



Introducing Databases in Context Through Customizable Visualizations

Suzanne W. Dietrich^{1*}, Don Goelman², Jennifer Broatch¹, Sharon Crook³, Becky Ball¹, Kimberly Kobjek¹ and Jennifer Ortiz⁴

¹School of Mathematical and Natural Sciences, Arizona State University, Phoenix, AZ, United States, ²Department of Computing Sciences, Villanova University, Villanova, PA, United States, ³School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, United States, ⁴Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, United States

OPEN ACCESS

Edited by:

Kevin Buffardi,
California State University, Chico,
United States

Reviewed by:

Bob Edmison,
Virginia Tech, United States
Mohammed Fawzi Farghally,
Virginia Tech, United States

*Correspondence:

Suzanne W. Dietrich
dietrich@asu.edu

Specialty section:

This article was submitted to
Digital Education,
a section of the journal
Frontiers in Education

Received: 20 July 2021

Accepted: 27 August 2021

Published: 09 September 2021

Citation:

Dietrich SW, Goelman D, Broatch J,
Crook S, Ball B, Kobjek K and Ortiz J
(2021) Introducing Databases in
Context Through
Customizable Visualizations.
Front. Educ. 6:719134.
doi: 10.3389/feduc.2021.719134

The goal of the Databases for Many Majors project is to engage a broad audience in understanding fundamental database concepts using visualizations with color and visual cues to present these topics to students across many disciplines. There are three visualizations: introducing relational databases, querying, and design. A unique feature of these learning tools is the ability for instructors in diverse disciplines to customize the content of the visualization's example data, supporting text, and formative assessment questions to promote relevance to their students. This paper presents a study on the impact of the customized introduction to relational databases visualization on both conceptual learning and attitudes towards databases. The assessment was performed in three different courses across two universities. The evaluation shows that learning outcomes are met with any visualization, which appears to be counter to expectations. However, students using a visualization customized to the course context had more positive attitudes and beliefs towards the usefulness of databases than the control group.

Keywords: databases, visualizations, contextualization, interdisciplinary, assessment

INTRODUCTION

The ubiquity of databases extends to all sectors of society and academia. Data analysts are in high demand. One aspect of data analysis is the use of database technology, which has been typically taught at universities in upper-level courses in computer science or business, and is therefore not accessible to the variety of university major programs that could benefit from the knowledge. Database technology is also an essential component of data science, providing a powerful data wrangling and analysis tool in a variety of application domains. The goal of the Databases for Many Majors project is to make fundamental database concepts accessible to learners in diverse disciplines, including the capability to use discipline-specific examples. The motivation for the project began with the introduction of database fluency classes for non-majors (Goelman, 2008). The initial funding focused on developing an engaging curriculum with collaborative learning opportunities to provide students of many backgrounds with an introduction to relational databases and querying. The project chose the use of a dynamic presentation that guides the student, who has no prior knowledge of databases, through the concepts using visual cues and color. Thus, the learning tools are called visualizations, since they are communicating abstract information in a visual manner. Collaborative learning modules were also developed in support of the visualizations. Subsequent funding enhanced the existing visualizations with a self-assessment framework, added a new visualization on how to design a relational database, and emphasized the customizations of the

visualizations for various STEM disciplines. Another facet of this latter phase of the project included an examination of the accessibility of the visualizations (Bingham et al., 2018), resulting in a revised framework for the visualizations that satisfies the AA performance level of the Web Contrast Accessibility Guidelines (W3C, 2008) and included a visual cue in addition to color to focus the user's attention.

The project has developed three visualizations that sequentially build knowledge, which are freely available to download from the project Web site (Dietrich and Goelman, 2021). The first visualization, called IntroDB, introduces the concept of a relational database and how it differs from a spreadsheet in the storage and retrieval of information; another, named QueryDB, covers the operations supported by the industry standard query language for retrieving information from a relational database; and the third, labeled DesignDB, addresses how to design a relational database for correctly storing the important concepts and their relationships needed for an application. These learning tools have a similar structure, providing topics in a navigation bar on the left-hand side to be explored by the student in succession. When selected, a topic tells its brief story, covering the essential concepts using highlighting to gain the student's attention to connect the textual content on the page with its supporting visuals. Each topic is broken down into a sequence of named steps, which are points in the story at which the student can restart the visualization within the topic for replay. The names of the steps provide an outline of the topic's story, providing a segmentation of the visualization into meaningful pieces (Spanjers et al., 2012). Some of the steps within the visualizations are interactive, giving students the ability to review the story's presentation by selecting a component that displays a corresponding visual. The visualizations also have a speed control to adjust the timing of the content display. All three visualizations also include a formative self-assessment feature, called a checkpoint, for students to check their learning (Dietrich and Goelman, 2017). Literature illustrates that formative self-assessment provides feedback during learning and improves student comprehension of learning outcomes (Cromley et al., 2016). There are twenty questions, mostly multiple-choice with some true/false, in each visualization for students to quiz themselves on their knowledge gained from the visualization.

To reach students in diverse disciplines, one of the overarching aims of the project from its inception is the ability to customize a visualization. Thus, an instructor can change the content of the example data, accompanying text, and formative assessment questions within the visualization to introduce databases in the context of their class content. The visualizations can be assigned for viewing outside of class, and in-class activities can further promote the relevance of databases in context. Note that a prior paper established that the default database visualizations, which use a university example of students taking courses, improved learning (Dietrich et al., 2015). Since that study, there are now several customizations available for students to view, including Astronomy, Computational Molecular Biology, Environmental Science, Forensics, Geographic Information Systems, Neuroinformatics, and Sports Statistics.

This paper studies the impact of the customized visualizations. Specifically, the two primary research questions being asked are: 1) Does the use of a contextualized, customized visualization impact student learning? 2) Does it affect student attitudes and beliefs towards learning databases? To assess the latter question, the study modified a validated "Computer Science Attitude Survey" for computer science (Heersink and Moskal, 2010) to databases. The attitude survey included constructs on confidence, interest, and usefulness.

The paper first briefly covers related work on visualizations, formative assessment, and contextualization in *Related Works Customization of IntroDB* presents an overview of the IntroDB visualization that was used in the study and discusses its customization, including screenshots that show the default and a customized visualization. The design of the customization impact study is presented in *Customization Impact Study*, including more details on the learning outcomes and attitudes surveys. *Customization Impact Study Results* examines the results of the study, including its limitations. Finally, the paper concludes with a discussion of customizing the visualizations to increase the relevance of databases in context.

