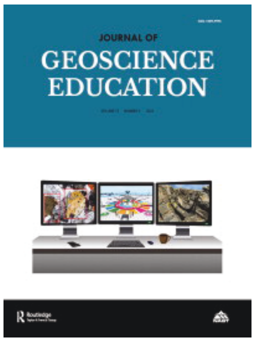

Journal of Geoscience Education

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujge20

An evaluation of instructional strategies for improving student understanding of the elastic rebound theory of earthquakes with spatial visualization

Michael R. Brudzinski, Thomas F. Shipley & Joy Ham

To cite this article: Michael R. Brudzinski, Thomas F. Shipley & Joy Ham (08 Apr 2024): An evaluation of instructional strategies for improving student understanding of the elastic rebound theory of earthquakes with spatial visualization, Journal of Geoscience Education, DOI: 10.1080/10899995.2024.2335844

To link to this article: https://doi.org/10.1080/10899995.2024.2335844

JOURNAL OF GEOSCIENCE EDUCATION
https://doi.org/10.1080/10899995.2024.2335844

Routledge
Taylor & Francis Group

An evaluation of instructional strategies for improving student understanding of the elastic rebound theory of earthquakes with spatial visualization

Michael R. Brudzinski[a], Thomas F. Shipley[b] and Joy Ham[b]

[a]Department of Geology and Environmental Earth Science, Miami University, Oxford, Ohio, USA; [b]Department of Psychology and Neuroscience, Temple University, Philadelphia, Pennsylvania, USA

ABSTRACT
Recent studies have identified an incomplete student understanding of how elastic rebound causes earthquakes. We hypothesized that realistic imaging of spatial patterns in ground motions over the course of the earthquake cycle would improve student understanding. Incorporating spatial change information in the form of both motion vectors and before-during-after contrasts should require most students to change an existing mental model or develop a new model. Using a quasi-experimental design, we developed instructional interventions for presenting variations in ground motion, including map views of fence bending and GPS velocity vectors. We measured the impact on student performance based on assignment questions related to the ground motion at different points in the earthquake cycle following several interventions in four undergraduate courses from introductory to upper level over 4 years. The first round of study was a free-response format and then multiple-choice answers were created from the most common answers, including new "worked example" questions inquiring about the reasons answers were correct or incorrect. We identified two key misconceptions based on student answer choices: (a) difficulty in recognizing velocity vector patterns when presented in a new reference frame, and (b) difficulty in reasoning that the fault must be locked for the strain to accumulate and produce an earthquake. Our analysis indicates the largest performance increases occur with simple animations that demonstrate the bending, breaking, and rebending of a fence, along with associated GPS vectors, plotted successively in different reference frames. This suggests difficulties in understanding elastic rebounds can be mitigated when spatial patterns are presented in a context with repeated opportunities to make predictions combined with animations to support mental models that connect the spatial patterns with ground movement.

ARTICLE HISTORY
Received 26 May 2023
Revised 21 March 2024
Accepted 24 March 2024

KEYWORDS
Earthquakes; elastic rebound; deformation; undergraduate

Introduction

The overarching goal of geoscience education is to improve students' mental models of the Earth. Geoscience learning requires developing complex multidimensional mental models of how spatial patterns in the world reveal geological processes acting over a range of spatial and temporal scales. The approach to improving students' spatial reasoning skills proposed by Davatzes et al. (2018) considers learning as a system and has four important features: (a) students are more likely to improve an erroneous model when they can externalize their mental model and compare it to a correct model, (b) a strong test of a mental model is to have it make a prediction which can be compared to a correct answer, (c) the spatial difference between the prediction and the correct answer provides spatial feedback to the student about what was wrong with their prediction, which can guide improvement of their model, and (d) a cycle of prediction and spatial accommodation (change in the mental model) can yield an improved mental model through refinement or reconstruction. Awareness and recognition of the spatial error causes students to change the model so that the next prediction will be more accurate. This aligns with general delta-rule models of learning (e.g., Rescorla & Wagner, 1972) where learning occurs when there is a mismatch between expected and actual outcomes, as well as the idea of accommodation from Piaget and many other modern models of learning (e.g., Dole & Sinatra, 1998). However, the spatial approach differs in two important ways:

1. Spatial feedback goes beyond binary feedback (right/wrong) to guide how to improve an answer. Feedback that contains spatial information (e.g., location, direction, or shape) provides specific guidance about how to correct a model to reduce error (e.g., the magnitude and direction of change). With this approach, spatial feedback is derived by comparing a model prediction to a correct answer. This principle comes out of research on the use of spatial feedback by Gagnier et al. (2017) to improve learning using block diagrams—a common and difficult

CONTACT Michael R. Brudzinski brudzimr@MiamiOH.edu Department of Geology and Environmental Earth Science, Miami University, 250 S. Patterson, Oxford, OH 45056, USA.

3D diagram (Kali & Orion, 1996)—and work by Resnick, Davatzes, et al. (2017a, 2017b) and Resnick, Newcombe, et al. (2017) to improve learning using spatial diagrams of large magnitudes of time, an area where students often make significant errors (Libarkin et al., 2007).

2. Spatial accommodation is the constructing and reconstructing of mental models to accurately incorporate spatial information. Spatial accommodation is the complement to spatial feedback in that it is the adjustment needed for the mental model to align with the feedback received. Feedback may cause a small adjustment to an existing model, significant reconstruction of an existing model, or the creation of a new mental model when none exists. Accommodation in these three cases can require different types of support for students (Kastens et al., 2009; Lombardi et al., 2021).

In this study, we focus on mental models of earthquakes, as there are a variety of preconceived notions about these hazardous events and their societal impact (e.g., Aydin & Coskun, 2011; Kirikkaya et al., 2011; Santos-Reyes et al., 2017; Sözen, 2019; Tsai, 2001). In particular, we focus on the concept of elastic rebound, as recent studies have identified that student understanding of elastic rebound and how it causes earthquakes is often incomplete (Dolphin & Benoit, 2016; Hubenthal, 2018a). A variety of different models of elastic rebound have been proposed for instruction due to the difficulty in creating a visualization of the actual process that operates on faults due to the wide range of scales involved (Hubenthal, 2018b; Hubenthal et al., 2011; LaDue et al., 2018). For example, rocks near a fault are bent, on the order of a few meters displacement near the fault relative to rocks far from the fault, but this bending can occur over tens of kilometers of distance. In addition, the bending near active faults builds up gradually over tens to hundreds of years, but when an earthquake occurs, the rocks move in a matter of seconds. People typically see only the results of an earthquake, such as the offset of a fence built across a fault, but not the gradual motion that precedes such an earthquake. Students may learn how to characterize faults based on their offset but have a harder time learning what leads to those offsets, and hence what is the true physical cause of earthquakes.

The increasing precision of geophysical instruments has enabled geoscientists to measure gradual motions along faults in greater detail. One of the most common strategies is to attach a high precision GPS receiver to a rock exposure and then record its daily position over many years using satellite signals to establish its geolocation. When a network of these GPS stations is installed around a fault, the process of rock bending can be observed in the variations of ground motion from near the fault to far away. We hypothesized that realistic imaging of spatial patterns in ground motions over the course of the earthquake cycle can help students understand the elastic rebound concept through visualization of the movement (Brudzinski, 2018). However, GPS motions are not typically used in instruction with novices, and the lack of ability to visually perceive this type of data relative to a bent fence could limit its usefulness. How the spatial patterns in GPS-derived ground motions could be used to help students visualize the elastic rebound concept has not been investigated. Hence, we sought to present variations in ground motion near a fault at multiple points in the elastic rebound process to evaluate whether the ground motion illustrations support students developing mental models that incorporate the key aspects of elastic rebound: friction on the fault creates locking, rocks bending to accumulate elastic energy, and the earthquake causing the rocks near the fault to catch up with the motions far from the fault.

One challenge with learning about Earth motions around faults is that there is no convention about what frame of reference to use for GPS data. Different reference frames are used frequently by experts in tectonics studies and with GPS data in particular (e.g., Cox & Hart, 1991; Freymueller, 2021). Early work in psychology on reasoning about frames of reference has found that people can flexibly adopt alternative frames of reference (e.g., imagining themselves moving through a scene or the objects and surfaces of a scene moving around themselves—Hegarty & Waller, 2004). Although multiple frames of reference are employed by the mind (cf. Wade, 1996 for an overview of multiple frames of reference in vision), some frames are more likely to be adopted than others, and there are individual differences in which frame of reference individuals prefer (Liben, 1991). Here multiple opportunities to make predictions and get feedback in different geometries and frames of reference were provided to students. The repetition was designed to offer an opportunity to develop flexibility in using different frames of reference for motion.

Research question, hypothesis, and study design

Considering that recent studies have identified incomplete student understanding of how elastic rebound causes earthquakes (Dolphin & Benoit, 2016; Hubenthal, 2018a), the specific research question this study seeks to address is: How can we support students' spatial reasoning to improve their understanding of the elastic rebound concept and its relationship to earthquakes? Our hypothesis in response to this question is: Spatial feedback accompanied by realistic visualizations of the spatial patterns in GPS-derived ground motions over the course of the earthquake cycle will improve student understanding of the elastic rebound concept and the processes that lead to earthquakes. We chose to use a quasi-experimental design seeking to measure the impact of different instructional interventions on student performance on assignment questions related to the ground motion at different points in the earthquake cycle. To help support the quantitative approach, we collected data from many questions in several courses over several years to ensure a large enough sample size to analyze the results statistically. We collected data from the novice, intermediate, and advanced levels of the undergraduate curriculum at an R2 institution in the Midwest and chose to take a cross-sectional view of the student performance instead of a longitudinal view as we found few students with multiple exposures that would allow us to evaluate their performance after receiving multiple forms of intervention.

We designed instructional interventions and assignment-based assessments to evaluate (a) how well students understand the relationship between pre-earthquake deformation and the eventual earthquake fault motion and (b) how well students can flexibly reason about ground movement from a variety of reference frames. To help students develop a model of how earthquake motions related to pre-earthquake ground motions, we asked students to reason about how the ground would move before, during, and after a quake, allowing them to make predictions before we provided feedback. Students were given multiple opportunities to make predictions and get feedback. To help students develop flexibility with their mental model, we varied the fault geometry, the frame of reference for the motions, and whether students were given ground displacement or velocity information. The frame of reference was varied by adjusting where ground motion was zero relative to the fault.

