

RESEARCH ARTICLE

Open Access



Puzzle-based versus traditional lecture: comparing the effects of pedagogy on academic performance in an undergraduate human anatomy and physiology II lab

Lucas Stetzik*, Anthony Deeter, Jamie Parker and Christine Yukech

Abstract

Background: A traditional lecture-based pedagogy conveys information and content while lacking sufficient development of critical thinking skills and problem solving. A puzzle-based pedagogy creates a broader contextual framework, and fosters critical thinking as well as logical reasoning skills that can then be used to improve a student's performance on content specific assessments. This paper describes a pedagogical comparison of traditional lecture-based teaching and puzzle-based teaching in a Human Anatomy and Physiology II Lab.

Methods: Using a single subject/cross-over design half of the students from seven sections of the course were taught using one type of pedagogy for the first half of the semester, and then taught with a different pedagogy for the second half of the semester. The other half of the students were taught the same material but with the order of the pedagogies reversed. Students' performance on quizzes and exams specific to the course, and in-class assignments specific to this study were assessed for: learning outcomes (the ability to form the correct conclusion or recall specific information), and authentic academic performance as described by (Am J Educ 104:280–312, 1996).

Results: Our findings suggest a significant improvement in students' performance on standard course specific assessments using a puzzle-based pedagogy versus a traditional lecture-based teaching style. Quiz and test scores for students improved by 2.1 and 0.4 % respectively in the puzzle-based pedagogy, versus the traditional lecture-based teaching. Additionally, the assessments of authentic academic performance may only effectively measure a broader conceptual understanding in a limited set of contexts, and not in the context of a Human Anatomy and Physiology II Lab.

Conclusion: In conclusion, a puzzle-based pedagogy, when compared to traditional lecture-based teaching, can effectively enhance the performance of students on standard course specific assessments, even when the assessments only test a limited conceptual understanding of the material.

Keywords: Puzzle, Lecture, Teaching, Pedagogy, Anatomy, Physiology, Performance, Concept

Background

Throughout the past 25 years, studies involving teaching methods that differ from the standard lecture/memorization/testing format have shown promise in increasing the level of authentic intellectual performance demonstrated by students [1–5]. Thus, in more recent studies it has been argued that traditional teaching methods convey information and content while lacking

sufficient development of critical thinking skills and problem solving [6]. The framework used to score non-traditional teaching methods has focused on learning outcomes, and content-specific conceptual understanding [6, 7]. In an effort to assess the broader effects of pedagogy on a student's critical thinking skills, this study has implemented a framework with which to determine the level of students' authentic intellectual performance [8–13]. The authentic intellectual performance framework used for this study was based on the following

* Correspondence: lstetzik@gmail.com
University of Akron, Akron, OH, USA

standards; higher-order thinking, depth of knowledge, connectedness to the world beyond the classroom, substantive conversation, and social support for student achievement.

Previous studies have established definitions for project-, problem-, and puzzle-based teaching methods in which each method builds upon those before it [6]. The least abstract method, project-based, includes working in teams and dealing with uncertainty and changing conditions. Built upon this method is problem-based learning; this method involves acquiring domain-specific knowledge and reasoning with domain-specific methods. The most abstract method, puzzle-learning, builds upon both of these to develop critical thinking and logical reasoning independent of a specific domain [6, 14–17]. The first two methods have been studied extensively, but little has been done with regards to puzzle-based learning [18, 19].

Our research aims to assess the impact of puzzle-based learning on conventionally valued academic knowledge such as: course specific material and course specific conceptual understanding, as well as broadly valued skills, such as conceptual reasoning independent of course material. Non-traditional lecture, memorization, and testing methods enable in-class activities to engage students in a wide range of intellectual skill sets both conventionally academic and broadly applicable [20–22]. In-class activities that use puzzle-based teaching can require students to connect to their prior knowledge, explain, interpret, and apply newly acquired knowledge, as well as develop personal perspective and understanding [4, 5].

The objective of this research was to compare the effects of traditional teaching methods and puzzle-based teaching on authentic intellectual performance, and learning outcomes.

Methods

The Internal Review Board at the University of Akron deemed acquisition of all data involving human participants exempt from internal review board approval as the proposed methods did not pose a threat of harm to the participants, the methods were in keeping with standard educational practice, and the identities of each participant remained anonymous throughout the study. The participants chosen for this research were students taking the Human Anatomy and Physiology II Lab at the University of Akron. Seven sections of this lab, 185 students in total, were taught using standard teaching practices as well as puzzle-based methods, and their level of authentic intellectual performance was measured in addition to their standard course assessments. Two teaching assistants were utilized among seven sections of the course (Fig. 1).