RELATED WORKS

The development and use of the database visualizations supported by visual cues and self-assessment were guided by recommended practices in the literature. Visualizations have been shown to target a wide range of learners, engaging student interest and providing motivation to learn (Byrne et al., 1999). In addition, "dynamic visualization of ideas can enhance cognitive meaning" (Wetzel et al., 1994) and enable students "to build concrete mental models" (Ben-Ari et al., 2011). These advantages of visualizations must be mindful of results in both the fields of program visualization (PV) and algorithm visualization (AV) indicating: "how students use AV technology has a greater impact on effectiveness than what AV technology shows them" (Hundhausen et al., 2002). Naps et al., 2002 introduced an engagement taxonomy for visualizations (viewing, responding, changing, constructing, and presenting), and offered that "responding significantly improves learning over just viewing". This engagement taxonomy provided a framework for evaluating visualization systems (Urquiza-Fuentes and Velázquez-Iturbide, 2009), which indicated that viewing can improve knowledge acquisition, and adding the responding engagement can also improve student attitude. Another survey of program visualization systems introduced the 2DET engagement taxonomy correlating content ownership and direct engagement as a "refined framework that could be used to structure future research" (Sorva et al., 2013). Content ownership is broken down into the following categories: given content, own cases, modified content, and own content. The levels of the direct engagement dimension include: no viewing, viewing, controlled viewing, responding, applying, presenting, and creating. The categorization of the database visualizations in the Databases for Many Majors project according to the 2DET framework is (given content, responding). The students are given

the visualizations with predefined behavior, and the students respond to questions about the visualized concepts. Specifically, the final topic of the visualization is a formative self-assessment, which the students find quite helpful (Dietrich and Goelman, 2017).

Literature illustrates that formative self-assessment provides feedback during learning and improves student comprehension of learning outcomes (McMillan and Hearn, 2008) especially with goal orientation that promotes a mastery goal in cognitive theory. The literature also recommends that performance goals be explicit for formative self-assessment and that feedback during the assessment is always present and visual (Koorse et al., 2014). Research shows that self-assessment that is required and has clear goals promotes better student attendance, participation, and performance (DePaolo and Wilkinson, 2014). Therefore, the recommendation for incorporating the project visualizations as an assignment is to require students to achieve a performance goal (minimum score) on the self-assessment and to submit a screenshot for verification. The project promotes active learning opportunities for use in the classroom and provides sample activities on its Web site (Dietrich and Goelman, 2021), based on recommendations in the literature that visualizations be supported by additional activity to further engage learners (Naps et al., 2002; Urquiza-Fuentes and Velázquez-Iturbide, 2009).

Recall that the goal of the project is to provide an engaging introduction to fundamental database concepts to students in diverse majors using visualizations. A major objective for the visualizations is the ability to customize the content of the visualization's example in context for various disciplines. These goals are supported by educational theory, such as Situated Learning Theory (Lave and Etienne Wenger, 1991), and more recently, the MUSIC Model of Academic Motivation (Jones, 2009). Situated Learning Theory proposes that the learning pedagogy connect knowledge with authentic context and active participation. The MUSIC Model of Academic Motivation specifies key components of empowerment, usefulness, success, interest, and caring (Jones, 2009) for instructional design. The goal of the MUSIC components is to increase student motivation to learn, which increases student learning. By contextualizing the visualizations, the intention is to provide students with the application of these concepts in an authentic setting to increase student's attitudes towards the usefulness of learning databases.

In the field of algorithm visualization, there was a study that investigated whether the presentation of the algorithm within a situated storyline instead of an abstract representation might improve their understanding and retention of the algorithms (Hundhausen et al., 2004). The results of the algorithm visualization study did not show a significant effect with respect to understanding and retention, which was an unexpected result. One possible explanation posed was that the participants' computer science understanding increased in the time between the tracing sessions used in the study (Hundhausen et al., 2004).

The formative assessment feature of the database visualizations incorporates the responding engagement, which

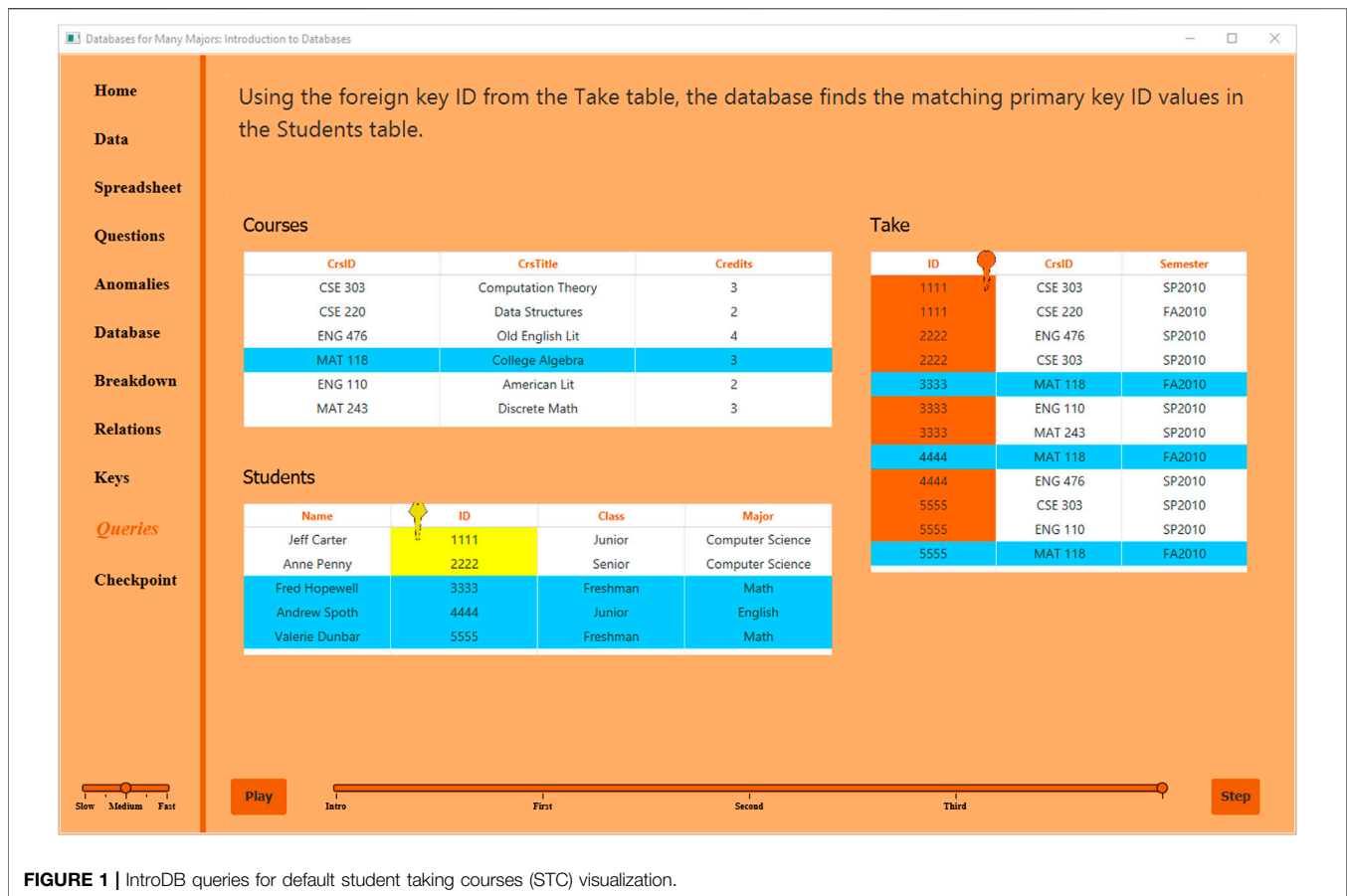
should improve knowledge acquisition and attitude (Urquiza-Fuentes and Velázquez-Iturbide, 2009). Contextualization of the visualizations to authentic context are supported by both Situated Learning Theory (Lave and Etienne Wenger, 1991) and the MUSIC Model of Academic Motivation (Jones, 2009). Therefore, the study in this paper is investigating whether the contextualizing of the IntroDB visualization affects student learning and attitudes towards databases.