Methods

Participants

We investigated the impact of interventions on student learning and understanding of elastic rebound in four courses at an R2 institution in the Midwest: introductory geology (Intro Level), solid earth geohazards (Early Majors), geophysics (Upper Level), and seismology (Upper Level), similar to Brudzinski et al. (2019). All four courses were taught by the first author over three years, with the Intro Level courses taught three times, Early Majors courses taught four times, and Upper Level courses taught three times (Figure 1). The average enrollments were 39 at the Intro Level, 30 for the Early Majors, and 13 for the Upper Level. The Intro Level courses were taught fully online and the others were held in person, and all occurred before the pandemic. We did not collect information on a student's major within the assignment, but these can be estimated from departmental data: In the Intro Level courses, which fulfilled the physical science requirement of the liberal education plan, ~1% of the students were undergraduate Geology and Environmental Earth Science majors, and were spread across

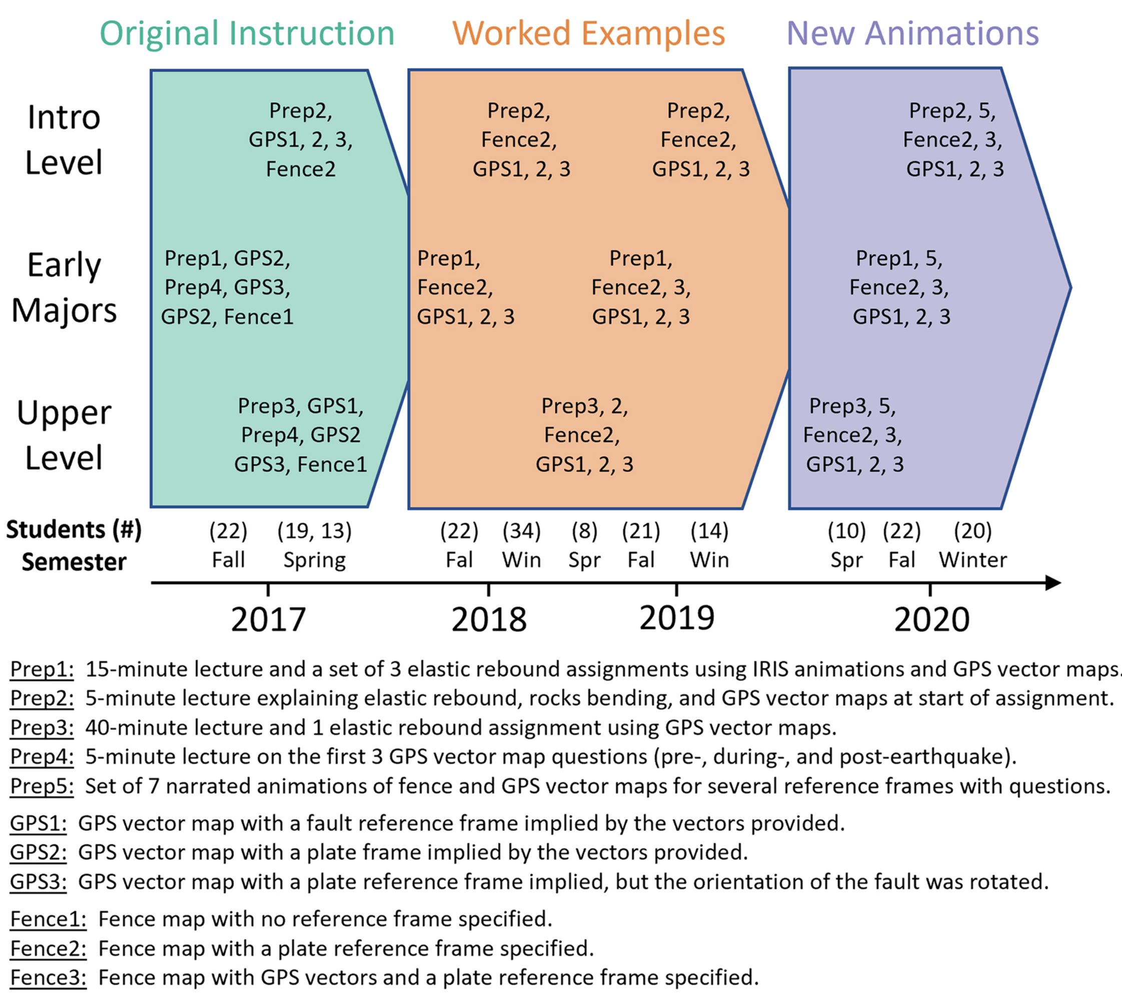

Figure 1. Summary of courses, assignments, and instructional interventions. Each block of text in the colored instruction type shows the order of instruction and assignment questions for each course offering. For Early Majors and Upper Level there were two instruction components each followed by assignments in the open answer Original Instruction. Numbers in parenthesis are the number of students participating in the study. Time axis shows when the course offerings occurred, with two occurring in Spring 2017.

all four years of undergraduate training. In the Early Majors courses, ~90% of the students were undergraduate Geology and Environmental Earth Science majors (primarily in their second year of training). In the Upper Level courses, 66% of the students were undergraduate majors and 33% were first- or second-year graduate students. In all, 67 students from the Intro Level, 87 students from the Early Majors, and 31 students from the Upper Level agreed to participate in the study.

Assignment

The assignment was constructed to utilize two types of map (overhead) view spatial representation of the elastic rebound patterns around a right-lateral, strike-slip fault at different times of the earthquake cycle: a fence that bends and breaks over time and GPS vectors showing the surface velocity at a given time. Earlier versions started with the GPS vectors in the assignment (Figure 1), but anecdotal feedback from students indicated that the fence diagrams were important to understand the spatial patterns implied by the GPS vectors. We have interpreted this as reflecting the challenge of using only discrete vectors to understand the spatial gradient in motion around a fault. The full version of the assignment developed over the course of this study can be found on SERC (Brudzinski, 2023).

Fence map

Figure 2 shows how the fence map is designed to illustrate simple fault motion, and the small arrows on either side of the fault indicate one side is moving in the opposite direction compared to the other side of the fault. When given a diagram that shows a red and blue fence immediately after they were built across a fault, students are asked to predict the expected pattern of how the fence will change shape at three points of time in the future: 100 years after the fence was built but before an earthquake, immediately after an earthquake, and 100 years after an earthquake. Correct answers for these three points in time are shown in Figure 2(a–c). These questions were designed to help students consider how rocks near a fault bend and then rebound as a key part of the earthquake process over time. Asking students to make specific predictions and then receive spatial information on how their choice was right or wrong was a key implementation of the spatial feedback approach.

During the first two versions of this assignment (Figure 1), students were asked to sketch their predictions. Answers were scored by the instructor (Brudzinski). Once it was clear that there was a consistent set of incorrect answers, these questions were designed for students to select from multiple choices devised from the sketches (Figure 2). We noted that many of the incorrect answers would have been the appropriate spatial pattern for a different time frame. The multiple-choice version of the assignment was implemented in the Moodle learning management system and followed the active-learning strategy of Sit and Brudzinski (2017). This automated grading strategy allows students to immediately identify if the answer they chose is correct, receive pre-written feedback tailored to the wrong answer choice, and then re-answer for diminishing partial credit. Some common incorrect answers for the pre-earthquake time frame are shown in Figure 2(d) (incorrect because the fault requires motion to be in opposite directions on opposite sides of the fault) and Figure 2(g) (incorrect because the equal motions on opposite sides of the fault indicate there is no friction on the fault), with answer frequency shown in Table 1. For during the earthquake, common incorrect answers are shown in Figure 2(e) (incorrect because the fence moved in the wrong direction relative to the arrows that show the fault motion) and Figure 2(h) (incorrect because the fault slips during an earthquake so each side snaps back to how they were before the rocks were bent). For the post-earthquake time frame, common incorrect answers are shown in Figure 2(f) (incorrect because the rocks near the fault should not move faster than those away from the fault) and Figure 2(i) (incorrect because the fault should be stuck again and causing rocks to bend again).

GPS velocity vectors

Figure 3 shows the same situations as Figure 2, but using GPS velocity vectors to illustrate the direction and speed of motion instead of the resulting displacement (e.g., the deformed fence). Scientists use sparse GPS velocity measurements such as these to estimate the rock bending and assess the strain accumulation process that leads to future earthquakes (Segall & Davis, 1997). In the first two implementations (Figure 1), students were given a map view of the same right-lateral, strike-slip fault and asked to draw the velocities expected at each GPS site for before, during, and after an earthquake. To help provide a reference point for their velocities, a velocity vector was provided for the furthest left GPS site. Multiple choice answers were derived from the student drawings, and subsequent implementations used these choices. Before the earthquake, the fault is "locked" due to friction and will not allow the two sides of the fault to move with different velocities. To still accommodate the opposite movements on the two sides further away from the fault, the velocities must change from near the fault to far away from the fault. The changes in velocity with increasing distance from the fault (Figure 3(a)) are what causes the fence (and rocks inside the Earth below it) to bend. During the earthquake, the two sides of the fault can finally move and "catch up" to the rest of the rocks on their side further away from the fault by snapping back to their original configuration. This rapid "unbending" of the fence (and rocks below it) in Figure 3(b) is what causes the earthquake. Many years after the earthquake, the fault would be stuck again and the rocks would be bending as they were before the earthquake, so the correct pattern (Figure 3(c)) is the same as Figure 3(a). We provided different scale bars to address the different rates of motion during the gradual strain accumulation and during the earthquake. After being asked to choose answers for each of the before, during, and after earthquake patterns in the assignment, students are asked again but for a different reference frame, where the right side of the map is held fixed such that both the fault

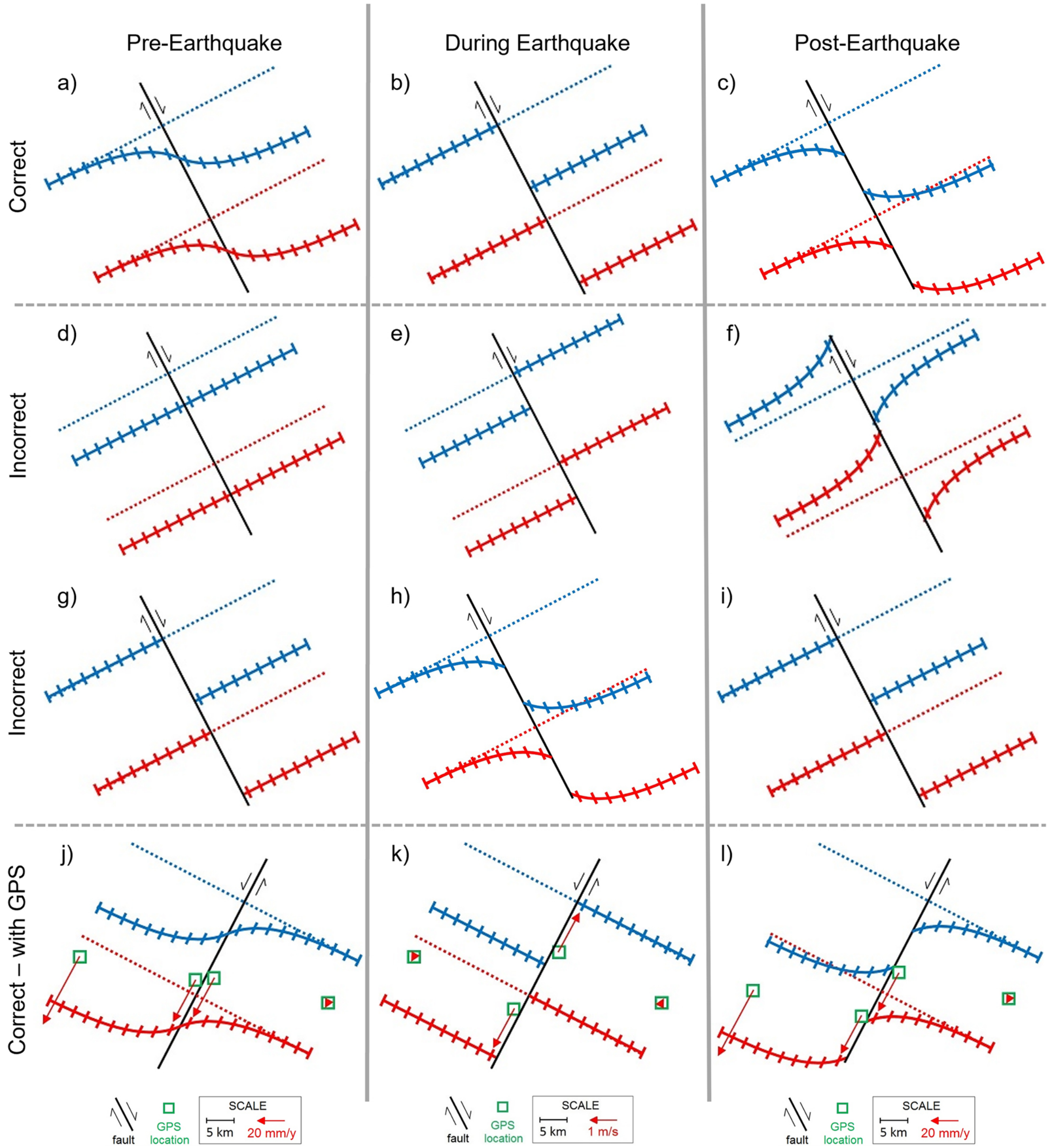

Figure 2. Map-view diagrams of a strike-slip fault illustrating elastic rebound motions over time. Black line is the surface trace of the fault, black arrows show the relative motion on either side of the fault. Two fences (blue and red) were built across a fault, and the dotted lines show the original position of the fences when built. (a) Deformed fence 100 years later, before a big earthquake when the rocks would have been bent near the fault. (b) Fence is broken during a big earthquake as sides of the fault finally slip past each other and rocks on either side un-bend. (c) Fence is deformed again 100 years after the earthquake as rocks near the fault are bent again. (d) Incorrect answer before the earthquake because the fault requires motion to be in opposite directions on opposite sides of the fault. (e) Incorrect answer during the earthquake because the fence moved in the wrong direction relative to the arrows that show the fault motion. (f) Incorrect answer after an earthquake because the rocks near the fault should not move faster than those away from the fault. (g,i) Are incorrect for before or after the earthquake because it is the correct answer for during an earthquake. (h) Is incorrect for during the earthquake because it is the correct answer for after the earthquake. (j–l) Show the correct answers for questions with the combined fence and GPS vectors before, during, and after an earthquake, respectively.

and the left side of the fault moves with large apparent velocities. Figure 3(g) illustrates the pattern before and after the earthquake. Finally, the students are asked to choose a configuration where the fault is rotated with one side of the map held fixed. Figure 3(j) illustrates this pattern before and after the earthquake.

