**Supplemental Information**

Table S1. Catalog of paper aquifer models. Models can be accessed at [http://aquifer.geology.buffalo.edu](http://aquifer.geology.buffalo.edu/)

|  |  |
| --- | --- |
| The Basics | Hydraulic Head |
| [Example Aquifer – The basics](http://aquifer.geology.buffalo.edu/index.php/2019/10/27/example-foldable-aquifer-model/) | [Hydraulic Head Across Aquifers (Easy)](http://aquifer.geology.buffalo.edu/index.php/2020/02/21/hydraulic-head-across-aquifers-easy/) |
| [Porosity and Grain Packing](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/porosity-and-grain-packing/) | [Hydraulic](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/hydraulic-head-across-aquifers/) [Head Across Aquifers (Hard)](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/hydraulic-head-across-aquifers/) |
| [Groundwater Storage](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/groundwater-storage/) | [Three Point Problem](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/three-point-problem/) |
| Hydraulic Conductivity | Well Hydraulics |
| [Heterogeneity and Wells](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/heterogeneity-and-wells/) | [Quantifying Hydraulic Conductivity: Slug Test](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/quantifying-hydraulic-conductivity-slug-test/) |
| [Effective Hydraulic Conductivity: The basics](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/effective-hydraulic-conductivity-the-basics/) | [Steady State Pumping: Thiem Equation](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/steady-state-pumping-thiem-equation/) |
| [Classic Transmissivity Problem](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/classic-transmissivity-problem/) | [Pumping near a constant head boundary](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/pumping-near-a-constant-head-boundary/) |
| Darcy’s Law | [Pumping near a no flow boundary](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/pumping-near-a-no-flow-boundary/) |
| [Darcy Columns](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/darcy-columns/) | [Image well and max pumping rate](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/image-well-and-max-pumping-rate/) |
| [Average Linear Velocity of Groundwater](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/average-linear-velocity-of-groundwater/) | [Impact on wetland from groundwater withdrawal](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/impact-on-wetland-from-groundwater-withdrawal/) |
| [Darcy’s Law and Radial Coordinates](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/darcys-law-and-radial-coordinates/) |  |
| [Groundwater Flow Between Wells](http://aquifer.geology.buffalo.edu/index.php/2019/11/27/groundwater-flow-between-wells/) |  |

Table S2. Anonymous Learning Survey Results for Darcy’s Law and Storage from Austin Peay State University. These tables summarize the most understood and least understood concepts as reported by students through a survey at the end of class that consisted of two questions: “What did you understand the most today?” and “What did you understand the least today?”. Note n values differ since these are completed at the end of class for those in attendance. Differences in most and least understood counts for pre-paper models are due to one question being left blank.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Pre-model** | **Topic** | **Count** | **%** |  | **Post-model** | **Topic** | **Count** | **%** |
| ***Darcy's Law*** | **Darcy's Law** | 6 | 18% |  | ***Darcy's Law*** | **Darcy's Law** | 4 | 16% |
| ***Most Understood*** | **Concepts Related to Darcy's Law** | 0 | 0% |  | ***Most Understood*** | **Concepts Related to Darcy's Law** | 0 | 0% |
|  | **Other** | 28 | 82% |  |  | **Other** | 21 | 84% |
|  |  | **34** |  |  |  |  | **25** |  |
|  |  |  |  |  |  |  |  |  |
| **Pre-model** | **Topic** | **Count** | **%** |  | **Post-model** | **Topic** | **Count** | **%** |
| ***Darcy's Law*** | **Darcy's Law** | 12 | 36% |  | ***Darcy's Law*** | **Darcy's Law** | 10 | 40% |
| ***Least Understood*** | **Concepts Related to Darcy's Law** | 4 | 12% |  | ***Least Understood*** | **Concepts Related to Darcy's Law** | 8 | 32% |
|  | **Other** | 17 | 52% |  |  | **Other** | 7 | 28% |
|  |  | **33** |  |  |  |  | **25** |  |
|  |  |  |  |  |  |  |  |  |
| **Pre-model** | **Topic** | **Count** | **%** |  | **Post-model** | **Topic** | **Count** | **%** |
| ***Storage*** | **S, Sy, Ss** | 6 | 18% |  | ***Storage*** | **S, Sy, Ss** | 5 | 20% |
| ***Most Understood*** | **Concepts Related to S, Sy, Ss** | 2 | 6% |  | ***Most Understood*** | **Concepts Related to S, Sy, Ss** | 1 | 4% |
|  | **Balloon Activity** | 8 | 24% |  |  | **Paper Model Act.** | 1 | 4% |
|  | **Other** | 18 | 53% |  |  | **Other** | 18 | 72% |
|  |  | **34** |  |  |  |  | **25** |  |
|  |  |  |  |  |  |  |  |  |
| **Pre-model** | **Topic** | **Count** | **%** |  | **Post-model** | **Topic** | **Count** | **%** |
| ***Storage*** | **S, Sy, Ss** | 6 | 17% |  | ***Storage*** | **S, Sy, Ss** | 4 | 16% |
| ***Least Understood*** | **Concepts Related to S, Sy, Ss** | 8 | 23% |  | ***Least Understood*** | **Concepts Related to S, Sy, Ss** | 4 | 16% |
|  | **Balloon Activity** | 4 | 11% |  |  | **Paper Model Act.** | 0 | 0% |
|  | **Other** | 17 | 49% |  |  | **Other** | 17 | 68% |
|  |  | **35** |  |  |  |  | **25** |  |