Human Anatomy and Physiology II Lab is a prerequisite for all medical related degrees including but not limited to: nursing, pre-med, exercise science, biochemistry, and biology at the University of Akron. The proportion of students representing these specific programs was not controlled, but the intended majors of each student were documented in case there was a strong program related bias observed in any of the six sections. Such a bias was not apparent in the results and so this was not included as a factor in the final analysis.

Single subject/cross-over

As a condition for conducting this study in the Human Anatomy and Physiology II Lab, it was required by the University of Akron that all the course material be presented in a specific order. This order is designed such that the material prior to the mid-term exam is more challenging than the material after the mid-term exam. Consequently, there could be a bias toward improved scores on standard course assessments in the second half of the semester independent of any manipulations to teaching-style (*i.e.* scores on the final exam may be higher than the mid-term exam).

Knowing that this bias could effect the interpretation of our results by showing improvement in any teaching style used in the second half of the semester, a single subject/cross-over design was implemented (Fig. 1) to compensate for this bias. Each teaching assistant taught two sections of the course using a puzzle-based pedagogy for the first half of the semester. After the midterm practical exam, these sections then switched to a traditional lecture-based pedagogy for the remainder of the semester. The remaining sections started with a lecture-based pedagogy and then moved to a puzzle-based pedagogy after the midterm practical exam.

By using the single subject/cross-over design, it can be determined if the order in which the teaching style was used has a significant effect on student performance. Such significance would potentially confirm the suspected bias towards improved scores on standard course assessments in the second half of the semester, and would be the result of a decreased difficulty in the course material. Additionally, by using this design, it can be determined if teaching-style, independent of the order in which the teaching-style was used, had a significant effect on student performance. Taken together, this means that if both order and teaching-style are significant that there is a bias towards improved scores in the second half of the semester, and that teaching-style did have a significant effect on student performance independent of this bias.

Traditional/lecture design

The traditional lecture-based portion of the course used a projector screen and Microsoft Power Point to display

Human Anatomy and Physiology II Lab					I		
Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7	II
Teaching Assistant 2				Teaching Assistant 1			
Lecture		Puzzle		Lecture		Puzzle	
Week 1	Course Introduction Lecture chapter 1 Prelab assignments chapter 1	Course Introduction Puzzle chapter 1		Course Introduction Lecture chapter 1 Prelab assignments ch. 1	Course Introduction Puzzle chapter 1		
Week 2	Lecture chapter 2 Prelab assignments chapter 2	Puzzle chapter 2		Lecture chapter 2 Prelab assignments ch. 2	Puzzle chapter 2		
Week 3	Lecture chapters 4, 5 Prelab assignments chapter 4, 5	Puzzle chapters 4, 5		Lecture chapters 4, 5 Prelab assignments ch. 4, 5	Puzzle chapters 4, 5		
Week 4	Quiz chapters 1, 2 Lecture chapter 6 Prelab assignments chapter 6	Quiz chapters 1, 2 Puzzle chapter 6		Quiz chapters 1, 2 Lecture chapter 6 Prelab assignments ch. 6	Quiz chapters 1, 2 Puzzle chapter 6		III
Week 5	Quiz chapters 4, 5 Lecture chapter 7 Prelab assignments chapter 7	Quiz chapters 4, 5 Puzzle chapter 7		Quiz chapters 4, 5 Lecture chapter 7 Prelab assignments ch. 7	Quiz chapters 4, 5 Puzzle chapter 7		
Week 6	No labs this week due to national holiday	No labs this week due to national holiday		No labs this week due to national holiday	No labs this week due to national holiday		
Week 7	Quiz chapters 6, 7 Midterm review / mock practical	Quiz chapters 6, 7 Midterm review / mock practical		Quiz chapters 6, 7 Midterm review / mock practical	Quiz chapters 6, 7 Midterm review / mock practical		
Week 8	MID-TERM EXAM MID-TERM EXAM MID-TERM EXAM						IV
★ Point of Pedagogy Cross-over	Puzzle	Lecture		Puzzle	Lecture		
Week 9	Puzzle chapter 8	Lecture chapter 8 Prelab assignments chapter 8		Puzzle chapter 8	Lecture chapter 8 Prelab assignments chapter 8		
Week 10	No labs this week due to Spring break	No labs this week due to Spring break		No labs this week due to Spring break	No labs this week due to Spring break		
Week 11	Puzzle chapter 9	Lecture chapter 9 Prelab assignments chapter 9		Puzzle chapter 9	Lecture chapter 9 Prelab assignments chapter 9		V
Week 12	Quiz chapters 8, 9 Puzzle chapter 10	Quiz chapters 8, 9 Lecture chapter 10 Prelab assignments chapter 10		Quiz chapters 8, 9 Puzzle chapter 10	Quiz chapters 8, 9 Lecture chapter 10 Prelab assignments chapter 10		
Week 13	Puzzle chapter 11	Lecture chapter 11 Prelab assignments chapter 11		Puzzle chapter 11	Lecture chapter 11 Prelab assignments chapter 11		
Week 14	Quiz chapters 10 & 11 Puzzle chapter 12	Quiz chapters 10 & 11 Lecture chapter 12 Prelab assignments chapter 12		Quiz chapters 10 & 11 Puzzle chapter 12	Quiz chapters 10 & 11 Lecture chapter 12 Prelab assignments chapter 12		
Week 15	Quiz chapter 12 Final review / mock practical	Quiz chapter 12 Final review / mock practical		Quiz chapter 12 Final review / mock practical	Quiz chapter 12 Final review / mock practical		
Week 16	FINAL EXAM FINAL EXAM FINAL EXAM						