CUSTOMIZATION OF INTRODB

IntroDB is the Introduction to Relational Databases visualization that builds on the student's assumed prior experience with spreadsheets, indicating their usefulness to answer certain types of questions. Moreover, it indicates the complications of other types of questions that require copying and manipulating the data to find an answer. In addition, the spreadsheet data may have issues, called anomalies, when updating, deleting, and inserting data in the spreadsheet. To avoid these anomalies, databases break down data related to specific concepts into tables, also called relations, without unnecessary repetition of information. Each table contains a primary key to uniquely identify a row in the table. These relations must be combined to answer complex questions. Therefore, the primary key of a table may be included in another table to create the association for linking the data. When a primary key of one table is included in another table, it is called a foreign key in that table. Queries then use these primary-foreign key relationships to combine relations when needed to answer a question.

Figures 1 and **2** show the last step of the Queries topic in both the default non-customized visualization, named Students Taking Courses (STC), and its customization for Computational Molecular Biology (CMB), respectively. In the STC example, the query finds the students who have taken the course titled College Algebra. The screenshot shows the last step, highlighting the orange foreign key ID in Takes referencing the yellow-gold primary key ID in Students. For CMB, the query finds the specimens that have been sequenced for the Leptin gene. The screenshot highlights the SpecID foreign key in SequenceData linking to the SpecID primary key in Specimen. The orange values must appear in the highlighted yellow-gold values. The light blue values illustrate to the learner the relevant values in the data.

The customization for IntroDB defines the enterprise and data for the example used throughout the suite of visualizations. This is not a trivial task. The customizer must design the scenario and queries in the context of their discipline and within the constraints of the visualization. There is a tool that assists the customizer in building their scenario within these constraints. The data example must be isomorphic with respect to the order of columns in the spreadsheet, whereas the row data is fully customizable, including the number of rows. Seventeen of the twenty formative self-assessment questions must be changed with respect to the contextualized example. To customize the visualization, an instructor runs the visualization in customization mode. The visualization automatically stops at



steps when customization is possible. The changes are saved in a file, which is submitted to the project leaders for curation before being incorporated with the visualization on the project's Web site.

CUSTOMIZATION IMPACT STUDY

The project employed both summative and formative assessment throughout the years (Dietrich et al., 2020). The initial summative evaluation of the learning outcomes for IntroDB and QueryDB illustrated that the visualizations, based on the default student taking courses example, improved student knowledge of fundamental database concepts through an evaluation of pre- and post- tests for several non-majors' courses across two universities (Dietrich et al., 2015). This initial study was performed before the introduction of the formative self-assessment questions. After the checkpoints were introduced, a qualitative content analysis of the open-ended question on the student perspective survey indicated that students found the checkpoints quite helpful and an important learning component of the visualizations (Dietrich and Goelman, 2017). This paper takes a first look at evaluating the impact of the customizations on conceptual learning as well as student attitudes towards databases.

Study Participants

The assessment was performed in three different courses across two universities: Computational Molecular Biology (CMB), Database Fluency (DBF) and Environmental Science (ENV). A total of 85 students participated in the study across the three courses. The Computational Molecular Biology course introduces the mathematical skills used in molecular biology, genomics, and bioinformatics. The Database Fluency class demystifies key database concepts: the relational model, normalization, the Entity-Relationship model and SQL. The Environmental Science class explores the structure, development and dynamics of terrestrial ecosystems, with a focus on the exchange of energy and materials between the atmosphere, soils, water, biosphere, and anthrosphere. This study only included the IntroDB visualization, which was covered in all three classes. A summary of the demographics of the courses are presented in **Table 1**, indicating whether the course was required or elective, the level of the students in the course, whether the completion of the IntroDB visualization was required, and whether the visualization was completed in-class or out-of-class. There were 27 different majors across the 85 students in the study, listed in alphabetical order: Accounting; Applied Math for the Life and Social Sciences; Biochemistry; Biology; Biological Sciences (Concentrations in Biomedical Sciences; Conservation Biology and Ecology; Genetics, Cell, and Developmental Biology; Neurobiology, Physiology, and Behavior); Biomedical

