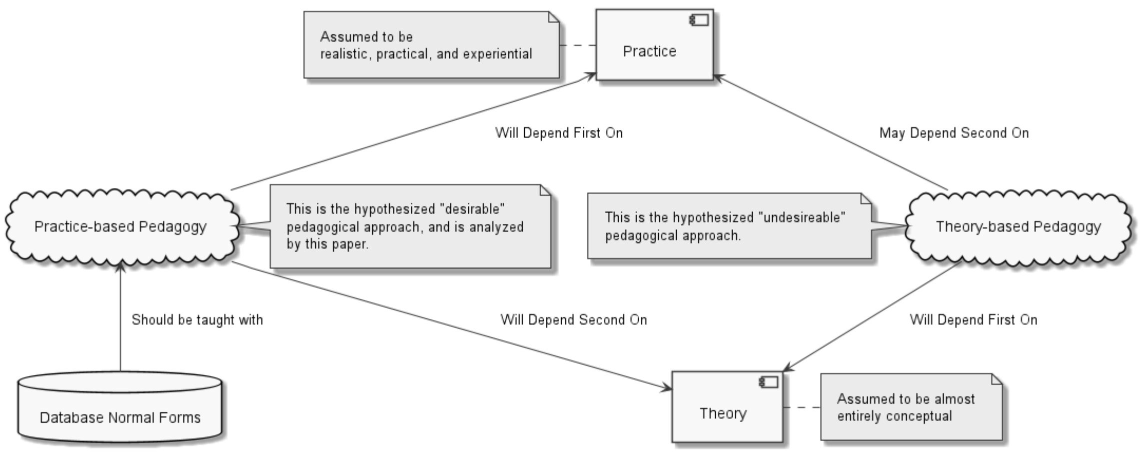


Fig. 1. Conceptual map describing different pedagogical approaches and some related examples.

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.

4.1 Preparing Experiment

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

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

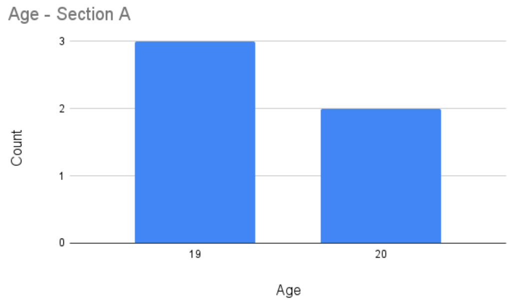
4.2 Creating Survey

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

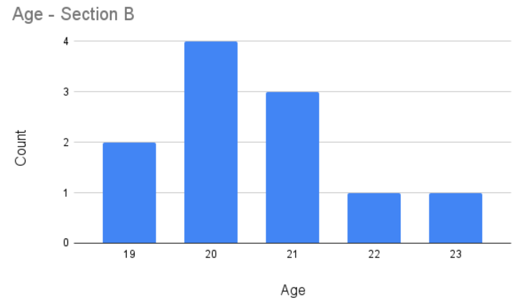
4.3 Conducting Experiment

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

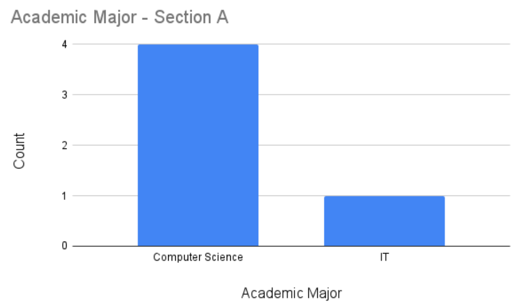
Section A was provided with a Theory-first approach to normal forms. Definitions on relational database terminology were provided upfront, followed by specific, pointed, examples of designs that violated the normal form being presented. Finally, the students thoroughly observed an example of a design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