Fig. 1 Course Structure and Single Subject/Cross-Over Design. A summary of the Human Anatomy and Physiology II Lab course schedule including the implementation of a single subject/cross-over design. I) Course offered at the University of Akron, II) Individual sections or "classes" each approximately 25–30 undergraduate students (185 in total). Each section was assigned to meet at the same unique specified time and classroom on campus at the University of Akron. The actual meeting times, and section numbers have been omitted to ensure anonymity of the participants in keeping with the requirements set forth by the Internal Review Board at the University of Akron. III) Weeks 1–8 of an academic semester, including indications of how sections were divided among the teaching assistants, and which sections began the study with either a traditional lecture or puzzle-based pedagogy. IV) The mid-term exam and, as indicated by the black five pointed star, the point of pedagogy cross-over. V) Weeks 9–16 of an academic semester, including indications of which sections cross-over to either a traditional lecture or puzzle-based pedagogy

images and key information related to the week's lab. The teaching assistant read through a prepared set of notes related to the content of the chapter, as they displayed the slide show. Figure 2 is an example of a PowerPoint slide and some of the notes that were presented to the class. Once the lecture was over, the students had the remainder of the class period to complete a prelab activity (Fig. 3) usually consisting of approximately twelve fill-in-the-blank questions, using their lab manual and the models in the classroom.

The students were asked to briefly describe any additional applications of the concepts learned in that week's activity, as a measure of authentic intellectual performance. Every picture in the lab manual had one or more models of the same anatomical structure physically present and available during the class period. At the end of the class period, each student's prelab was graded for correctness and returned to the student so that they could study from it in the following week.

Puzzle design

Content specificity, and conceptual generality

As previously stated, a key feature of puzzle-based learning that distinguishes it from problem-based or project-based learning is the emphasis on domain-independent critical thinking and abstract reasoning [6, 16, 17]. This is not to say that a puzzle-based learning course is about presenting and discussing a variety of puzzles, but rather that it is about presenting and understanding

problem-solving principles through puzzles that serve both as an illustration of broad concepts and of course specific material [6, 17].

This study's puzzle design used the Human Anatomy and Physiology II Lab course content as a vehicle for the development of domain-neutral basic reasoning skills. In general, the participants would have to familiarize themselves with a set of "rules" given in the form of background information regarding a real-world medical application of the course material.

For example if the chapter material focused on the lymphatic system, a real-world medical application for the lymphatic system was the diagnosis of lymphoma (a cancer specific to the lymphatic system) through histological test results. As previously mentioned, the "rules" for this puzzle were not directly stated; the participant would come to know them by a conceptual understanding of the chapter material in the lab manual, and any supplementary material provided in the form of a photocopied handout on the day of class. In the case of the lymphoma puzzle, such an understanding should have yielded the following rules:

- 1) The cancer will only affect anatomical structures specific to the lymphatic system.
- 2) The lymphatic system has a particular directional flow by which the cancer can spread.
- 3) When viewed with basic staining techniques, some types of lymphatic cancer cells have one distinctive