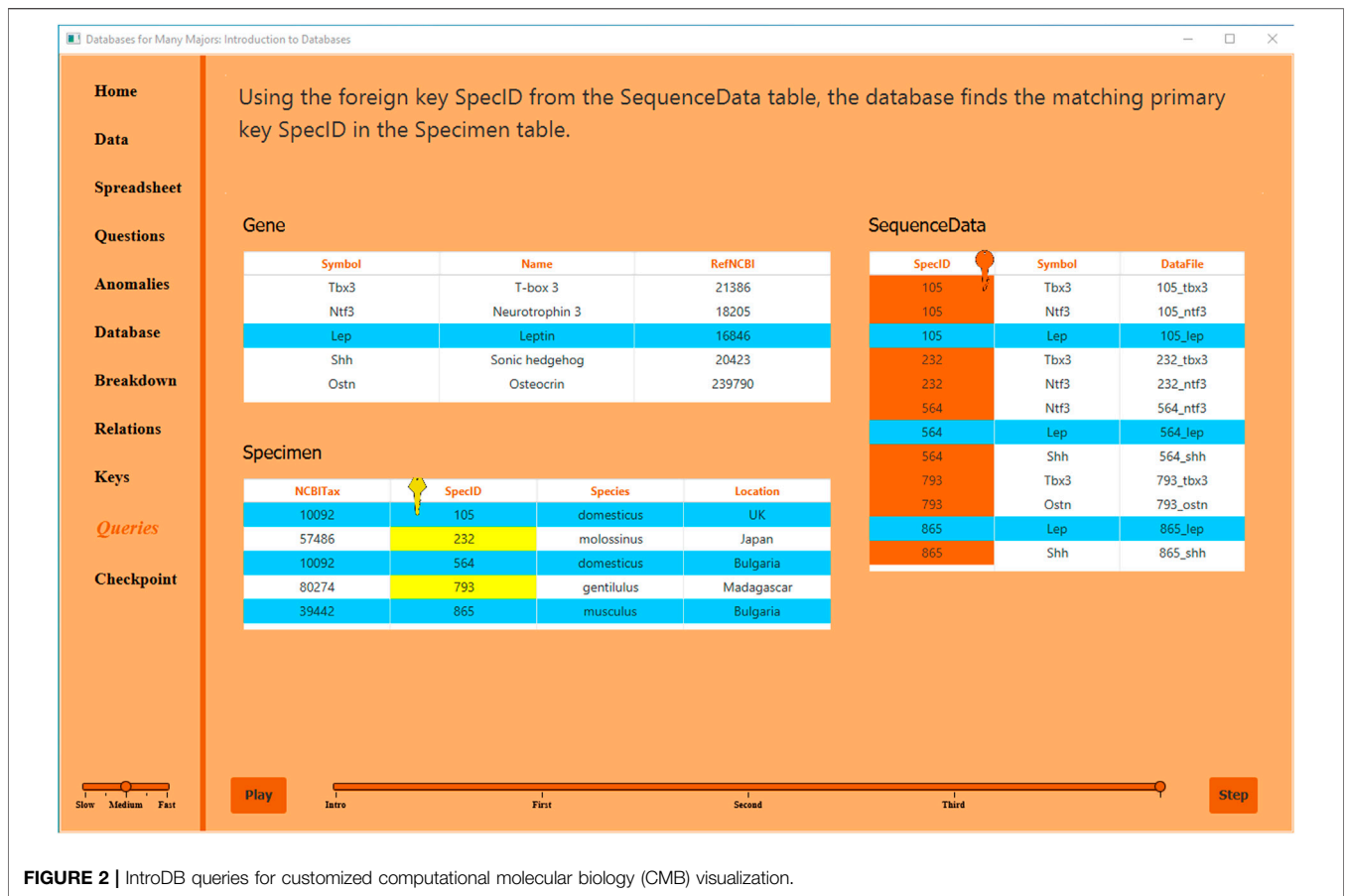


FIGURE 2 | IntroDB queries for customized computational molecular biology (CMB) visualization.

TABLE 1 | Study demographics.

Group	CMB	DBF	ENV
# of students	45	20	20
Course Required?	Elective	Elective	Required
Student Level?	Upper	All	Upper
Assignment?	Required	Extra credit	Extra credit
In or Out of class?	Out	Out	Out

Informatics; Business Analytics; Chemical Engineering; Communication; Computational Mathematics; Computer Science; Computer Systems Engineering; Earth and Environmental Studies; Economics; English; Environmental Science; Finance; Forensic Science; Global Health; History; Informatics; Interdisciplinary Studies; Management; Mathematics; Microbiology; Molecular Biosciences and Biotechnology; Real Estate.

Study Design

Each course assigned the IntroDB visualization to their class and randomly assigned (either alphabetically or random number generator) half the class to use the non-customized visualization (Students Taking Courses) and the other half to a customized version. The CMB group used the Computational

Molecular Biology customization, and similarly the ENV group used the Environmental Science customization. The DBF group was comprised of a diverse set of majors. Half of the students were assigned the non-customized visualization and the other half were allowed to select a customization of their choice from those available on the Web site, including the default visualization. Besides the Computational Molecular Biology and Environmental Science customizations, at the time of the study there were additional customizations available in Astronomy, Forensics, Geographic Information Systems, and Sports Statistics. All students were asked to complete the visualizations, and the participation in the study was voluntary. Students completed the survey assessments outside of class using an online survey.

Evaluation Tools and Methods

The impact of both the conceptual “content” learning and attitudes towards databases were assessed and compared between the customization and non-customization groups. All students were given a 20-question learning assessment, based on IntroDB topics and aligned with learning objectives formatively assessed by the checkpoint questions, and a 25-question survey of attitudes and beliefs about databases.

The learning assessment survey consisted of twenty multiple-choice questions involving a scenario of travel reservations. The learning objectives included: identifying unnecessary repetition of