Table S3. Likert Scale rankings for “On a scale of 1 (not at all) to 6 (a lot) rate the extent to which you've developed your understanding of Darcy's Law.” Note n values differ as some students did not complete these quizzes, which were given for a completion grade.

|  |  |  |  |
| --- | --- | --- | --- |
| **Pre-model** | **Rankings** | **Quiz 10b** | **Quiz 24** |
| ***Darcy's Law*** | **1** | 0 | 0% | 0 | 0% |
|  | **2** | 2 | 11% | 0 | 0% |
|  | **3** | 3 | 16% | 0 | 0% |
|  | **4** | 7 | 37% | 7 | 24% |
|  | **5** | 6 | 32% | 12 | 41% |
|  | **6** | 1 | 5% | 10 | 34% |
|  |  | **19** |  | **29** |  |
| **Post-model** | **Rankings** | **Quiz 10b** | **Quiz 24** |
| ***Darcy's Law*** | **1** | 0 | 0% | 0 | 0% |
|  | **2** | 0 | 0% | 0 | 0% |
|  | **3** | 3 | 11% | 0 | 0% |
|  | **4** | 13 | 48% | 6 | 24% |
|  | **5** | 6 | 22% | 9 | 36% |
|  | **6** | 5 | 19% | 10 | 40% |
|  |  | **27** |  | **25** |  |

Table S4. Percentage of students scoring in the “High”, “Medium”, and “Low” ranges pre- and post-model use for homework, quiz, and exam questions. “High” was defined as 90-100%, “Medium” as 65-89%, and “Low” as < 65%. Since each assessment was worth varying numbers of points (3-8) these ranges were chosen as they worked with all possible score outcomes with each assessment. Note n value differs for the pre-model calculation quiz as one iteration of the course did not take a calculation quiz.

|  |  |  |
| --- | --- | --- |
|  | **Darcy's Law Homework** | **Darcy's Law Calculation Quiz** |
|  | ***Pre*** | ***Post*** | ***Pre*** | ***Post*** |
| ***"High"*** | 43% | 48% | 35% | 52% |
| ***"Medium"*** | 26% | 45% | 13% | 17% |
| ***"Low"*** | 31% | 7% | 52% | 31% |
|  |  |  |  |  |
|  | **Darcy's Law Midterm Exam** | **Darcy's Law Final Exam** |
|  | ***Pre*** | ***Post*** | ***Pre*** | ***Post*** |
| ***"High"*** | 14% | 14% | 20% | 41% |
| ***"Medium"*** | 29% | 31% | 46% | 45% |
| ***"Low"*** | 57% | 55% | 34% | 14% |
|  |  |  |  |  |
|  | **Storage Homework** |  |  |
|  | ***Pre*** | ***Post*** |  |  |
| ***"High"*** | 14% | 17% |  |  |
| ***"Medium"*** | 26% | 52% |  |  |
| ***"Low"*** | 60% | 31% |  |  |

**Details of Flipped Classroom Methodology & Results at Austin Peay State University**

The “Average Linear Velocity of Groundwater” model (Table S1) is utilized as an activity to help introduce Darcy’s Law. After building their models, students discuss the components of Darcy’s Law and identify them on the model. Then they calculate discharge and specific discharge of the gravel aquifer with a given distance between the wells. Finally, they are given effective porosity and asked to calculate average linear velocity. A few weeks later, the “Groundwater Storage” models (Table S1) are utilized as an activity to further understand the concept of storage and the differences in aquifers. After building their models, students calculate and compare differences in storage between confined and unconfined aquifers as described. The paper aquifer models’ activities both occur in the first half of the course, and while these two class days are the only ones that specifically utilize the models, the instructor encourages students to regularly use them to review concepts and when they work on homework assignments. Additionally, the instructor has models available during office hours and regularly talks through concepts with students while using these. Note that although the second half of the Spring 2020 course was taught remotely, which included the midterm and final exams, the activities occurred prior to this so the aquifer models were utilized in person during both the Spring 2020 and Fall 2021 courses.

As described in the Results & Discussion, data collection allowed for comparisons between the pre-paper model (Fall 2015, Spring 2017, and Fall 2018; n = 35) and post-paper model (Spring 2020 and Fall 2021; n= 29) iterations of the course. Students’ persistence, defined in this context as students’ motivation to complete all course assignments, was assessed through analyses of attempts of homework, quiz, and exam questions that specifically targeted the concepts of Darcy’s Law and Storage. However, generally all students attempted every question analyzed on exams and the online nature of the exams in Spring 2020 is potentially impacting this so only homework and quiz persistence will be discussed. Students’ learning was assessed through their scores on homework, quiz, and exam questions that specifically targeted the concepts of Darcy’s Law and Storage. Note that all questions and assignments analyzed were either identical or similar questions and assignments in all iterations of the course. Finally, students’ perceptions of their own learning were assessed through analyses of anonymous learning surveys at the end of each class period with a paper aquifer model activity, and specific questions from two online quizzes related to their understanding of the concepts of Darcy’s Law and Storage.