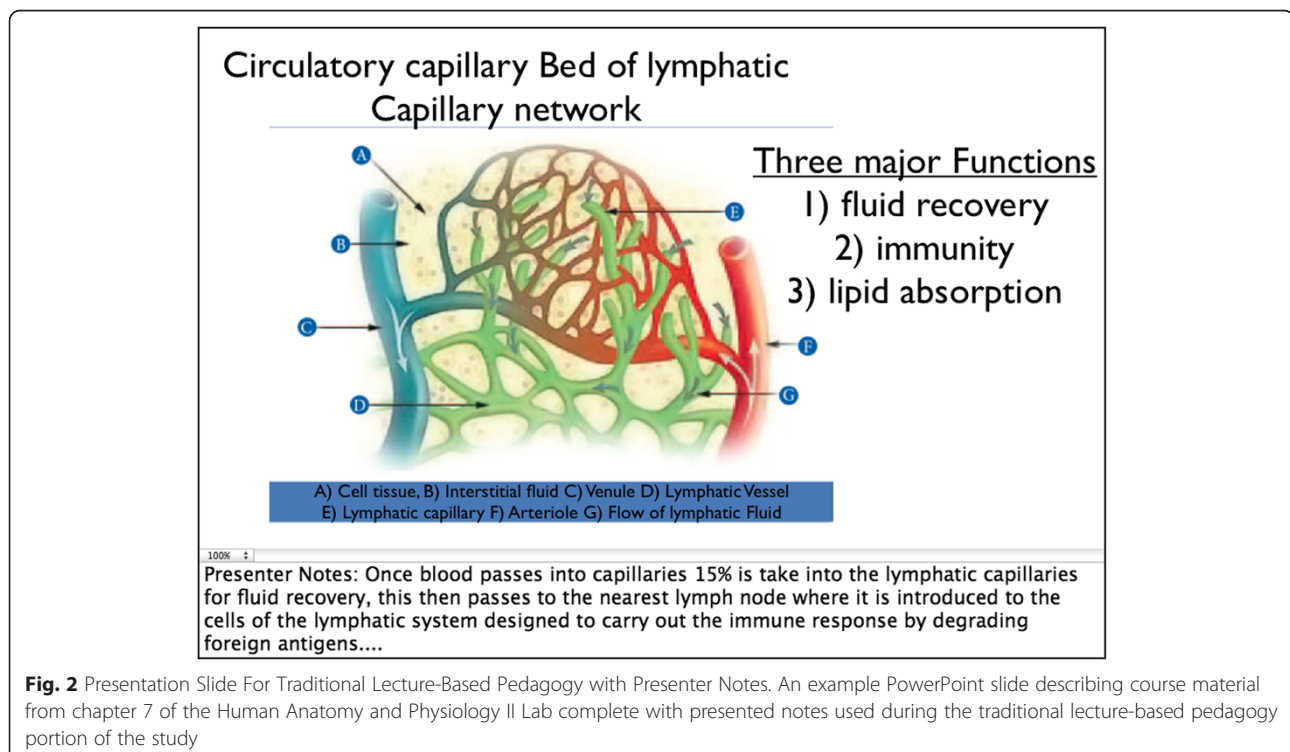



Fig. 2 Presentation Slide For Traditional Lecture-Based Pedagogy with Presenter Notes. An example PowerPoint slide describing course material from chapter 7 of the Human Anatomy and Physiology II Lab complete with presented notes used during the traditional lecture-based pedagogy portion of the study



7 CHAPTER Prelab

Name: _____
 Lab Day/Time: ____ / ____ Section#: _____

- 1) Helper T cells (TH) secrete _____ in response to non-self antigens and are necessary for most immune responses.
- 2) A lymphatic capillary with large opening involved in lipid absorption is called a:
 - A. Lacteal
 - B. Lymphocyte
 - C. Occipital
 - D. None of the above
- 3) T or F : The three primary functions of the lymphatic systems are fluid recovery, immunity, and lipid absorption.
- 4) Lipid absorption occurs in the _____ within villi of the small intestine.
- 5) In order for cytotoxic T cells to function they must be activated by:
 - A. APC
 - B. Monocyte
 - C. Eosinophil
 - D. None of the above
- 6) Elephantiasis is an extreme example of _____ as a result of blockage of lymph flow.
- 7) The _____ is the largest lymphatic organ.
- 8) _____ are phagocytic machines that also play a role in the immune system by acting as an antigen presenting cell (APC).
- 9) About how many lymph nodes does a person have? _____.
- 10) _____ filter lymphatic fluid returned from tissue and contain lymphocytes, plasma cells, macrophages and reticular cells.
- 11) Name the six major types of lymphocytes. _____

- 12) Where are most of the formed elements produced? _____.