In Figure 3 we also show common wrong answers students chose for the GPS vector portion of the assignment (Table 1), which were based on the early phase of the research and reflect a high probability of erroneous sketches. The choices indicate that some students have a mental model in which the Earth is moving at the same rate everywhere before an earthquake (e.g., Figure 3(d)). We interpret this choice to reflect a student's mental model where the fault is stuck, but it does not address the reality that there must be a variation in motion across the fault to build up the energy for an earthquake to occur. Other students chose Figure 3(e), but equal motions on opposite sides of the fault would indicate there is no friction on the fault. If an earthquake is going to occur, friction must be present to cause bending of the rocks that accumulate the elastic energy. During an earthquake, students also chose a version of this

Table 1. Proportion of students who selected correct and incorrect answers on their first attempt at questions for the different timing of the earthquake cycle and different visualizations.

Timing	Visualization	Correct	Incorrect 1	Incorrect 2	Incorrect 3
Pre	Fence	2a: 65%	2g: 10%	2f: 8%	2d: 6%
During	Fence	2b: 56%	2h: 26%	2f: 8%	2e: 7%
Post	Fence	2c: 47%	2i: 20%	2a: 11%	2f: 10%
Pre	GPS1	3a: 57%	3b: 20%	3e: 10%	3d: 9%
During	GPS1	3b: 69%	3f: 18%	3e: 6%	3a: 5%
Post	GPS1	3c: 83%	3e: 7%	3d: 5%	3b: 4%
Pre	GPS2	3g: 33%	3b: 32%	3e: 17%	3d: 11%
During	GPS2	3h: 82%	3a: 10%	3e: 5%	3d: 2%
Post	GPS2	3i: 60%	3b: 18%	3e: 11%	3d: 7%
Pre	GPS3	3j: 64%	3k: 20%	3eROT: 11%	3dROT: 5%
During	GPS3	3k: 84%	3j: 6%	3eROT: 6%	3dROT: 3%
Post	GPS3	3i: 75%	3b: 13%	3eROT: 7%	3dROT: 5%

ROT: rotated version of this figure was the answer choice for the GPS3 reference frame.
The correct and incorrect answer labels in the table refer to the images in Figures 2 and 3.

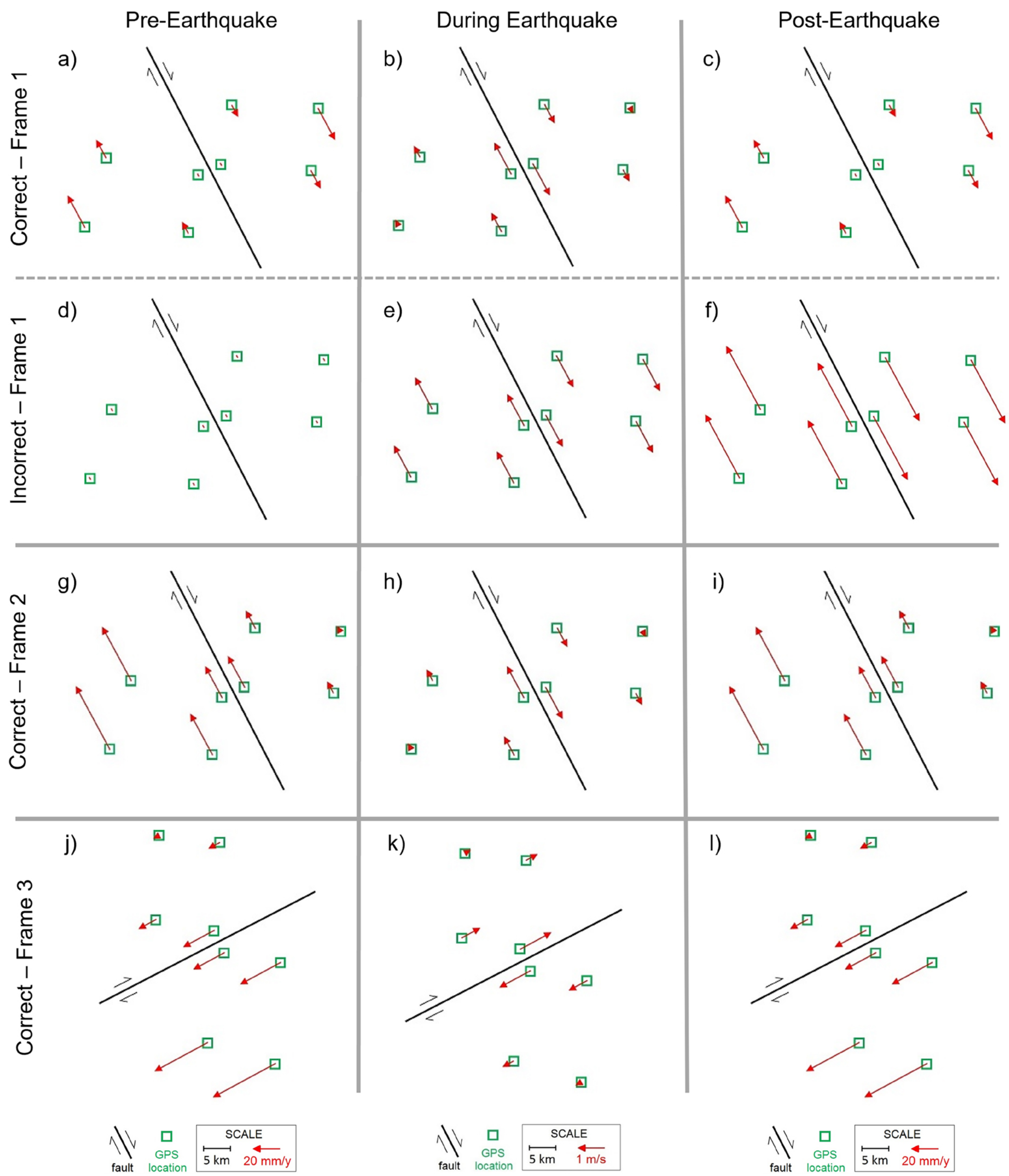

Figure 3. Map-view diagrams similar to Figure 2 showing motions (red arrows) that would be observed using GPS sites (squares). Correct GPS site motions for a fault-centric reference frame are shown for (a) before a large earthquake, (b) during the earthquake, and (c) after the earthquake (same as a). Note the different legends provided at the bottom for before/after *vs.* during the earthquake. (d–f) Common incorrect answers for before, during, or after an earthquake. Correct answers for a different reference frame (right side fixed) are shown for (g) before, (h) during, and (i) after an earthquake. Correct answers for a different fault configuration and the reference frame set so the upper side is fixed for (j) before, (k) during, and (l) after an earthquake.

but with larger arrows (Figure 3(f)), presumably to indicate the larger amount of motion during earthquakes.

The incorrect answers highlight some incomplete mental models common among students (Table 1). When students who chose Figure 3(e) were told it was incorrect and why, they often indicated that they did not realize the fault was stuck together before an earthquake. Students also struggled initially to choose Figure 2(a,c) presumably due to difficulty envisioning how the rocks are bending. This highlighted the limitations of this teaching approach (Figure 1), including that students could not watch the time progression when deciding on an answer, and that it required students to mentally connect the map to a broad physical area. Another key issue is that most students will not have prior knowledge of GPS motions or experience with sparse velocity vectors, which limits their ability to infer patterns from them. We noted that this type of data differs from the fence in that it shows velocity instead of displacement, instantaneous motion *vs.* a historical accumulation, and punctuated measurements as opposed to a continuous line segment. In addition, GPS velocities can be reported with various frames of reference to highlight motion relative to different points of view, but the lack of interactive maps makes it difficult for students to learn this mental manipulation skill. These potential challenges highlight that while GPS velocity patterns are a boon to geoscientists for measuring deformation and providing opportunities to utilize real data in the classroom, it is not straightforward how to use them in instruction. We decided to add a set of questions after the fence diagram question set that showed both the fence and the GPS station vectors to help students make connections between these data types (Figure 2(j–l)) (Fence 3 in Figure 1).

Worked examples

To help students improve their mental models, we employed the principle of spatial feedback in the context of the "worked examples" instructional strategy that arose from design science work in the GET-Spatial collaborative network (Davatzes et al., 2018; GET-Spatial, 2017). Worked examples have been shown to improve mathematics learning in situations where instructors can anticipate common errors (e.g., reporting the product of two negative numbers to be negative) and take students through the problem step by step so the student can deliberately contrast their answer with the correct answer. Davatzes et al. (2018) applied this approach to the principle of spatial feedback by having the students make a spatial guess at an answer and then providing students with the correct spatial location and an expert's reasoning for the answer. Thus, worked examples are useful when students are unsure how to adjust their model. The spatial error signal and the expert's reasoning could offer information about how and why to spatially adjust a mental model to make better future predictions. Worked examples are most likely to be effective when mental models are close to being correct and will benefit from spatial feedback that allows adjustment of a model.

We employed worked examples that asked students to consider the reasoning for wrong and then right answers after each of the situations and data types. After a question asking the students to choose the correct fence bending pattern before an earthquake, students were given Figure 2(d) as a wrong answer and asked why it was wrong (Answer: The fault requires motion to be in opposite directions on opposite sides of the fault), Figure 2(b) as a wrong answer and asked why it was wrong (Answer: The equal motions on opposite sides of the fault indicate there is no friction on the fault), and Figure 2(a) as a correct answer and asked why it was correct (Answer: Because friction on the fault does not allow the two sides to move in opposite directions). After the fence during earthquake question, students were given Figure 2(e) as a wrong answer (Answer: The fence moved in the wrong direction relative to the arrows that show the fault motion) and Figure 2(b) as a right answer (Answer: The fault slips during an earthquake so each side snaps back to how they were before the rocks were bent). After the fence post-earthquake question, students received Figure 2(f) as a wrong answer (Answer: The rocks near the fault should not move faster than those away from the fault) and Figure 2(c) as a right answer (Answer: The fault should be stuck again and causing rocks to bend again). Students were asked to consider a similar set of wrong and right answers after each of the GPS vector situations.