Students’ persistence for Darcy’s Law was measured through the two homework assignments (1-2 problems each) that immediately followed the class period in which the topic was introduced, and an online calculation quiz that occurred 1-2 weeks after this introduction. During the pre-paper model iterations, a mean of 10% of students did not submit either of the homework assignments, while all students in the post-paper model courses submitted both assignments. The online calculation quiz had a mean of 3% of students not attempting the quiz for both pre- and post-paper model iterations. For the pre-paper model, one of the three iterations of this course did not take a calculation quiz, so n = 23 for this quiz. However, all students were given an opportunity to redo the quiz and explain where they made errors originally to earn up to full credit. While the percentage of students who attempted this opportunity was nearly identical (pre- 30%, post-31%), the outcomes of these attempts differed. In the pre-paper model iterations, 71% of the students who attempted the redo raised their grade to a C or higher, and only 29% earned the full credit. In the post-paper model iterations, all of the students who attempted the redo raised their grade to a C or higher, and 78% earned the full credit. Although the assignments were returned and cannot be analyzed further, anecdotally the instructor noted this difference due to either a lack of explanation or only attempting to redo part of this multi-step calculation in the pre-paper model courses.

Students’ persistence for Storage was measured through the homework assignment that immediately followed the class period in which the topic was introduced. The percentage of students who did not attempt this homework assignment dropped from 20% pre-paper models to 7% post-paper models.

Students’ learning for Darcy’s Law was measured through the homework assignment that immediately followed the class period in which the topic was introduced, an online calculation quiz that occurred 1-2 weeks after this introduction, and questions from the midterm and final exams related to both Darcy’s Law concepts and calculations. Only one homework assignment was analyzed for learning since only half of the second assignment was related to Darcy’s Law, and only raw scores were available. The mean score for the homework increased slightly from pre-paper model (72%-all; 78%- attempted only) to post-paper model iterations (84%); however, the percentage of students who scored in the “low” range decreased (Table S4). An increase of more than a letter grade occurred in the online calculation quiz from pre-paper model (57%-all; 60%- attempted only) to post-paper model iterations (73%-all; 75%- attempted only), and again a decrease in “low” scores occurred. On the midterm exam, mean scores from Darcy’s Law calculations were the same (56%) and the percentage of students with “low” scores was similar; however, mean scores on fill-in-the-blank and true/false questions on the related concept of terms and units for hydraulic conductivity increased from 59% (pre-paper models) to 74% (post-paper models) and percentage of “low” scores decreased. Additionally, mean scores from three multiple choice final exam questions related to Darcy’s Law and requiring calculations also increased from 60% (pre-paper models) to 76% (post-paper models) while “low” score percentage decreased.

Students’ learning for Storage was measured through the homework assignment that immediately followed the class period in which the topic was introduced and questions from the midterm and final exams related to concepts of Storage. The mean score for the homework increased from pre-paper model (55%-all; 69%- attempted only) to post-paper model iterations (72%-all; 77%- attempted only) and the percentage of students who scored in the “low” range decreased (Table S4). Mean scores were similar for fill-in-the-blank and short answer midterm exam questions (pre- 52%, post- 57%) and multiple-choice and matching final exam questions (pre- 62%, post- 60%) related to Storage concepts.

Students’ perceptions of their own learning for both Darcy’s Law and Storage were measured through Anonymous Learning Surveys at the end of the class day in the post-paper model iterations including the paper model activity associated with the topic and various questions from online quizzes. Specifically, these included Likert scale questions on understanding of Darcy’s Law approximately one week after the activity and then again at the end of the semester, and a least and most understood topic selected from the nine major topics covered in the course. Storage was not specifically listed but included in the broad topic of “Hydraulic parameters, such as porosity, hydraulic conductivity, and storage” and included in a Likert scale question of understanding at the end of the semester, as well as one of the listed topics. Results from the Anonymous Learning Surveys (Table S2) indicated a minimal change in understanding of storage, with 24% most understood as storage or related to storage and a decrease from 40% pre-paper model to 32% post-paper model for least understood as storage or related to storage. Similarly, Darcy’s Law had minimal change as most (pre- 18%, post- 16%) and least (pre- 36%, post- 40%) understood. However, there was an increase in the least understood related to Darcy’s Law from 12% pre-paper models to 32% post-paper models, with the majority of these mentioning either specific discharge or average linear groundwater velocity. Results from the Likert scale questions for Darcy’s Law (Table S3) also indicated similarities in learning perception pre- and post-paper models, with minimal increases the week after the activity. The hydraulic parameters questions, which included storage, and were only included at the end of the semester quiz, also had similar results pre- and post-paper model activities. Minimal increases were seen in the most understood selections for Darcy’s Law (pre- 24%, post- 32%), and in both instances were 0% selected for the least understood