Fig. 3 In-class Prelab Assignment For Traditional Lecture-Based Pedagogy. The prelab activity for chapter 7 of the Human Anatomy and Physiology II Lab, for use in the traditional teaching portion of the class

morphology, are highly metastatic, and thus have the ability to spread throughout major anatomical features, while other types of cancer cells have a different type of morphology and are only able to spread through minor substructures of the lymphatic system.

While this may appear to be a domain specific problem, the puzzle “rules” directed the participants to develop a set of basic reasoning skills independent of the chapter content. Stated more abstractly, the rules could be simplified to:

- 1) This is a closed system, the limits of which are defined.
- 2) This closed system has sequential order through which it functions on two size dependent scales.

- 3) Idea 1 affects the system on one size scale; Idea 2 affects the system on a different size scale. Idea 1 and 2 are mutually exclusive.

In-addition to learning critical thinking and basic reasoning skills, a puzzle-based in-class activity reinforces the specific chapter material by putting it into a contextual framework. Where a traditional lecture would state the directional flow of the lymphatic system on two size scales, its various anatomical structures, and its physiological relevance, a puzzle-based in-class activity reinforces that same information by making it impossible to complete the assignment without a conceptual understanding of the material.

Assessments of authentic intellectual performance (AIP)

At the end of the puzzle activities and the prelab activities, participants were asked to identify any additional

applications of the concepts from that activity's chapter material. This provided a means through which to assess the participants' broader conceptual understanding of the material. This broader conceptual understanding should, in theory, translate to "real-world" aptitude [9–12]. These assessments were scored using the authentic intellectual performance criteria outlined by Newmann and Wehlage [10]. These standards were developed in order to represent the quality of intellectual work without being linked to any one teaching method [8, 13, 10]. Each of the following standards; higher-order thinking, depth of knowledge, connectedness to the world beyond the classroom, substantive conversation, and social support for student achievement; was designed as a continuous, analog representation of quality.

In general, scores were given on a scale of 0–5 in which answers that demonstrated little conceptual understanding beyond the domain specific context of the puzzle material were given a 1–2, answers that demonstrated conceptual understanding beyond the domain specific content but lacked an application of the concept outside of the domain specific content received a 3, answers that demonstrated a conceptual understanding beyond the domain specific content and an application of the concept that was beyond the domain specific content of the puzzle but were still within the domain specific content of the course received a 4, and answers that demonstrated a conceptual understanding beyond the domain specific content and an application of the concept that was beyond the domain specific content of the puzzle and the course received a 5.

For example, in the puzzle about the lymphatic system, the participants provided some variation of the following answers and received the following scores:

The Lymphatic system has lymph nodes all over the body that move lymphatic fluid and nutrients in the cells. (Score 1 or 2)

The lymphatic system is a mechanism for biological cleaning and nutrient transport. (Score 3)

The lymphatic system moves waste out of the cell and moves nutrients into the cell, much the way the act of eating and digestion moves nutrients into the body and waste out of the body (score 4)

The lymphatic system is like the water treatment center for a city, it moves clean fluid into homes and keeps the residents alive, similar to how the lymphocytes keep cells of the tissue alive, and then the pipes connected to the sewage system remove waste and keep people in their homes clean, the way lymphatic fluid removes wastes from cells. (Score 5)

If the student failed to answer the question they received no credit.

One important feature of this assessment is that it demonstrates a broader conceptual understanding independent of the puzzle content [10]. We assume that this broader understanding would be difficult to achieve without the use of basic reasoning. This assumption implies that if a puzzle/project/problem requires the participant to use and develop basic reasoning, this basic reasoning will facilitate the connection of domain specific concepts to domain independent concepts [6, 16, 17].

Standard course assessments

On the days the class had a quiz, the quiz tested a subset of course material defined by the course syllabus, and then the day's activity began once all the quizzes were collected by the teaching assistant. The quiz format was specific to the structure of the Human Anatomy and Physiology II Lab course, prior to the inclusion of this study. Students were presented with a Power Point slideshow in which each slide had one of the figures from the student's lab manual with its structures labeled by an alphabetic letter. The teaching assistant then informed the class which letters needed to be identified as anatomical structures, or asked a question relating to physiological function and included what number-line they should write the anatomy term next to on their quiz sheets. The assessments had a total of 30 questions, and in general 5–8 of these questions would be related to a physiological concept while the remainder of the questions were anatomical identification.