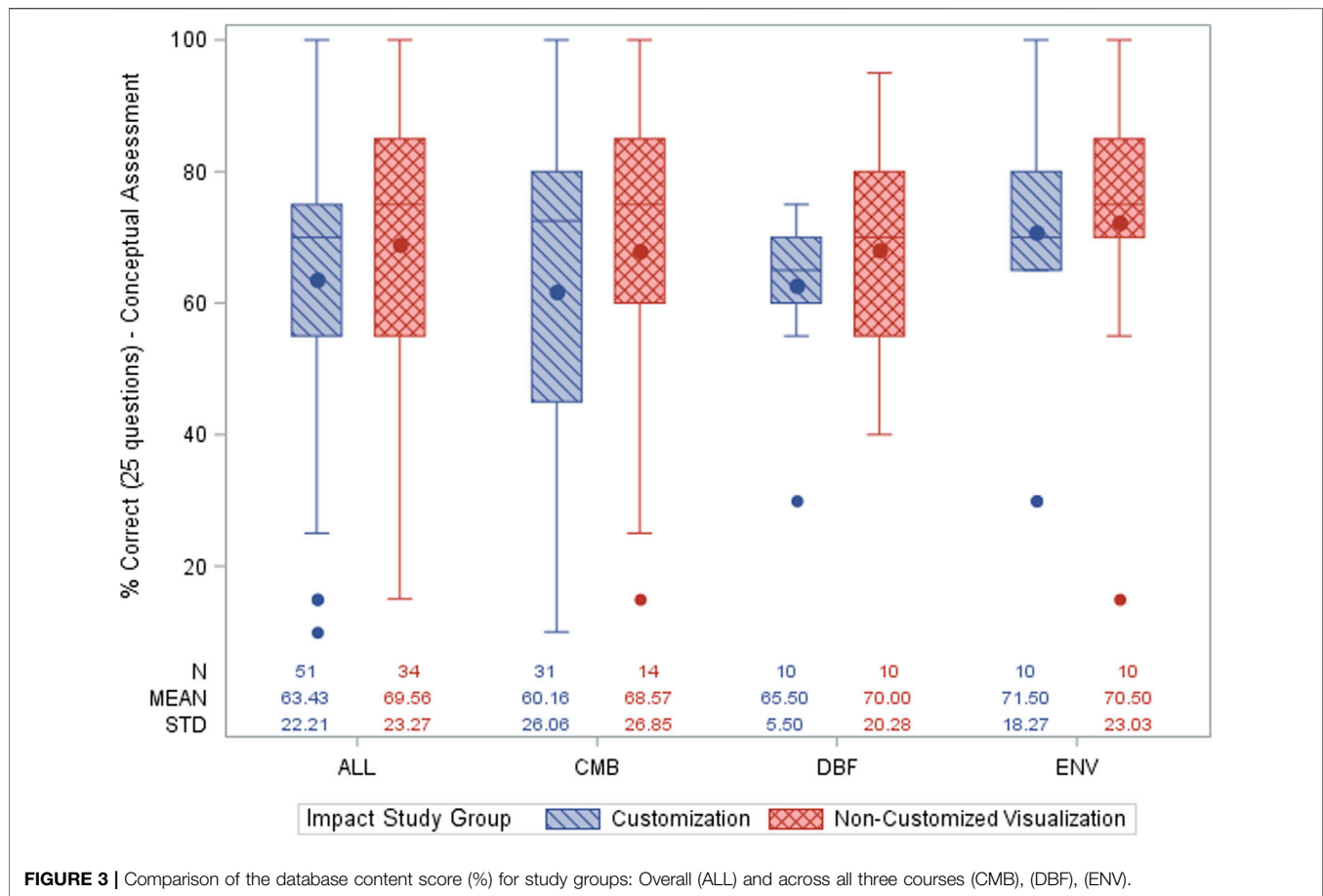


FIGURE 3 | Comparison of the database content score (%) for study groups: Overall (ALL) and across all three courses (CMB), (DBF), (ENV).

information and update/delete/insert anomalies in a spreadsheet; identifying whether anomalies exist in the given relational database; identifying primary and foreign keys of the tables in the database; and identifying the tables needed to answer a query. This survey is comparable to the original learning outcomes survey for IntroDB (Dietrich et al., 2015), which also has an item analysis of student responses (Dietrich et al., 2020). The learning outcomes are also consistent with the checkpoint questions for IntroDB (Dietrich and Goelman, 2017) in which students are asked similar questions over the spreadsheet and its corresponding relational design that are introduced within the visualization itself.

The attitudes and beliefs survey was based on a validated “Computer Science Attitude Survey” to assess attitudes and beliefs in computer science (Heersink and Moskal, 2010). The survey wording was altered by replacing “computer science/computing” with “databases,” and shortened to reflect only the validated constructs of interest to this study: confidence, interest, and usefulness. Specifically, those constructs are students’ beliefs in their confidence in their own ability to learn databases, their interest in databases, and the usefulness of learning databases. All 25 attitude and belief questions were measured on a 5-point Likert scale: Strongly Agree (5), Agree (4), Neutral (3), Disagree (2), Strongly Disagree (1). The three constructs are a summation of the individual item scores, forming a Likert-type scale score.

The survey included many negatively worded questions (e.g., “I do NOT use databases in my daily life.”) that were oppositely scored when constructing the scale score. A two-factor analysis of variance and two-sample *t*-test were used to assess the difference between the study groups (Customization v. Non-Customization) on the database content assessment and attitude assessment across all three courses.

CUSTOMIZATION IMPACT STUDY RESULTS

Figures 3–6 provide the comparison of the content, confidence, interest, and usefulness scores, respectively, for the study groups overall (ALL), and for each course (CMB, DBF, ENV). For each construct, the sample size (N), mean, and standard deviation for each group are presented at the bottom of its associated graph. Table 2 summarizes the associated *p*-values for each construct for the study groups. There was no significant impact of the customized visualization on the learning of the database content nor the confidence or interest constructs. Hence, this study reveals that the use of a customization may not impact the students’ interest or confidence in databases. There was a significant impact on the reported usefulness of databases for students that utilized the customized visualization, as shown by