New animations

We also considered the possibility that some students would not have a preexisting mental model of the ground motions to allow flexible changes of frame of reference or one that was close enough to be improved with feedback from the static worked examples. To provide support for students who did not have a preexisting mental model, we specifically developed a set of narrated animations showing spatial patterns of bending and associated velocity vectors over time to illustrate the key spatial and temporal patterns associated with the elastic rebound process (Figure 4). This instructional approach continued to include the worked example questions during the assignment but used the animations at the beginning of the assignment. The full set of questions and animations as part of this intervention can be found on the SERC page for this assignment (Brudzinski, 2023).

Instructional interventions

A variety of instructional strategies have been employed depending on the course level. Elastic rebound is only briefly covered at the Intro Level. Students received a 5-min lecture on the key concepts, explaining elastic rebound and rock bending. They only received a brief description of GPS data at the beginning of the assessment as part of Prep2 (Figure 1).

The Early Majors students received 15 min of lecture instruction as well as three assignments focused on elastic rebound in Prep1. In this instruction, students were initially introduced to elastic rebound theory *via* a lecture about the 1906 San Francisco earthquake, including much of the information described by Dolphin (2018). The lecture focuses on the alternating phases of gradual rock bending over long time frames and rapid fault motion over seconds and

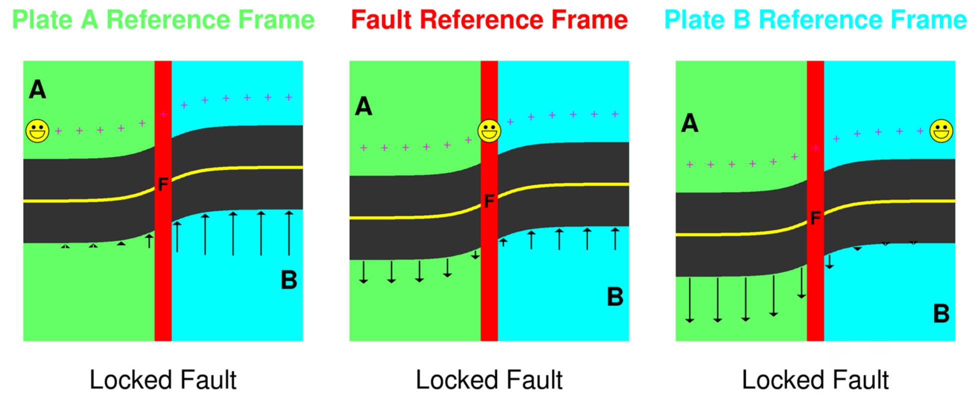

Figure 4. Still image from a narrated animation created to illustrate the spatial patterns of deformation and velocities of motion associated with elastic rebound before and during an earthquake. The animation specifically sought to highlight the invariant aspects of the relative motion patterns across reference frames by contrasting three different reference frames, where the emoji indicates the fixed reference point as plate A (left), the fault (middle), or plate B (right).

includes animations with accompanying explanations that help students visualize these concepts (IRIS, 2016). The pre-assignment lecture also introduces students to the concept of geodesy, where scientists measure changes in the Earth's surface over time with things like precise GPS instruments. Students then complete three assignments that incorporate the basic concept of elastic rebound and GPS data.

The Upper Level students received a longer, 40-min lecture on the same concepts as the Early Majors course but in more detail as part of Prep3. Prep3 also had a single, longer elastic rebound assignment using GPS data. This assignment integrates many of the aspects of the Early Majors assignments but asks students to work more with adjusting reference frames when considering GPS vector data. These first three instruction strategies were in place before the study.

To address the instructional goals, we developed two new interventions. The first (Prep4) was an additional 5-min lecture on the first three GPS questions (pre-EQ, during-EQ, post-EQ), providing an opportunity for students to be directly introduced to the fence bending versions of these situations. This was given to the Early Majors and Upper Level students in the Original Instruction stage of the study after the students had seen the first set of GPS questions (Figure 1). Based on the desire to support students without a mental model of the spatial patterns of elastic rebound, we created a set of seven narrated animations of a gradually bending and breaking fence and GPS vectors showing different reference frames over the earthquake cycle as well (Prep5). These animations were incorporated into the beginning of the assessment so we were able to include a few open-ended questions to ask students to reflect on what they are observing. The full set of instructional resources for each of these interventions can be found on the SERC page for this assignment (Brudzinski, 2023).

Statistical analyses

Our analysis focused on student performance on questions that asked the participants to identify the correct spatial pattern pre-/during/post-earthquakes for the four different visualizations (1 Fence Map, and 3 GPS reference frames) that were persistent throughout all of the instructional interventions. Longitudinal and hierarchical data were analyzed using a hierarchical (multilevel) linear modeling approach. The value of multilevel models is they account for hierarchically structured data where each participant that we observed is nested in their particular course offering, allowing examination of group-level effects (i.e., method of instruction) while accounting for random effects of the semester that each group-level contains (Flunger et al., 2021). The analysis essentially constructs two statistical models. One model that is based on main effects (e.g., course level) is compared to a second model that contains both main effects and interaction (e.g., how do the responses to questions differ in a course for different instruction types). If the interaction model can explain a significant amount of the variance in the data, we can conclude that instructional intervention had an influence, and we then will do a simple slopes moderation analysis to investigate the nature of the influence.

Following practice in this area for each analysis we report: (a) An estimate of the interaction model's fit to explain the data as an R^2, which reflects the proportion of the variability in the data the model accounts for; (b) A measure of how much variance each model accounts for using AIC (Akaike information criterion), which is a measure of the goodness of fit between data and model. We also report BIC (Boolean information criterion), which is a different measure of fit that penalizes greater numbers of free parameters in a model; (c) The likelihood that the interaction model accounts for more variance than the main effects model in the form of a chi-square test comparing AIC values for each mode; and (d) The statistical significance of the interaction in the model, which is judged by whether or not the slope term (b) differs from zero and is reported as a *t*-test. A negative b indicates improvement in the goodness of fit due to a reduction in variance. We conduct our follow-up simple slope moderation analyses if, within the interaction model, there is a significant interaction and the chi-square detects a significant reduction in AIC, but the reader will note in some cases BIC increases.

The simple slopes moderation analysis examines whether there is a difference in slopes between each level of the independent variables in the interaction. To test this, it applies a t-test to see if there is a significant difference between the slopes of a specific pair of cases (e.g. is there a difference between the effect of two instructions for a specific question). The measure differs from a familiar t-test that compares means in that it uses slopes to estimate differences and requires an estimated degree of freedom to apply the t-distribution. A positive β indicates improvement from the first to the second instruction.

A mixed-effects ANOVA would be unable to entirely differentiate between the effect of the instruction-method and the particular semester that the method occurred in, resulting in an increased chance of Type-1 error (Peugh, 2010). All multilevel models were constructed using the "lmer" command in the "lme4" package in R-studio (Bates et al., 2015).

For a detailed analysis of student responses to Worked Example questions, binomial probabilities were calculated using BINOM.DIST() in Excel from the number of students and the probability of guessing each answer by chance. Focusing on guessing by chance as the null hypothesis, p-values indicated whether rates of answer selection were greater than expected or less than expected by chance ($p < .05$), or indistinguishable from chance ($p < .05$). For readability of the main text, detailed statistical results and comparisons are shown in Tables 2–4.

Group-level data are primarily visualized using violin plots, which indicate the average of the collected data as a black dot, as well as an envelope that allows the reader to see the distribution of data values. One advantage of these plots is that the reader can directly compare changes in distribution patterns. In contrast, comparisons of bar charts with standard error bars require making assumptions, which are often incorrect, about the underlying distributions.

To help with gauging the scale of the effects from the different instruction types, we also report the normalized gain from one instruction type to another (the base). This was calculated by taking the average score of one instruction type minus the average score of the base, and then dividing by the possible improvement, which is 1 minus the average score of the base.

Results

Overall, we found that the new teaching strategies improved student performance on the spatial pattern questions, as illustrated in violin plots colored by instruction type for each time frame and visualization. In nearly all cases, we find a shift of the distribution from Original Instruction upwards to higher scores with Worked Examples and New Animations (Figure 5). For example, the violin plots illustrate that student scores with the Original Instruction were skewed strongly toward low scores on both pre- and post-earthquake questions for the Fence Map and first GPS vector frame. For most questions, the violin plots show student scores during the Worked Examples and New Animations interventions are skewed strongly toward high scores on all three time frames of the earthquake cycle. The exception to this pattern was the distribution on the pre-earthquake question for the second GPS vector frame that skewed low, which was an indication that students were struggling with changing reference frames. To statistically examine these differences across the teaching strategies, and

Table 2. Results of the simple slopes moderation analysis probing the interaction between instruction and question type.

Instruction	Question type	β	*SE*	t^{++}	Normalized gains (%)
Original instruction *vs.* worked examples	Fence map	0.22	0.04	5.52***	34
Original instruction *vs.* worked examples	GPS vectors frame 1	0.44	0.04	11.31***	66
Original instruction *vs.* worked examples	GPS vectors frame 2	0.14	0.04	3.59***	21
Original instruction *vs.* worked examples	GPS vectors frame 3	0.3	0.04	7.64***	58
Original instruction *vs.* new animations	Fence map	0.33	0.04	7.49***	62
Original instruction *vs.* new animations	GPS vectors frame 1	0.49	0.04	11.19***	79
Original instruction *vs.* new animations	GPS vectors frame 2	0.28	0.04	6.37***	54
Original instruction *vs.* new animations	GPS vectors frame 3	0.36	0.04	8.21***	78
Worked examples *vs.* new animations	Fence Map	0.12	0.39	2.99**	42
Worked examples *vs.* new animations	GPS vectors frame 1	0.05	0.39	1.41	38
Worked examples *vs.* new animations	GPS vectors frame 2	0.14	0.39	3.64***	43
Worked examples *vs.* new animations	GPS vectors frame 3	0.07	0.39	1.68	48

A positive β indicates improvement from first to second instruction. GPS vector reference frames are shown in Figure 3.
$^{**}p < .01$; $^{***}p < .001$; $^{++}$Estimated degrees of freedom ranged from 728 to 744.

Table 3. Results of the simple slopes moderation analysis probing the interaction between instruction and time frame.

Instruction	Time frame	β	*SE*	t^{++}	Normalized gains (%)
Original instruction *vs.* worked examples	Pre-earthquake	0.2	0.04	5.55***	29
Original instruction *vs.* worked examples	During earthquake	0.25	0.04	7.1***	52
Original instruction *vs.* worked examples	Post-earthquake	0.37	0.04	10.44***	56
Original instruction *vs.* new animations	Pre-earthquake	0.35	0.04	8.71***	61
Original instruction *vs.* new animations	During earthquake	0.31	0.04	7.68***	73
Original instruction *vs.* new animations	Post-earthquake	0.44	0.04	11.05***	73
Worked examples *vs.* new animations	Pre-earthquake	0.15	0.03	4.33***	46
Worked examples *vs.* new animations	During earthquake	0.06	0.03	1.61	42
Worked examples *vs.* new animations	Post-earthquake	0.07	0.03	2.09*	38

$^{*}p < .05$; $^{***}p < .001$; $^{++}$Estimated degrees of freedom ranged from 529 to 549.

Table 4. Results of the simple slopes moderation analysis probing the interaction between instruction and course level.

Instruction	Course level	*β*	*SE*	t^{++}	Normalized gain (%)
Original instruction *vs.* worked examples	Intro level	0.3	0.04	6.96***	46
Original instruction *vs.* worked examples	Early majors	0.33	0.04	8.08***	60
Original instruction *vs.* worked examples	Upper level	0.08	0.07	1.16	28
Original instruction *vs.* new animations	Intro level	0.42	0.05	8.35***	65
Original instruction *vs.* new animations	Early majors	0.4	0.05	8.57***	72
Original instruction *vs.* new animations	Upper level	0.23	0.07	3.53***	81
New animations *vs.* worked examples	Intro level	0.12	0.04	2.96**	35
New animations *vs.* worked examples	Early majors	0.07	0.04	1.78	32
New animations *vs.* worked examples	Upper level	0.15	0.07	2.03*	73

*$p < .05$; **$p < .01$; ***$p < .001$; ++Estimated degrees of freedom ranged from 194 to 195.