The mid-term and final lab practical exams were also a format specific to the structure of the Human Anatomy and Physiology II Lab course, prior to the inclusion of this study. In these assessments, a total of 20 anatomical models or histological images were assigned individual stations throughout the exam room. Students completed a total of 2, 1-min rotations for each station in sequential order. Each station had 2–3 questions that either required the student to identify a particular anatomical structure indicated by a numbered sticker on the model or image, or to answer a question about physiological function related to that station's model or image. The assessments had a total of 50 questions, 30 of which were anatomical identification questions, and 20 of which were physiological function questions.

All standard course assessments were graded for correctness. To receive full credit, the answers needed to be complete and all vocabulary terms needed to be spelled correctly. Incorrect spellings different by a single vowel received $\frac{3}{4}$ credit; incorrect spellings by a non-sensational consonant (C instead of K or S, etc.) received $\frac{1}{2}$ credit. All other misspellings received no credit.

Statistical analysis

Analyses were done with IBM SPSS Statistics version 21. Multivariate analysis of variance (MANOVA) was used to determine if there were between-subjects effects of pedagogical styles on the scores of standard course assessments. MANOVA was used because this method can analyze significant differences in student performance on standard course assessments as single factor, effectively combining the tests and quizzes to look for an overall effect with regards to teaching-style, order, and teaching-style + order (Table 1). A similar MANOVA analysis that separated the standard assessments into quizzes and tests was used because this method can analyze significant differences in student performance on tests and quizzes separately to look for an effect with regards to teaching-style, order, teaching-style + order (Table 2). Additionally, this analysis was split by TA to determine the effect of subtle difference in instruction on teaching-style, order, and teaching-style + order (Table 3). For the AIP assessments a univariate analysis of variance was used to determine between-subjects effects with regards to teaching-style, order, and teaching-style + order. For all analyses results were considered statistically significant when $p < 0.05$.

It is important to note that the significant effect of order, and teaching-style + order demonstrated a statistically significant improvement in student performance on standard course assessments in the second half of the semester (data not shown). However, the single subject/cross-over design of the study removes this bias when analyzing the effect of teaching style exclusively.

Results

Overall, our findings indicate a significant improvement in students' performance on standard course specific assessments using a puzzle-based pedagogy versus a traditional lecture-based teaching style. Quiz and test scores for students improved by 2.1 and 0.4 % respectively in the puzzle-based pedagogy, versus the traditional lecture-based teaching. Mean quiz and test scores for puzzle-based students were 85.2 and 77.7 respectively, while the mean quiz and test scores for traditional lecture-based teaching were 83.1 and 77.3 respectively (Fig. 4). The significance of teaching-style independent

of order in combination with the cross-over design suggests that the significant bias of order was effectively compensated for (Table 1). If the bias of order were not effectively compensated for there would be no significance in teaching-style independent of order as scores on all standard assessments increased in the second half of the semester independent of teaching-style. Interestingly when analyzing the effect of teaching-style on the specific outcomes of tests and quizzes separately neither is significant (Table 2). However, when these results are further analyzed for the effect of teaching-style on the specific outcomes of test and quizzes with respect to individual TAs the test scores for TA2 remain significantly improved by 3.04 % (Fig. 5, Table 3).

Additionally, in terms of the assessments for authentic intellectual performance (AIP), our findings indicate a significant difference ($p = 0.001$) between puzzle-based and traditional lecture-based pedagogies (Fig. 6). However, based on anecdotal accounts of interactions between the teaching assistants and participants, we also report that these assessments contained a number of complicating factors that may have contributed to a fundamental inaccuracy of the significance in this study, and failure to assess the participant's potential for broader reasoning and conceptual understanding independent of domain-specific content.

Discussion

Taken together, the results of the assessments for learning outcomes and authentic intellectual performance help to inform the relative effectiveness of puzzle-based pedagogies as a tool for "open" or "experiential" learning theories in comparison with more didactic theories, the primary vehicle of which is a traditional lecture based pedagogy [8]. Previous studies have demonstrated the effectiveness of implementing "open/experiential" learning through a puzzle-based pedagogy in mathematics and computer science courses [6, 16, 17]. This study expands the academic context of "open/experiential" learning theories into a Human Anatomy and Physiology II Lab that is traditionally assumed to be rooted in a didactic theory, through dry anatomical identification and rote memorization. By moving "open/experiential" learning into this new academic territory it could be possible, as previous studies have suggested, that educational techniques such as puzzle-based and problem-based pedagogies are effective in a broad range of academic contexts [8]. The findings of this study suggest a significant improvement in performance with puzzle-based versus traditional lecture-based pedagogy, on standard Human Anatomy and Physiology II Lab assessments. Additionally, there is an arguably unreliable significance in assessments of authentic intellectual performance.