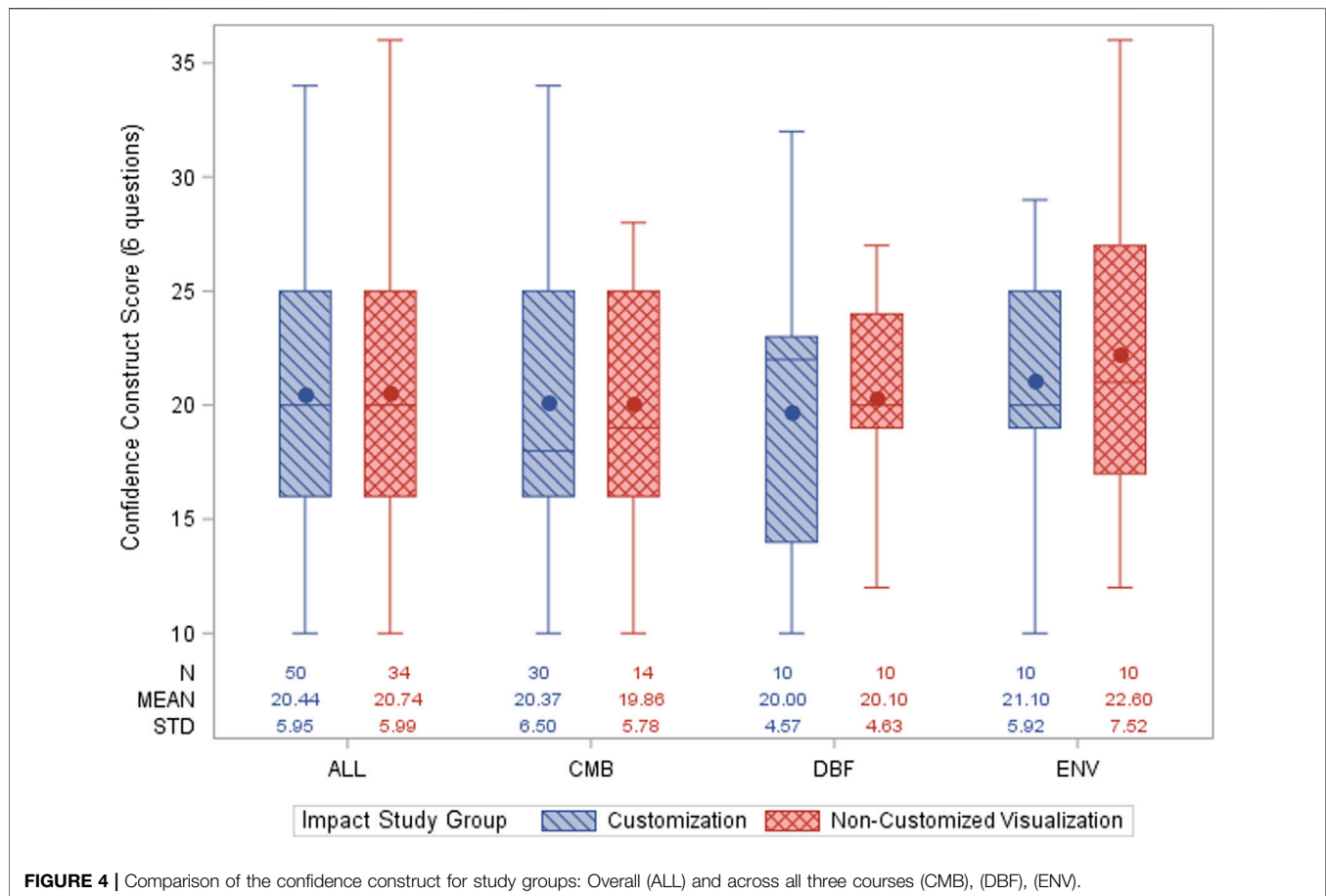


FIGURE 4 | Comparison of the confidence construct for study groups: Overall (ALL) and across all three courses (CMB), (DBF), (ENV).

the bold p -values in **Table 2**. This difference in the perceived usefulness of databases is primarily driven by the CMB course. The students that utilized the CMB customization reported significantly higher (more positive) attitudes and beliefs toward the usefulness of databases than students that did not utilize the CMB customization, whereas DBF and ENV students were similar. The usefulness construct consists of the following six questions: 1) Developing database skills will NOT play a role in helping me achieve my career goals. 2) Knowledge of database will allow me to secure a good job. 3) My career goals do NOT require that I learn database skills. 4) Developing database skills will be important to my career goals. 5) Knowledge of database skills will NOT help me secure a good job. 6) I expect that learning to use database skills will help me achieve my career goals.

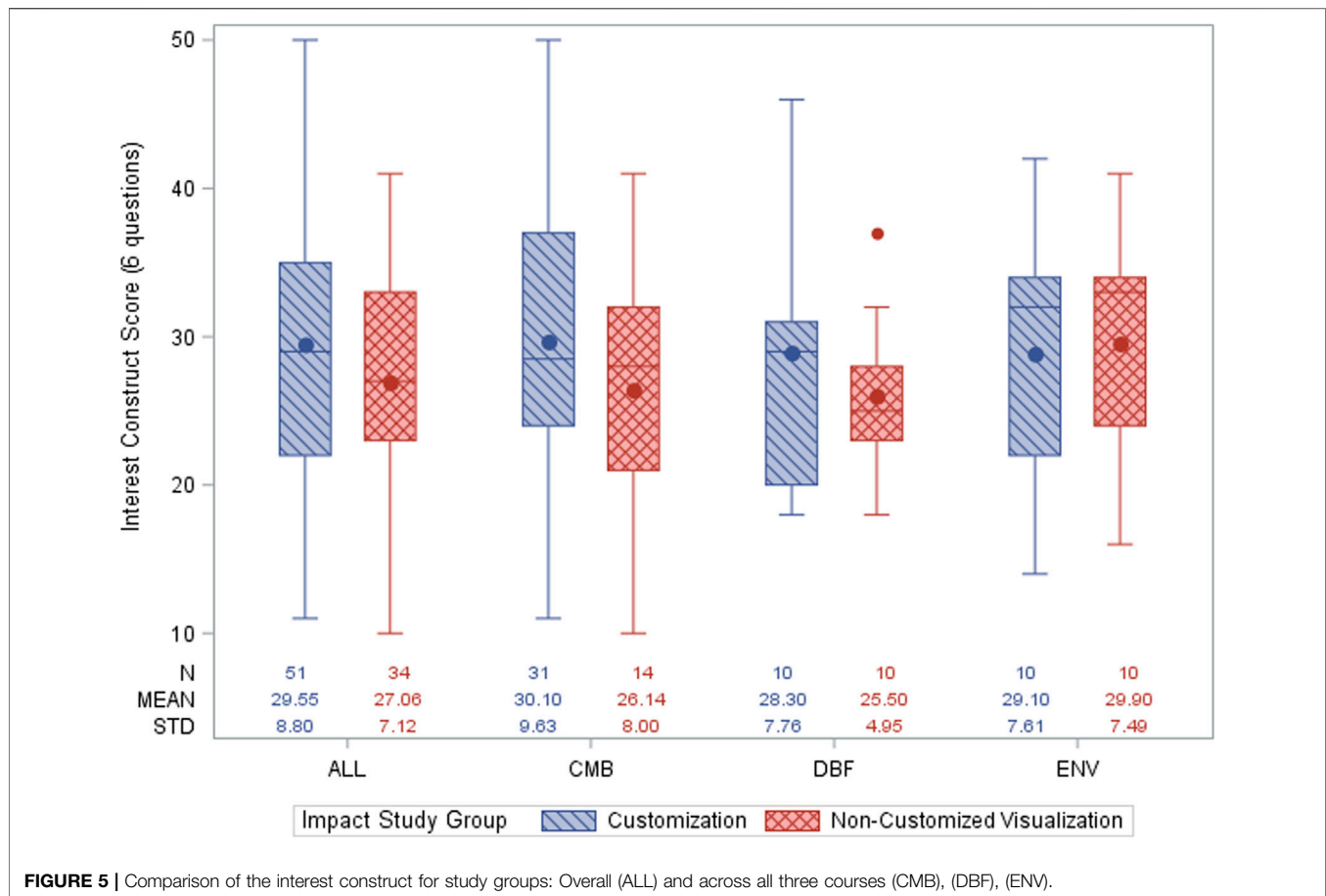
Limitations of the Study

The impact study was limited to three courses and 85 total students, and primarily focused on the computational molecular biology and environmental science IntroDB customizations. Furthermore, there were three different instructors for these classes that had very different course objectives for database coverage. Similar to many social science experiments, another limitation was based on the students following their random assignment of visualization for the study. While most students followed the given assignment, some CMB students reportedly self-

selected into the customization group potentially biasing the results. Recall that the CMB class showed a more positive attitude and beliefs toward the usefulness of databases when using a customized visualization.

Discussion

The design of the impact study purposely included both learning and attitude components. We were unsure whether the conceptual learning would illustrate statistically significant results but were hopeful that the contextualization of the visualization example and self-assessment would improve learning. We also hoped that student attitudes towards databases would show significant improvement with a visualization customized for the course content. The impact study illustrated that the learning outcomes were met regardless of which visualization was used, which appears counter to expectations. We were somewhat surprised that only the Computational Molecular Biology class found databases more useful when viewing the customized visualization, yet the interest and confidence constructs did not show any significant difference. Again, the diverse nature of the classes could have influenced the results. Perhaps the CMB class found it more useful because the assignment was required? Maybe the students in the DBF class on database fluency already were interested in databases because of their perceived usefulness



and were confident in their ability to learn databases because they were enrolled in the course? Regardless, the instructors that have customized the visualizations have shared that they appreciate the opportunity to present database concepts in the context of their discipline, and the study provided some preliminary results showing that in one class, students who utilized the customized visualization found databases more useful.