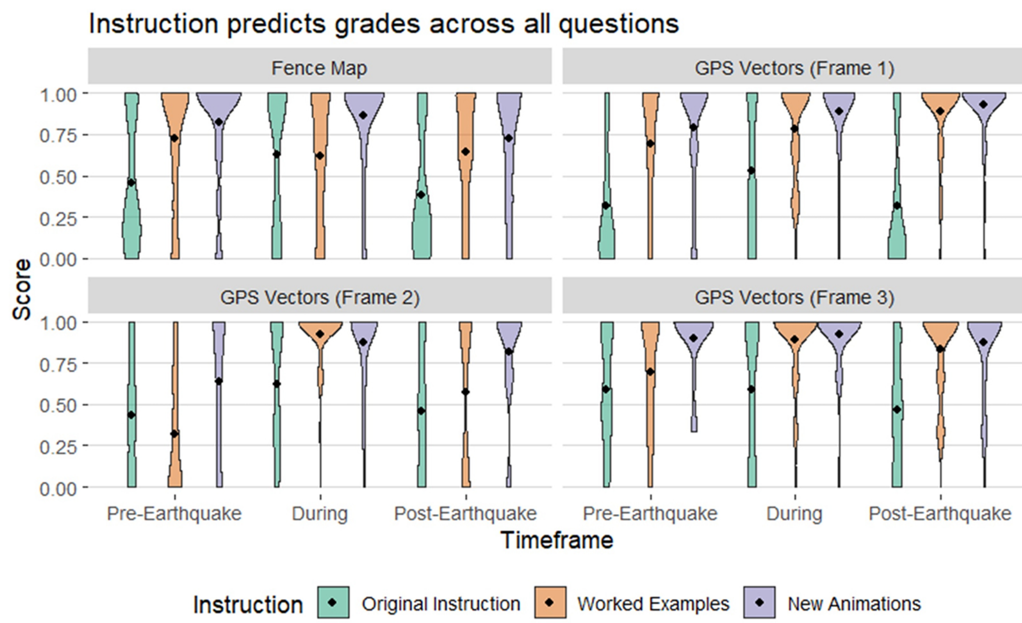

Figure 5. Violin plots showing the distribution of student scores for what the fence would look like before, during, and 100 years after an earthquake (upper left panel), and the successive three sets of questions on how the ground would be moving using GPS vectors pre-, during, and post-earthquake. The GPS vector reference Frame 1 is shown in Figure 3(a–f), reference Frame 2 in Figure 3(g–i), and reference Frame 3 in Figure 3(j–l). The average score for each question is shown as a black dot, and the envelope shows the distribution of scores, colored by the instruction type.

how the strategies may have supported reasoning about the different phases of an earthquake, we constructed three multilevel models.

Influence of instruction type for different spatial visualizations

First, we asked whether the Worked Examples and New Animations instruction types predicted better performance than the Original Instruction for the four different question types (e.g., Fence Map, GPS Vector Frames 1, 2, and 3). We constructed a multilevel model with question score as the outcome, an interaction between instruction and question type as the predictor, course level as a main effect, and student and semester as random effects. The model fit statistics improved (decreased) from the main effects model (AIC = 1712.8, BIC = 1770.8) to the interaction model (AIC = 1675.8, BIC = 1768.7). The improvement in model fit was significant [$\chi^2(6) = 48.99$, $p < .001$]. We detected a significant interaction between instruction and question type ($b = -0.16$, SE = 0.05, $t(2231) = -3.15$, $p = .002$, 95% CI: [−0.26, −0.07]). Here the fit significantly improves (AIC decreases) from the model with only main effects to the one with an added interaction term, with the interaction model explaining 38% of the variance in the data ($R^2 = 0.23$, R^2adj = 0.38).

We then ran a simple slopes moderation analysis so that we could evaluate how scores on each question varied at each level of instruction and question type (for details see Bauer, 2006). The simple slopes demonstrated that both Worked Examples and New Animations predicted better performance than Original Prep across all four question types (Figure 6; Table 2), including Fence Map, GPS vectors frames 1, 2, and 3. The normalized gains in performance from the Original Prep to Worked Examples were 35% on the Fence Map questions and 48% on the GPS questions. The normalized gains in performance from Original Prep to New Animations were higher: 52% on the Fence Map questions and 67% on the GPS questions. Simple slopes also detected a significant improvement in performance for the New Animations intervention relative to Worked Examples, though only for the Fence Map and frame 2 of the GPS vectors

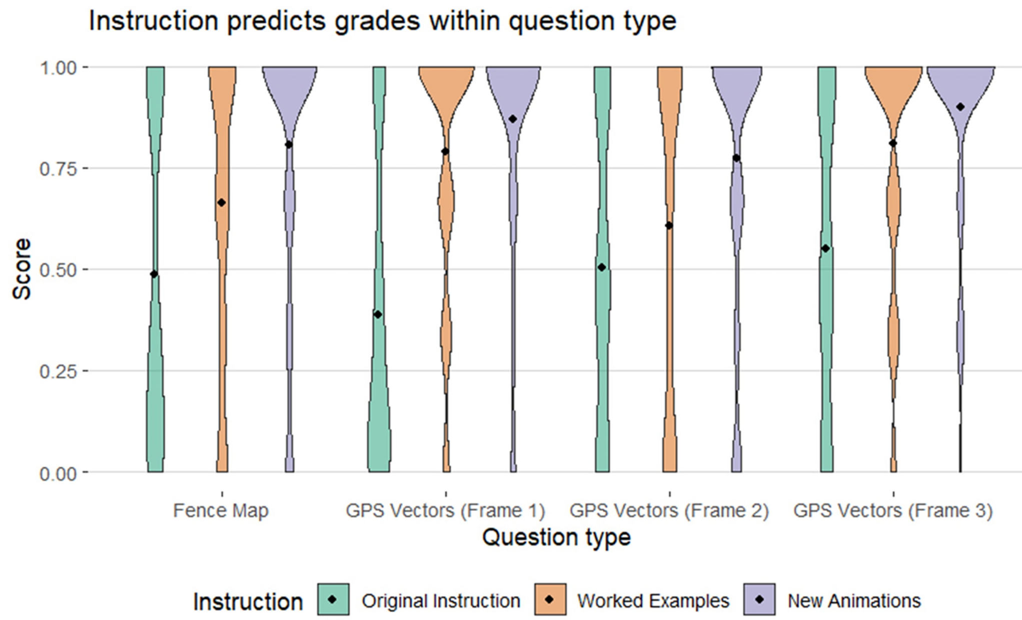

Figure 6. Violin plot showing the distribution of student scores when grouped by the visualization type (Fence Map, GPS Vector reference Frames 1, 2, and 3) and separated by the instruction type (colors). Format as in Figure 5. GPS vector reference frames are shown in Figure 3.

questions (Figure 6; Table 2), with a normalized gain of 43%. Although New Animations scores were higher than Worked Examples for several questions, the primary advantage of the New Animations intervention relative to the Worked Examples in this data set was to help with changing reference frames.

Influence of instruction type for different time frames

Next, we examined whether the Worked Examples and New Animations instruction types predicted better performance than Original Instruction for different question time frames (e.g. pre-earthquake and post-earthquake). We constructed a multilevel model with question score as the outcome, an interaction between instruction and question time frame as the predictor, course level as a main effect, and student and semester as random effects.

The model fit statistics improved (decreased) from the main effects model (AIC = 1679.9, BIC = 1732.1) to the interaction model (AIC = 1662.6, BIC = 1738). The improvement in model fit was significant [$\chi^2(4) = 25.29$, $p < .001$]. We detected a significant interaction between instruction and question type ($b = -0.1$, SE = 0.04, $t(2234) = -2.45$, $p = .01$, 95% CI: [−0.17, −0.02]). The interaction model explained 38% of the variance in the data ($R^2 = 0.24$, R^2adj = 0.38). Again, we then applied a simple slopes moderation analysis to probe the interaction at each level of instruction and question time frame.

Simple slopes demonstrated that Worked Examples and New Animations predicted better performance than Original Instruction across all three question time frames (Figure 7; Table 3), including pre-earthquake, during earthquake, and post-earthquake. The normalized gains in performance from Original Instruction to New Animations were 52% for pre-earthquake, 70% for during earthquake, and 69% for post-earthquake. Simple slopes analysis also detected that New Animations predicted significantly better performance compared to Worked Examples, though only for pre-earthquake and post-earthquake questions (Figure 9; Table 3), with a normalized gain of 27%.

Influence of instruction type for different course level

Finally, we wanted to examine whether the efficacy of the instruction type varied depending on how much experience students had with earth science in general. We constructed a multilevel model with question score as the outcome, an interaction between instruction and course level as the predictor, and student and semester as random effects. The model fit statistics improved (decreased) from the main effects model (AIC = 1771.4, BIC = 1812) to the interaction model (AIC = 1767.5, BIC = 1831.3). The improvement in model fit was significant [$\chi^2(4) = 11.91$, $p = .02$]. We detected a significant interaction between instruction and question type ($b = -0.19$, SE = 0.08, $t(194) = -2.3$, $p = .02$, 95% CI: [−0.34, −0.03]). The interaction model explained 33% of the variance in the data ($R^2 = 0.2$, R^2adj = 0.33). Again, we then applied a simple slopes moderation analysis to probe the interaction at each level of instruction and question time frame.

The simple slopes analysis detected that Worked Examples improved performance on questions for students at Intro Level and Early Majors levels relative to Original Instruction with a normalized gain of 53%, compared to 28% for Upper Level students (Figure 8; Table 4). New Animations improved performance on questions for students at all levels relative to Original Instruction producing a normalized gain of 73% with the highest values at the Upper Level. New Animations improved scores over Worked Examples for Intro and Upper Level students, with normalized gains of 35 and 73%,

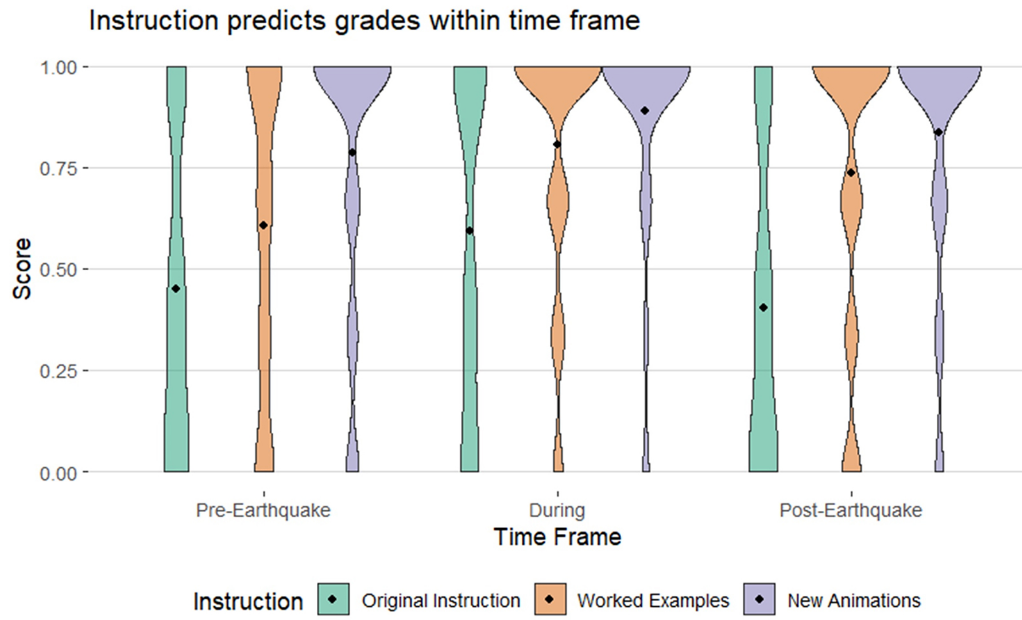

Figure 7. Violin plot showing the distribution of student scores when grouped by the earthquake time frame and separated by the instruction type (colors). Format as in Figure 5.

respectively. Notably, these findings support the notion that the Worked Examples approach is sufficient to help novices to perform at a level similar to advanced undergraduates, but the New Animations approach can also help advanced students improve significantly.