Table 1 MANOVA comparing scores on standard assessments between pedagogies

Effect	Wilks' Lambda	F	Hypothesis df	Error df	p-value
Teaching Style	.95	5.588 ^a	3.000	298.000	0.001*
Order	.88	13.459 ^a	3.000	298.000	0.000*
Teaching Style + Order	.83	16.353 ^a	3.000	298.000	0.000*

*significant at $p < 0.05$

^a exact statistic

Table 2 MANOVA comparing scores on standard assessments between pedagogies with tests for between-subjects effects on quizzes and tests

Multivariate tests					Between subject effects	
Effect	Wilks' Lambda	F	Hypothesis df	Error df	Dependent variable	p-value
Teaching style	.95	5.588 ^a	3.000	298.000	Quizzes	0.876
					Tests	0.846
Order	.88	13.459 ^a	3.000	298.000	Quizzes	0.006*
					Tests	0.001*
Teaching Style + Order	.829	16.353 ^a	3.000	298.000	Quizzes	0.000*
					Tests	0.186

*significant at $p < 0.05$

^a exact statistic

One explanation for the significant improvement in performance on the standard assessments between puzzle-based and traditional lecture-based pedagogies is that the puzzle-based pedagogy creates a contextual framework for the course material that improves students' ability to recall specific details. Previous studies have demonstrated the increase in student performance using a puzzle-based pedagogy in advanced mathematics courses and computer science courses [6, 16, 17]. The assessments for these courses require students to have a conceptual understanding of the material in order to grasp challenging abstract concepts.

The standard assessments for a human anatomy and physiology lab, however, are ~80 % anatomical identification and ~20 % physiology. Neither set of questions in the Human Anatomy and Physiology II Lab requires a deeper conceptual understanding of the material in order to perform well on the standard assessments. Rather, the assessments are geared toward rote memorization, devoid of conceptual understanding. Despite the vast differences between Human Anatomy and Physiology II Lab and the previous studies that have implemented puzzle-based learning in advanced mathematics and computer science courses [6, 16, 17] in

terms of course design and assessments, the contextual framework established by the puzzle-based pedagogy seems to aid in the recall of specific details improving student performance when compared to traditional lecture-based teaching.

Importance of individual characteristics

The significance of pedagogy in teaching assistant #1's sections presents a very important observation that is intuitively obvious and methodologically difficult to demonstrate. This observation being that the individual characteristics of the teaching assistant significantly impacts the performance of students. In this study the individual characteristics that may have affected students' performance are: familiarity with the course material, approachability, and style of fielding questions during class activities.

While neither teaching assistant had presented an in-class lecture or a puzzle activity for this course, Teaching assistant #2 (TA2) had previously taught this course once prior to this study. Additionally, because TA2 had prior experience with the material, their intellectual input was critical to the design of each in-class activity. Though both TAs met prior to instructing the chapter

Table 3 MANOVA comparing scores on standard assessments between pedagogies for ta1 and ta2 with tests for between subject effects on quizzes and tests

Multivariate tests										Between subject effects			
Effect	Wilks' Lambda		F		Hypothesis df		Error df		p-value		Dependent variable	p-value	
	TA1	TA2	TA1	TA2	TA1	TA2	TA1	TA2	TA1	TA2		TA1	TA2
Teaching style	.98	.97	1.421 ^a	3.016 ^a	2	2	167	183	.244	.005*	Quizzes	.107	.758
											Tests	.702	.041*
Order	.96	.86	.959 ^a	15.024 ^a	2	2	167	183	.030*	.000*	Quizzes	.205	.000*
											Tests	.159	.000*
Teaching Style + Order	.90	.85	.902 ^a	15.698 ^a	2	2	167	183	.000*	.000*	Quizzes	.000*	.205
											Tests	.068	.000*