CONCLUSION

The content customization of database visualizations provides a discipline-specific example for instructors to introduce fundamental database concepts in context. The contribution of this paper is the assessment and evaluation of the impact of the conceptual learning and attitudes towards databases between students who used a customized visualization versus those that used the default visualization. The assessment was performed in three different courses across two universities. The evaluation shows that learning outcomes are met with any visualization, yet students using a visualization customized to the course context had more positive attitudes and beliefs towards the usefulness of databases than the control group using the default visualization. Students' interest and perceived relevance/usefulness of STEM topics, e.g., databases, toward their future goals are positively associated with course grades and retention in STEM (Cromley et al.,

2016). Furthermore, "both students' interest and perceived relevance can be increased by changing teaching" (Cromley et al., 2016); specifically, by highlighting the integration of the material and establishing a connection between the two subjects.

The process of customization does require an initial time investment, which can be facilitated by supervising an undergraduate student project to customize the suite of three database visualizations. Instructors indicate that the time investment is worthwhile to be able to present the relevance of database concepts in the context of their course. Some customizers have added active learning exercises in support of their customization, which appear as additional information on the project Web site. Another customizer incorporated the database visualizations as part of their statistics course (Broatch et al., 2019) with learning activities using the dplyr module (Wickham et al., 2017) for the R statistical language (R Core Team, 2017).

The visualizations are also used in database courses for majors. Most students appreciate the visual introduction to the concepts that they will learn in much more depth. Some indicated that the availability of customizations provide an opportunity to review the concepts several times in different contexts.

Although the project funding has ended, the project leaders will continue the effort. There is an increasing demand in diverse fields for data and computation skills, which also form major

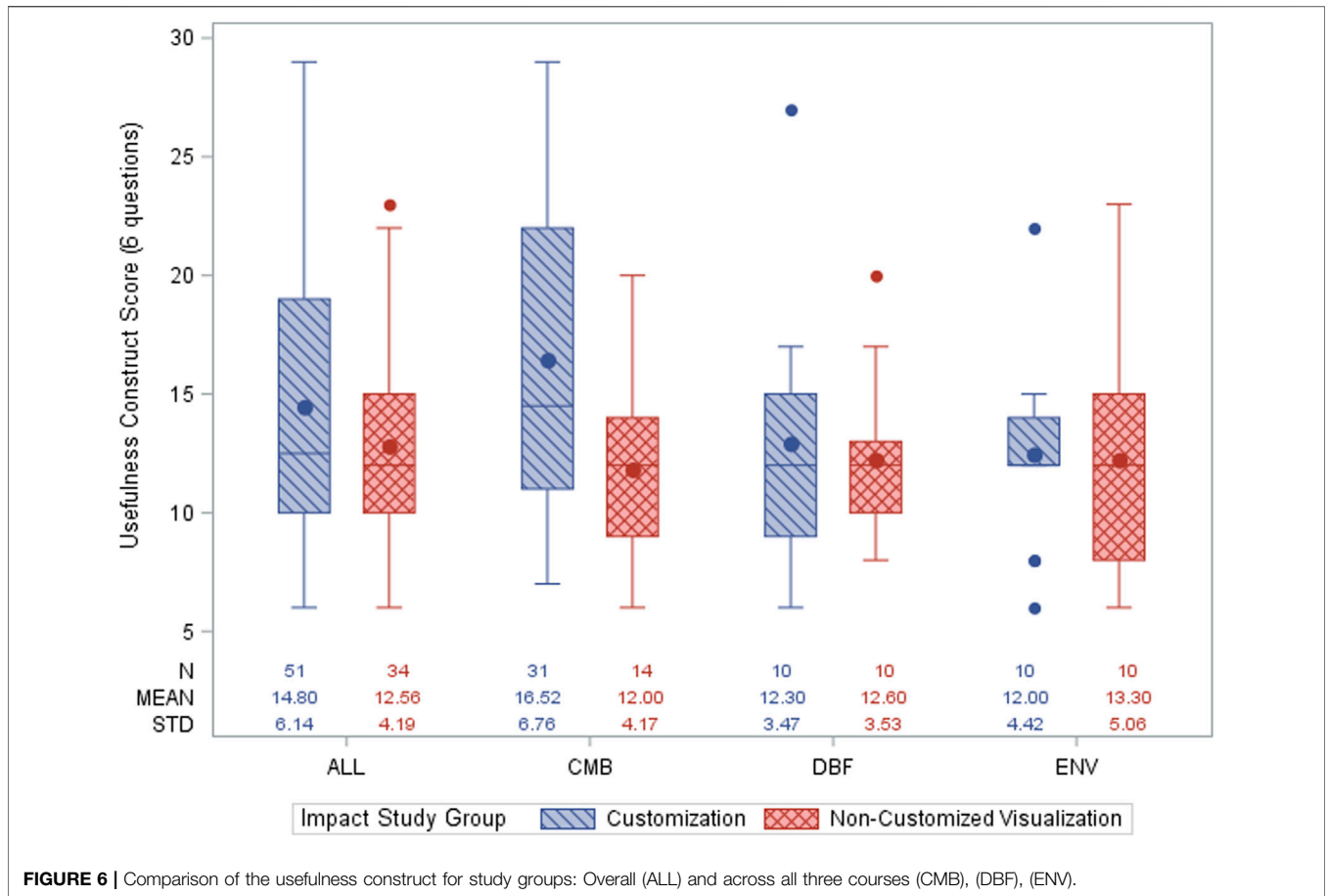


FIGURE 6 | Comparison of the usefulness construct for study groups: Overall (ALL) and across all three courses (CMB), (DBF), (ENV).

TABLE 2 | Summary of p-values.

Construct	ALL	CMB	DBF	ENV
Content	0.22	0.25	0.66	0.92
Confidence	0.82	0.79	0.97	0.5
Interest	0.17	0.14	0.45	0.82
Usefulness	0.048	0.009	0.89	0.58

components of data science. The knowledge of querying and designing databases is an empowering skill set that crosses discipline boundaries.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: The surveys and datasets analyzed for this study can be found in the Customization Impact Study component (<https://osf.io/npbya>) of the Open Science Framework site for the Databases for Many Majors project: <https://osf.io/qgepn>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Arizona State University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SWD, DG, and JB were responsible for writing up the article. JB and SWD formulated the study, and JB analyzed and interpreted the data. DG, SC, and BB customized the visualizations and participated in the customization study. KK and JO customized the visualizations. All authors contributed to a review of the article.