Students reasoning for correct answers

We then used the questions generated during the Worked Examples intervention to investigate how well students understood the reasons for the spatial patterns before, during, and after the earthquake. We focused on questions illustrating the correct spatial pattern and then asking students what is the primary reason for this pattern. As with other questions, students who chose a wrong answer were allowed to try again until they selected the correct answer. Figure 9 illustrates the results of this investigation for analogous fence (lighter bars) and GPS vector (darker bars) questions for the two intervention conditions (Worked Examples in orange and New Animations in purple). The results are strikingly different for the pre-earthquake questions (top panel) compared to those during and after the earthquake (middle and bottom panels). For the during and after earthquake questions, more than ~70% of students identified the correct answer on the first try: that slip during the earthquake enabled areas near the fault to move quickly to catch up, and that after the earthquake the fault should be stuck and causing rocks to bend again. There are slight improvements from the initial Fence Map question to the subsequent GPS Vector question during an assignment, as well as slight improvements from the Worked Example to the New Animations intervention. In contrast, fewer students correctly selected the correct explanation for the pre-earthquake pattern, that friction does not allow the sides to move, with accuracy ranging from 32 to 54% across the conditions. In addition, students did not select the correct explanation at appreciably higher rates than other answers on the second try. Students did not select a specific wrong answer at higher rates than others, suggesting that there was not an alternative misconception that students were drawn toward.

To put the pre-earthquake correct spatial pattern question associated with Figure 9 in context, it helps to review the two prior worked example questions that ask about the reasons why the spatial patterns would not occur before an earthquake. The first worked example asked why a translated fence without an offset across the fault (Figure 2(d)) was not the correct spatial pattern with the right answer being “The fault requires motion to be in opposite directions on opposite sides of the fault.” This is designed to highlight a condition necessary for elastic rebound that was not being met. The second worked example spatial pattern satisfies this condition because it shows a completely offset fence (Figure 2(g)), which is the correct spatial pattern immediately after an earthquake, but it is not the correct pre-earthquake spatial pattern. The right answer for why this spatial pattern is not correct is “The equal motions on opposite sides of the fault indicate there is no friction on the fault.” This helps to put the distractor answers for Figure 9 in context. For example, “the motions on either side of the fault must be in opposite directions” is an insufficiently correct answer in Figure 9 because even though this condition was present in Figure 2(g) it did not produce the correct spatial pattern as this condition could not explain why the fence is bent. Instead, the second worked example focused on Figure 2(g) and highlighted that friction on the fault was the missing condition, which should have guided students to the correct answer for the question in Figure 9 that “friction on the fault does not allow the two sides to move in opposite directions.”

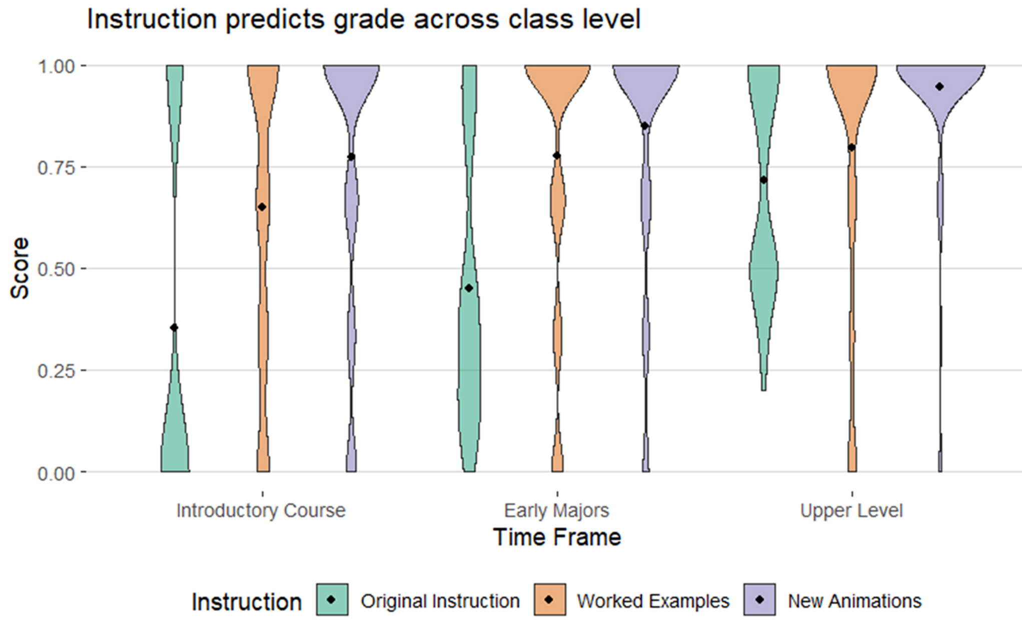

Figure 8. Violin plot showing the distribution of student scores when grouped by the course level and separated by the instruction type (colors). Format as in Figure 5.

To assess students' performance on the pre-earthquake questions relative to guessing by chance, we used a binomial test adjusting the probabilities as the number of choices changed from one try to the next—the probability of guessing the correct answer = 1/(4 − number of prior attempts) and the probability of guessing an incorrect answer = (1 − (1/(4 − number of prior attempts)))/3. On the first try, choosing the correct answer was reliably above chance ($p < .05$), but almost all of the incorrect answers were not reliably different from chance ($p < .05$), and a couple were chosen at rates above chance. Indeed, the answer most likely to be avoided on the first try (motion was building) was frequently selected on the second and third tries, at rates above chance or rates that did not differ from chance. In contrast, students who were wrong on their first try for the during and after earthquake questions selected the correct answer above chance ($p < .05$) on their second try for the majority of conditions. For all questions by the third try, all choices were not reliably different from chance ($p < .05$). However, 42% of the students made a third try for the pre-earthquake question, and <9% for the during and after earthquake questions.

These results indicate the majority of students were guessing when asked to explain the spatial (fence) and motion (GPS) patterns present before an earthquake, and understanding only improved slightly during the assignment or with the New Animations intervention. In contrast, students clearly understood that a similar answer explained the spatial and motion patterns after an earthquake and were able to discern from similar distractor answers (Figure 9, top *vs.* bottom). Taken together, these results suggest that fault locking due to friction on the fault is the most difficult concept for students to comprehend in regards to how elastic rebound causes earthquakes.

Discussion

Taken together, the New Animations instructional style predicted better performance for all questions compared to Original Instruction, and for some questions compared to the Worked Examples instructional style. Additionally, New Animations never predicted worse performance on any specific questions compared to Worked Examples when considering all of the students in the study. The learning benefits associated with both Worked Examples and New Animations declined as students gained more experience in earth science, but this was evident at all levels. This suggests that these instructional types are efficacious at improving learning outcomes even for advanced students who have already developed the key conceptual knowledge illustrated in the New Animations.

The results also indicate that the Worked Examples approach provides support for connecting fault motion to pre-fault deformation (e.g., Figure 7). A key motivation for utilizing Worked Examples was to address the issue of students choosing a spatial pattern that would be present for a different time frame by engaging them in a consideration of the conceptual reasoning for the correct spatial pattern. We suggest the improvement in performance from Worked Examples may be due to the spatial accommodation process as Worked Examples help students to initiate the reconstruction of an improved mental model. However, the support appears somewhat fragile as transfer to new orientations and new frames of reference was weak (e.g., Figure 6). Here the transfer appears to have been stronger when New Animations were provided with multiple representations that illustrated how changing the frame of reference would change the size of the pre-earthquake motion at a given location (Figure 4), but that the overall gradient remained the same. The literature on when animations help learning is mixed (Harrower

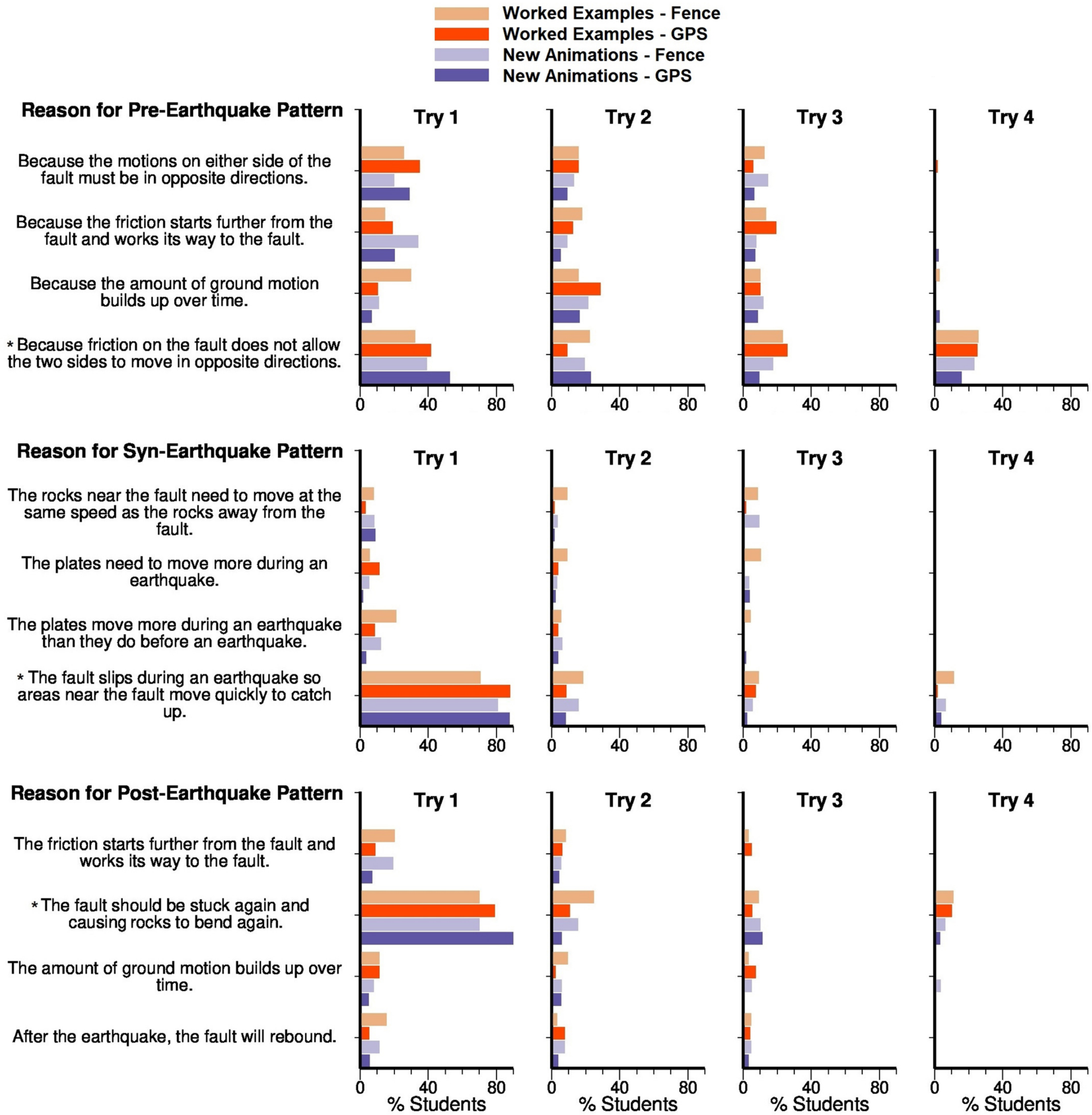

Figure 9. Percentage of students that selected each potential reason that would explain the correct spatial pattern. The correct answer is marked with a star. Students choosing a wrong answer were allowed to try again until selecting the correct answer. Orange colors indicate responses for the Worked Examples intervention and purple colors indicate those for the New Animations intervention. Lighter colors are for the initial Fence Map question and darker colors are for the subsequent GPS Vector question.