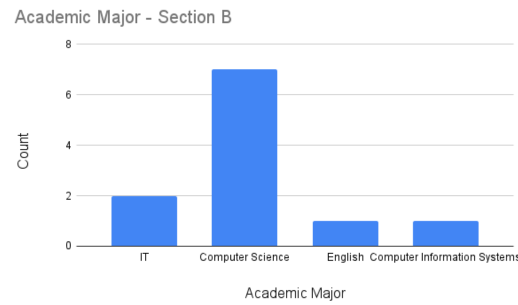
(a) Ages for Section A



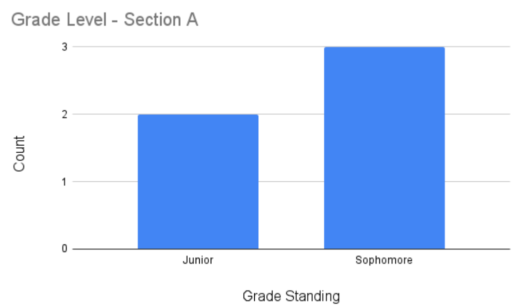
(b) Ages for Section B



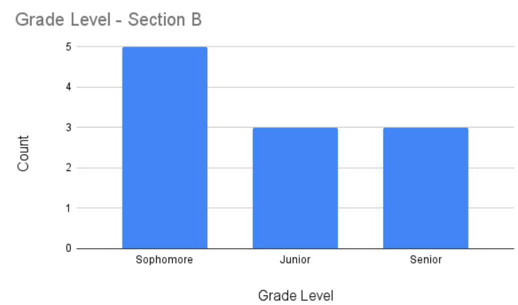
(c) Academic Majors for Section A



(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.

4.4 Gathering Results

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

5 RESULTS

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

5.1 Level of Comfort Before the Lecture

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

5.2 Level of Personal Comfort After the Lecture

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

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

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

5.4 Normal Form Complexity

Students were asked, on a scale from 1 - 5, what their impressions were on the complexity of database normal forms. Specifically, a statement posits: "I find third normal form easy to understand." A student selecting 1 would indicate strong disagreement, and a 5 would indicate strong agreement. Students in Section A were, on average, in agreement with this statement. The calculated average was 3.6 / 5, which rounded to the nearest integer would be: **4 - Agree**. Students in Section B were, on average, in agreement with this statement. The calculated average was 3.6 / 5, which rounded to the nearest integer would be: **4 - Agree**.

5.5 Test Readiness

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

5.6 Demonstration

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

5.7 Educational Approach

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

5.8 Predisposition to Theory

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

5.9 Discussion

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

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

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

6 CONCLUSION AND FUTURE WORK

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

REFERENCES

[1] Maria Jakovljevic and Piet Ankiewicz. 2015. Project-based pedagogy for the facilitation of webpage design. *Int J Technol Des Educ* 7 (may 2015). https://doi.org/10.1007/s10798-015-9312-5

[2] Limin Jao, Dawn Wiseman, Marta Kobiela, Allison Gonsalves, and Annie Savard. 2018. Practice-Based Pedagogy in Mathematics and Science teaching Methods: Challenges and Adaptations in Context. *Ontario Institute for Educational Studies* 20 (jun 2018).

[3] Katarina Slobodová Nováková and Zuzana Giertlová. 2016. New models of theoretical and practical education in urban environment (On example of experience-based pedagogy in Slovak Towns). *Procedia - Social and Behavioral Sciences* 228, 12 (June 2016), 305–310.

[4] Shannon Pella. 2015. Pedagogical Reasoning and Action: Affordances of Practice-Based Teacher Professional Development. *Teacher Education Quarterly* 42 (2015), 81–101. https://www.jstor.org/stable/10.2307/teaceducquar.42.3.81

[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. *Teaching in Higher Education* 5 (2000), 345–358. https://doi.org/10.1080/713699144

[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

[7] Hyontai Sug. 2019. A Method for Normalization of Relation Schema Based on Data to Abide by the Third Normal Form. *WSEAS TRANSACTIONS on MATHEMATICS* 19, 10 (December 2019). https://doi.org/10.37394/23206.2020.19.20

[8] Katerina Tzafilkou, Nicolaos Protogeros, and Adamandi Chouliara. 2019. Experiential learning in web development courses: Examining students' performance, perception and acceptance. *Education and Information Technologies* 24 (December 2019). https://doi.org/10.1007/s10639-020-10211-6