FUNDING

This work was supported in part by the National Science Foundation under Grant DUE-1431848, DUE-1431661, DUE-0941584, DUE-0941401.

REFERENCES

- Ben-Ari, M., Bednarik, R., Ben-Bassat Levy, R., Ebel, G., Moreno, A., Myller, N., et al. (2011). A Decade of Research and Development on Program Animation: The Jeliot Experience. *J. Vis. Languages Comput.* 22 (5), 375–384. doi:10.1016/j.jvlc.2011.04.0044
- Broatch, J. E., Dietrich, S., and Goelman, D. (2019). Introducing Data Science Techniques by Connecting Database Concepts and Dplyr. *J. Stat. Educ.* 27, 147–153. doi:10.1080/10691898.2019.1647768
- Byrne, M. D., Catrambone, R., and Tasko, J. T. (1999). Evaluating Animations as Student Aids in Learning Computer Algorithms. *Comput. Educ.* 33 (4), 253–278. doi:10.1016/S0360-1315(99)00023-8
- Cromley, J. G., Perez, T., and Kaplan, A. (2016). Undergraduate STEM Achievement and Retention: Cognitive, Motivational, and Institutional Factors and Solutions. *Pol. Insights Behav. Brain Sci.* 3 (1), 4–11. doi:10.1177/2372732215622648
- DePaolo, C. A., and Wilkinson, K. (2014). Recurrent Online Quizzes: Ubiquitous Tools for Promoting Student Presence, Participation and Performance. *Interdiscip. J. E-Learning Learn. Objects* 10, 75–91. Available at: <http://www.ijello.org/Volume10/IJELLOv10p075-091DePaolo0900.pdf>. doi:10.28945/2057
- Dietrich, S. W., Goelman, D., Borrer, C. M., and Crook, S. M. (2015). An Animated Introduction to Relational Databases for Many Majors. *IEEE Trans. Educ.* 58 (2), 81–89. doi:10.1109/TE.2014.2326834
- Dietrich, S. W., Goelman, D., Broatch, J., Crook, S. M., Ball, B., Kobojeck, K., et al. (2020). Using Formative Assessment for Improving Pedagogy. *ACM Inroads* 11 (December 2020), 427–434. doi:10.1145/3430766
- Dietrich, S. W., and Goelman, D. (2021). Databases for Many Majors. Available at: <http://databasesmanymajors.faculty.asu.edu> (Accessed August 20, 2021).
- Dietrich, S. W., and Goelman, D. (2017). “Formative Self-Assessment for Customizable Database Visualizations: Checkpoints for Learning,” in Paper presented at 2017 ASEE Annual Conference & Exposition, Jun 2017 (Columbus, Ohio. Available at: <https://peer.asee.org/27853>).
- Goelman, D. (2008). “Databases, Non-majors and Collaborative Learning,” in Proceedings of the 13th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE '08) (New York, NY, USA: Association for Computing Machinery), 27–31. doi:10.1145/1384271.1384281
- Heersink, D., and Moskal, B. M. (2010). “Measuring High School Students’ Attitudes toward Computing,” in Proceedings of the 41st ACM Technical Symposium on Computer Science Education (SIGCSE '10) (New York, NY, USA: Association for Computing Machinery), 446–450. doi:10.1145/1734263.1734413
- Hundhausen, C. D., Douglas, S. A., and Tasko, J. T. (2002). A Meta-Study of Algorithm Visualization Effectiveness. *J. Vis. Languages Comput.* 13, 259–290. doi:10.1006/jvlc.2002.0237
- Hundhausen, C. D., Patterson, R., Lee Brown, J., and Farley, S. (2004). “The Effects of Algorithm Visualizations with Storylines on Retention: An Experimental Study,” in 2004 IEEE Symposium on Visual Languages - Human Centric Computing, 226–228. doi:10.1109/VLHCC.2004.55
- Jones, B. D. (2009). Motivating Students to Engage in Learning: The MUSIC Model of Academic Motivation. *Int. J. Teach. Learn. Higher Educ.* 21 (2), 272–285.
- Koorsse, M., Olivier, W., and Greyling, J. (2014). School Evaluation. *J. Inf. Technol. Educ. Innov. Pract.* 13, 89–117. doi:10.1787/9789264117044-8-en
- Lave, J., and Etienne Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- McMillan, J. M., and Hearn, J. (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. *Educ. Horizons* 87, 140–149. Available at: <http://www.jstor.org/stable/42923742>.
- Naps, T. L., Rößling, G., Almstrum, V., Dann, W., Fleischer, R., Hundhausen, C., et al. (20032002). Exploring the Role of Visualization and Engagement in Computer Science Education. *SIGCSE Bull.* 35, 2131–2152. doi:10.1145/782941.782998
- R Core Team (2017). R: A Language and Environment for Statistical Computing. Available at: <https://www.R-project.org/> (Retrieved January, 2021).
- Sorva, J., Ville Karavirta, V., and Malmi, L. (2013). A Review of Generic Program Visualization Systems for Introductory Programming Education. *ACM Trans. Comput. Educ.* 13, 4Article 15, 2013. (Nov. 2013), 64. doi:10.1145/2490822
- Spanjers, I. A. E., van Gog, T., Wouters, P., and van Merriënboer, J. J. G. (2012). Explaining the Segmentation Effect in Learning from Animations: The Role of Pausing and Temporal Cueing. *Comput. Educ.* 59, 274–280. doi:10.1016/j.compedu.2011.12.024
- Urquiza-Fuentes, J., and Velázquez-Iturbide, J. Á. (2009). A Survey of Successful Evaluations of Program Visualization and Algorithm Animation Systems. *ACM Trans. Comput. Educ.* 9 (2), 1–21. Article 9 (June 2009). doi:10.1145/1538234.1538236
- Wetzel, C. D., Radtke, P. H., and Stern, H. W. (1994). *Instructional Effectiveness of Video media*. Hillsdale, NJ, USA: Lawrence Erlbaum Associates. Science, Engineering and Humanities and Social Sciences references.
- Wickham, H., François, R., Henry, L., and Müller, K. (2017). Dplyr: A Grammar of Data Manipulation. Available at: <https://CRAN.R-project.org/package=dplyr> (Retrieved January, 2021).
- World Wide Web Consortium (2021). Web Content Accessibility Guidelines (WCAG) 2.0. Available at: <https://www.w3.org/TR/WCAG20/> (Accessed August 20, 2021).

Author Disclaimer: Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Dietrich, Goelman, Broatch, Crook, Ball, Kobojeck and Ortiz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.