& Fabrikant, 2008 *vs.* Tversky et al., 2002). Here the animations may have been particularly helpful because the equivalence of static representations of motion gradients as vector arrows may be particularly opaque. Having a visualization with a clearly consistent motion to illustrate how changing a frame of reference will change the vector field but not the overall pattern of deformation may have allowed students to develop a more robust mental model that connects the vector field to the deformation field.

By changing which portion of the region had a zero velocity we effectively changed the frame of reference, which appeared to substantially disrupt performance. We know from prior work that people have preferred frames of reference (Hegarty & Waller, 2004), and that there are also individual differences in which frame of reference is adopted for a given problem (Liben, 1991). For example, individual differences in people's beliefs about the orientation of fluid within a tilted bottle appear to arise from whether the individual adopts a bottle-centered or gravity-centered frame of reference. That anyone would believe the fluid tilts with the bottle may be counterintuitive for some readers, nevertheless, a significant number of people consistently report expecting the water to tilt as the bottle is rotated (Liben, 1991), and furthermore, the tendency predicts some STEM outcomes (Liben, 1991) and is thus relevant for geosciences education (Kastens & Ishikawa, 2006). Within the context of a moving Earth, challenges in understanding hotspot motion appear to arise, at least in part, from failures to coordinate the frames of reference of the surface and the whole Earth. Errors in spatial predictions about where a hotspot will appear suggest reasoning about the movement of the Earth surface, which is our environment and thus likely to be experienced as the stable framework, is challenging for some (LaDue & Shipley, 2018). Here we aimed to develop flexibility in adopting new frames of reference through the practice of aligned examples where the frame of motion reference varied. Multiple opportunities to align was not as helpful as seeing the animation to transfer reasoning from one frame of reference to another.

In seismology, a common convention is to set the motion on one side of a fault to zero when the focus of a visualization is on how much deformation is occurring near the fault relative to the stable part far from the fault (e.g., interior of a tectonic plate). However, when displaying larger-scale data, such as a regional network of faults, both sides of a local fault may have non-zero motions despite the convention. Thus, facile reasoning from motion vectors requires developing pattern recognition that is both orientation independent and absolute magnitude independent. This is not a minor achievement, and we suspect it means that early understanding of fault systems from local motion data will be fragile. Some students will need to develop more robust mental models as they proceed toward professional careers that will require regularly working with such data. How to advance these skills is an area of research that could offer some evidence-informed guidance.

Although the difficulty in manipulating velocity vectors into a new reference frame was a clear hurdle for students in this analysis, the results indicated the most prominent misconception was the difficulty in recognizing the fault must be locked for the strain to accumulate and produce an earthquake. This was unexpected as prior work had suggested the difficulty in accepting that rocks bend was the primary misconception in understanding the elastic rebound theory (e.g., Hubenthal, 2018a). The Worked Examples intervention provided new information on how students conceptualized what is necessary to cause the spatial patterns involved in the elastic rebound process. In particular, it highlighted how students were able to identify the importance of bending to produce the spatial patterns but still struggled to identify the importance of fault friction to cause locking, which prevents the two sides of the fault from moving in opposite directions. The fact the students did not choose a particular wrong answer along with statistical evidence that a large portion of students were guessing indicates that the primary issue is with limited understanding of friction and fault locking. This indicates the spatial error signal was insufficient to help the students adjust their mental model. Thus, it appears many students are still missing a key piece of the process that generates the energy necessary to cause earthquakes. This is perhaps not surprising as theoretical descriptions of friction involve sophisticated physical principles that can be challenging to teach (Besson et al., 2007; Kurnaz & Eksi, 2015; Marone, 1998), and fault locking has been recognized as a problematic concept within geoscience (Lay & Schwartz, 2004; Wang & Dixon, 2004a, 2004b).

Limitations and future work

There are several limitations of this study associated with the fact that the assignments evaluated are all from a single instructor at a single institution, which limits our ability to evaluate how well the results would apply in other courses at institutions with a different student body. The initial assignment was designed to be open-answer and was only graded by the instructor so there are no inter-rater reliability estimates. Given all data was collected at a single institution, there were some students who received multiple exposures. None of the students in the Intro Level were found in the other two course types, five students in the Upper Level had taken the Early Majors course before, and two students in the Upper Level had taken the other Upper Level course before. The total number of students with multiple exposures is too small to fully evaluate the impact, but it may have contributed to better average performance in the Upper Level group.

Another limitation is the use of multiple-choice questions to evaluate student performance. This type of question is likely effective in the evaluation of student predictions of ground motions under various conditions and earthquake cycle timing, but it is likely to be less effective for evaluating student understanding of the concepts involved in elastic rebound. The distractor answers for the reasoning of right and wrong answers in the worked example questions were derived from the initial open-answer assessments and evaluated by the seismologist on the team (Brudzinski) for evidence of correct understanding. However, this approach does not ensure these distractor answers are tightly tied to specific misconceptions. As such, our ability to use student selections of these distractor answers as approximations of specific misconceptions is limited. Specifically, although it is clear that students struggled with the pre-earthquake questions, and from their answers we inferred this was due in part to failure to understand the role of friction, work is needed to delve deeper into students' mental models of pre-earthquake Earth processes than multiple choice examples allow. Future work to explore mental models and evaluate whether the questions and distractor answers were interpreted as intended could be accomplished *via* think aloud interviews or through additional expert validation.

The design allows us to assess students only within a limited time window for a limited range of items. As noted above there is fairly clear evidence that initial learning is relatively brittle, in the sense that surficial changes to questions, such as orientation of the fault, result in incorrect answers from students who had appeared to have mastered the question. Such sensitivity to surface features, and a lack of robust generalization, suggests that the learning might be transient. We are limited in estimating how long the gains seen for Worked Examples and New Animations persist. Future work will be needed to determine if the skills developed in these classes will transfer to future encounters with similar seismology concepts.

An important limitation that applies here, but also more broadly to designing new approaches to education in the geosciences, is that there is a lack of well-articulated theoretical accounts of how skills are acquired and the underlying cognitive processes that support them (e.g., St. John et al., 2020). Thus, we are hampered in interpreting the modest learning gains seen for the pre-earthquake question with a new frame of reference. This is likely due to the challenge of developing a generalized spatial thinking skill that allows concepts to be seen to apply in all orientations. It may also be that students are in the early stages of developing a broad enough set of concept exemplars so that concepts appear to be orientation independent. Here partnering with cognitive process modelers to track and simulate the spatial and temporal progression of concept learning might pay significant dividends as they have in the area of hand sample classification (Miyatsu et al., 2019; Nosofsky et al., 2018).

Implications

Our study indicates that richer ground motion data can be used in instruction when presented with appropriate support. A combination of animations that illustrate ground motion over time and multiple opportunities to practice making predictions appears to support the creation of mental models that connect the spatial displacements to the ground velocities. We noted that the new mental models were somewhat fragile given the application to new orientations and reference frames were weaker without the support of animations in the instructional materials. Nevertheless, these mental models are likely an important step to moving geoscience students from thinking of tectonic plates as undeformable objects, so should be included in the solid earth geology major curriculum.

Our results also indicate that many students' mental models lacked an understanding of the role of fault locking caused by friction as essential to cause the strain build up needed for successive earthquakes on the same fault. New approaches will likely be required to help students to build a mental model to use for a robust understanding of this aspect of the elastic rebound process. One possibility would be to use physical models (e.g. Dolphin & Benoit, 2016; Hubenthal, 2018a) to focus attention on demonstrating how a normal force pressing the two sides of the fault together is required for friction to cause fault locking. Perhaps a better option would be to use physics-based simulations (e.g. Wieman et al., 2008) to more fully enable students to explore how the patterns of ground motion over time would be different when the coefficient of friction on a fault is varied from frictionless to completely locked. The lack of bending resulting from the frictionless scenario may help students to understand how friction is a necessary component to store the bending energy that would eventually lead to an earthquake.

Conclusions

Recent studies have identified that student understanding of elastic rebound and how it causes earthquakes is often incomplete (Dolphin & Benoit, 2016; Hubenthal, 2018a). We hypothesized that realistic imaging of spatial patterns in ground motions over the course of the earthquake cycle could help students visualize the elastic rebound concept (Brudzinski, 2018). Incorporating change information, in the form of both motion vectors and before-during-after contrasts, into a mental model required most students to change an existing model or develop a new model. We evaluated strategies for supporting mental model changes when presenting variations in ground motion, such as map views of fence bending and GPS velocity vectors, to educate undergraduate students on the concept of elastic rebound. We collected student performance on questions related to elastic rebound following several intervention strategies in four courses from introductory to upper level at an R2 institution in the Midwest between Fall 2016 and Spring 2019. Assessments began as an open-answer format and then we created a multiple-choice version from the most common answers, including new "worked example" questions inquiring about the reasons certain answers were correct or incorrect.

The results identified two key challenges: (a) difficulty in recognizing velocity vector patterns when presented in a new reference frame, and (b) difficulty in reasoning the fault must be locked for elastic strain to accumulate and eventually produce an earthquake. Unexpectedly, we did not find difficulty with the concept that rocks bend despite prior work suggesting this was the primary misconception in understanding the elastic rebound theory (Hubenthal, 2018a). The struggles with recognizing motion patterns highlight that while GPS data are a boon to geoscientists in showing authentic spatial patterns in ground motion, this type of data differs from traditional bent fence plots by showing velocity instead of displacement, instantaneous motion *vs.* a historical accumulation, punctuated measurements as opposed to a continuous line segment, and various frames of reference can be chosen to highlight motion relative to different points of view. An important takeaway is that these challenges can be mitigated and richer ground motion data can be used in instruction when presented in a context with repeated opportunities to make predictions combined with animations to support mental models that connect the spatial displacements to the ground velocities. Despite substantial improvements in some areas due to instructional interventions, it is not clear that after instruction most students' mental model included an understanding of the role of fault locking caused by friction to produce successive earthquakes on the same fault. New approaches may be required to provide a model that students can assimilate and use for a robust understanding of this aspect of elastic rebound.

Acknowledgments

This article benefited greatly from in-depth discussions with the GET-Spatial collaborative network (GET-Spatial, 2017) in addition to input from S. Graham, M. Hubenthal, D. Sumy, two peer reviewers, an associate editor, and a research editor. We are grateful to SERC for hosting the final form of the assignment and associated teaching resources. Materials provided by IRIS Consortium were used in the instruction.

Disclosure statement

The authors report there are no competing interests to declare. The views and conclusions contained in this document are those of the authors and do represent the opinions of the USGS or NSF. The data used in this study was collected under Temple IRB protocol 23869 and Miami IRB protocol 03584e.

Funding

This work was supported by the National Science Foundation under Grants EAR-2025073 and SBE-1640800; and the United States Geological Survey under grants G20AP00104 and G20AP00066. The facilities of IRIS Consortium were supported by the National Science Foundation's Seismological Facility for the Advancement of Geoscience (SAGE) Award under Cooperative Agreement EAR-1724509.

ORCID

Michael R. Brudzinski http://orcid.org/0000-0003-1869-0700
Thomas F. Shipley http://orcid.org/0000-0001-8397-2486
Joy Ham http://orcid.org/0000-0002-9306-2117

References

Aydin, F., & Coskun, M. (2011). Gifted students' opinions about earthquake: A qualitative study. *International Journal of Physical Sciences*, *6*, 1863–1867.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*(4), 437–448. https://doi.org/10.3102/10769986031004437

Besson, U., Borghi, L., De Ambrosis, A., & Mascheretti, P. (2007). How to teach friction: Experiments and models. *American Journal of Physics*, *75*(12), 1106–1113. https://doi.org/10.1119/1.2779881

Brudzinski, M. R. (2018). Using GPS velocity vectors to illustrate elastic rebound. *In the Trenches*, *8*(1), 1–2. https://nagt.org/nagt/publications/trenches/v8-n1/196290.html

Brudzinski, M. R. (2023). Spatial visualization strategies for improving student understanding of the elastic rebound theory of earthquakes. Retrieved December 7, 2023, from https://serc.carleton.edu/teachearth/activities/264843.html

Brudzinski, M. R., Jaeger, A., & Shipley, T. F. (2019). Challenges in making meaning from ground-motion visualizations: The role of geoscience knowledge in interpreting dynamic spatiotemporal patterns. *Seismological Research Letters*, *90*(4), 1692–1701. https://doi.org/10.1785/0220180289

Cox, A., & Hart, R. B. (1991). *Plate tectonics: How it works*. John Wiley & Sons.

Davatzes, A., Gagnier, K., Resnick, I., & Shipley, T. (2018). Learning to form accurate mental models. *EOS*, *99*, 1–8. https://doi.org/10.1029/2018EO091643

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, *33*(2–3), 109–128. https://doi.org/10.1207/s15326985ep3302&3_5

Dolphin, G. (2018). The history of elastic rebound theory: How a big disaster helped us better understand how the Earth works. *In The Trenches*, *8*(1), 1–2. https://nagt.org/nagt/publications/trenches/v8-n1/196287.html

Dolphin, G., & Benoit, W. (2016). Students' mental model development during historically contextualized inquiry: How the 'Tectonic Plate' metaphor impeded the process. *Science Education*, *38*(4), 267–297.

Flunger, B., Trautwein, U., Nagengast, B., Lüdtke, O., Niggli, A., & Schnyder, I. (2021). Using multilevel mixture models in educational research: An illustration with homework research. *The Journal of Experimental Education*, *89*(1), 209–236. https://doi.org/10.1080/00220973.2019.1652137

Freymueller, J. T. (2021). GPS, tectonic geodesy. In *Encyclopedia of solid earth geophysics* (pp. 558–578). Springer International Publishing.

Gagnier, K. M., Atit, K., Ormand, C. J., & Shipley, T. F. (2017). Comprehending diagrams: Sketching to support spatial reasoning. *Topics in Cognitive Science*, *9*(4), 883–901. https://doi.org/10.1111/tops.12233

Geoscience Education Transdisciplinary Spatial Learning Network (2017). Retrieved November 24, 2023, from https://serc.carleton.edu/getspatial

Harrower, M., & Fabrikant, S. (2008). The role of map animation in geographic visualization. In M. Dodge (Ed.), *Geographic visualization: Concepts, tools and applications* (pp. 49–65). Wiley.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191. https://doi.org/10.1016/j.intell.2003.12.001

Hubenthal, M. (2018a). Exploring undergraduates' conceptions of elasticity, within a plate tectonics context, before and after experience with rock's elastic behavior. *Journal of Geoscience Education*, *66*(4), 261–277. https://doi.org/10.1080/10899995.2018.1493964

Hubenthal, M. (2018b). Bending rocks in the classroom. *In the Trenches*, *8*(1), 1–3. https://nagt.org/nagt/publications/trenches/v8-n1/bending_rocks.html

Hubenthal, M., Stein, S., & Taber, J. (2011). A big squeeze: Examining and modeling causes of intraplate earthquakes in the earth science classroom. *The Earth Scientist*, *27*(1), 33–39.

Incorporated Research Institutions for Seismology (2016). Fault: Strike-slip – High friction with elastic rebound. Retrieved August 21, 2016, from https://www.iris.edu/hq/inclass/animation/elastic_rebound_on_highfriction_strikeslip_fault

Kali, Y., & Orion, N. (1996). Spatial abilities of high school students in the perception of geologic structures. *Journal of Research in Science Teaching*, *33*(4), 369–391. https://doi.org/10.1002/(SICI)1098-2736(199604)33:4<369::AID-TEA2>3.0.CO;2-Q

Kastens, K. A., & Ishikawa, T. (2006). Spatial thinking in the geosciences and cognitive sciences. In C. Manduca, D. Mogk, & N. Stillings (Eds.), *Earth and mind: How geoscientists think and learn about the complex earth* (pp. 53–76). Geological Society of America Special Paper 413.

Kastens, K. A., Agrawal, S., & Liben, L. S. (2009). How students and field geologists reason in integrating spatial observations from outcrops to visualize a 3-D geological structure. *International Journal of Science Education*, *31*(3), 365–393. https://doi.org/10.1080/09500690802595797

Kirikkaya, E. B., Çakin, O., Imali, B., & Bozkurt, E. (2011). Earthquake training is gaining importance: The views of 4th and 5th year students on earthquakes. *Procedia-Social and Behavioral Sciences*, *15*, 2305–2313. https://doi.org/10.1016/j.sbspro.2011.04.098

Kurnaz, M. A., & Eksi, C. (2015). An analysis of high school students' mental models of solid friction in physics. *Educational Sciences: Theory & Practice*, *15*(3), 787–795.

LaDue, N., & Shipley, T. F. (2018). Click-on-diagram questions: A new tool to study conceptions using classroom response systems. *Journal of Science Education and Technology*, *27*(6), 492–507. https://doi.org/10.1007/s10956-018-9738-0

LaDue, N., Hubenthal, M., & Dolphin, G. (2018). Using models to develop deep understandings of earthquakes. *In the Trenches*, *8*(1), 1–3. https://nagt.org/nagt/publications/trenches/v8-n1/using_models.html

Lay, T., & Schwartz, S. Y. (2004). Comment on "Coupling semantics and science in earthquake research". *EOS, Transactions American Geophysical Union*, *85*(36), 339–340. https://doi.org/10.1029/2004EO360003

Libarkin, J. C., Kurdziel, J. P., & Anderson, S. W. (2007). College student conceptions of geological time and the disconnect between ordering and scale. *Journal of Geoscience Education*, *55*(5), 413–422. https://doi.org/10.5408/1089-9995-55.5.413

Liben, L. S. (1991). Adults' performance on horizontality tasks: Conflicting frames of reference. *Developmental Psychology*, *27*(2), 285–294. https://doi.org/10.1037/0012-1649.27.2.285

Lombardi, D., Shipley, T. F., Bailey, J. M., Bretones, P. S., Prather, E. E., Ballen, C. J., Knight, J. K., Smith, M. K., Stowe, R. L., Cooper, M. M., Prince, M., Atit, K., Uttal, D. H., LaDue, N. D., McNeal, P. M., Ryker, K., St. John, K., van der Hoeven Kraft, K. J., & Docktor, J. L. (2021). The curious construct of active learning. *Psychological Science in the Public Interest*, *22*(1), 8–43. https://doi.org/10.1177/1529100620973974

Marone, C. (1998). Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences*, *26*(1), 643–696. https://doi.org/10.1146/annurev.earth.26.1.643

Miyatsu, T., Gouravajhala, R., Nosofsky, R. M., & McDaniel, M. A. (2019). Feature highlighting enhances learning of a complex natural-science category. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *45*(1), 1–16. https://doi.org/10.1037/xlm0000538

Nosofsky, R. M., Sanders, C. A., Meagher, B. J., & Douglas, B. J. (2018). Toward the development of a feature-space representation for a complex natural category domain. *Behavior Research Methods*, *50*(2), 530–556. https://doi.org/10.3758/s13428-017-0884-8

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*(1), 85–112. https://doi.org/10.1016/j.jsp.2009.09.002

Rescorla, R. A., & Wagner, A. R. (1972). *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, classical conditioning II* (A. H. Black & W. F. Prokasy, Eds., pp. 64–99). Appleton-Century-Crofts.

Resnick, I., Davatzes, A., Newcombe, N. S., & Shipley, T. F. (2017a). Using relational reasoning to learn about scientific phenomena at unfamiliar scales. *Educational Psychology Review*, *29*(1), 11–25. https://doi.org/10.1007/s10648-016-9371-5

Resnick, I., Davatzes, A., Newcombe, N. S., & Shipley, T. F. (2017b). Using analogy to learn about phenomena at scales outside of human perception. *Cognitive Research: Principles and Implications*, *2*(1), 21. https://doi.org/10.1186/s41235-017-0054-7

Resnick, I., Newcombe, N., & Shipley, T. F. (2017). Dealing with big numbers: Representation and understanding of magnitudes outside of human experience. *Cognitive Science*, *41*(4), 1020–1041. https://doi.org/10.1111/cogs.12388

Santos-Reyes, J., Santos-Reyes, G., Gouzeva, T., & Velazquez-Martinez, D. (2017). Schoolchildren's earthquake knowledge, preparedness, and risk perception of a seismic-prone region of Mexico. *Human and Ecological Risk Assessment: An International Journal*, *23*(3), 494–507. https://doi.org/10.1080/10807039.2016.1188368

Segall, P., & Davis, J. L. (1997). GPS applications for geodynamics and earthquake studies. *Annual Review of Earth and Planetary Sciences*, *25*(1), 301–336. https://doi.org/10.1146/annurev.earth.25.1.301

Sit, S. M., & Brudzinski, M. R. (2017). Creation and assessment of an active e-learning introductory geology course. *Journal of Science Education and Technology*, *26*(6), 629–645. https://doi.org/10.1007/s10956-017-9703-3

Sözen, E. (2019). The earthquake awareness levels of undergraduate students. *Journal of Pedagogical Research*, *3*(2), 87–101. https://doi.org/10.33902/JPR.2019254175

St. John, K., McNeal, K. S., MacDonald, R. H., Kastens, K. A., Bitting, K. S., Cervato, C., McDaris, J. R., Petcovic, H. L., Pyle, E. J., Riggs, E. M., Ryker, K., Semken, S., & Teasdale, R. (2020). A community framework for geoscience education research: Summary and recommendations for future research priorities. *Journal of Geoscience Education*, *69*(1), 2–13. https://doi.org/10.1080/10899995.2020.1779569

Tsai, C. C. (2001). Ideas about earthquakes after experiencing a natural disaster in Taiwan: An analysis of students' worldviews. *International Journal of Science Education*, *23*(10), 1007–1016. https://doi.org/10.1080/09500690010016085

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, *57*(4), 247–262. https://doi.org/10.1006/ijhc.2002.1017

Wade, N. J. (1996). Frames of reference in vision. *Minimally Invasive Therapy & Allied Technologies*, *5*(5), 435–439. https://doi.org/10.3109/13645709609153705

Wang, K., & Dixon, T. (2004a). "Coupling" semantics and science in earthquake research. *EOS, Transactions American Geophysical Union*, *85*(18), 180–180. https://doi.org/10.1029/2004EO180005

Wang, K., & Dixon, T. (2004b). Reply to Comment on "Coupling semantics and science in earthquake research. *EOS, Transactions American Geophysical Union*, *85*(36), 339–340.

Wieman, C. E., Adams, W. K., & Perkins, K. K. (2008). PhET: Simulations that enhance learning. *Science*, *322*(5902), 682–683. https://doi.org/10.1126/science.1161948