*significant at $p < 0.05$

^a exact statistic

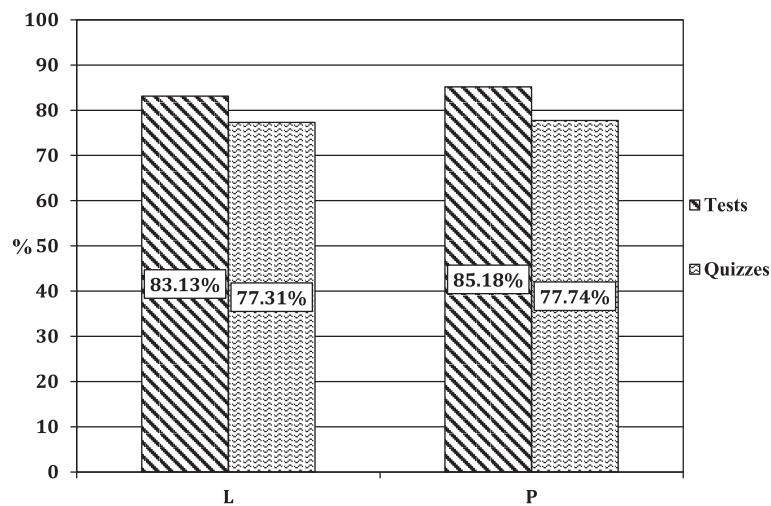


Fig. 4 Standard Course Assessments. Mean scores for standard assessments. Both tests and quizzes showing a significant improvement ($p < 0.05$) in the puzzle-based pedagogy (P) versus traditional lecture-based teaching (L). Tests are an average of final and mid-term exams. $N = 185$

material for that week to ensure there was a uniformity in instruction, it is possible that TA2's familiarity with that material may have influenced the ability answer activity related questions. Though both TAs did not directly provide any student with answers to any in-class activity, a greater familiarity with the material could possibly have give TA2 the ability to direct students into making conceptual connections more effectively. This is supported by anecdotally reported interactions between numerous participants and the lab coordinator/instructor, in which participants expressed a frustration and dissatisfaction with TA1's level of preparedness and familiarity with in-class activities.

Motivational effect of course expectation

The focus of Human Anatomy and Physiology Labs on anatomical identification and rote memorization is not only something that is commonly accepted and understood, but is expected by students enrolling in the course. Based on anecdotal reports of student-participant teaching assistant interactions, it was clear that the majority of students felt that the puzzle-based in-class activities were excessively difficult, and added little benefit to their understanding of the material. Many participants argued at great length with their teaching assistant about the puzzle activities and consistently expressed the view that Anatomy and Physiology was a

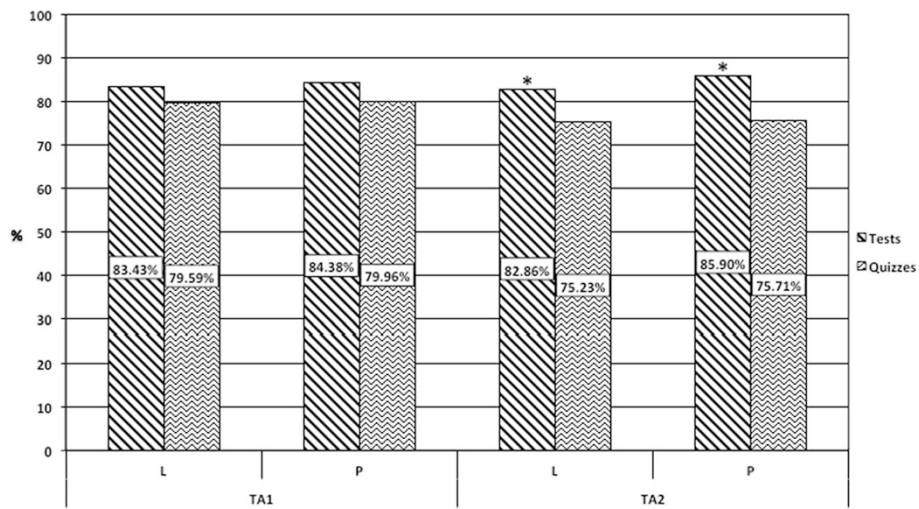


Fig. 5 Standard Course Assessments Separated by Tests, Quizzes, and TA. Mean scores for standard assessments. Tests for TA2 showing a significant improvement ($p < 0.05$) in the puzzle-based pedagogy (P) versus traditional lecture-based teaching (L). Tests are an average of final and mid-term exams. TA1 $N = 97$, TA2 $N = 88$

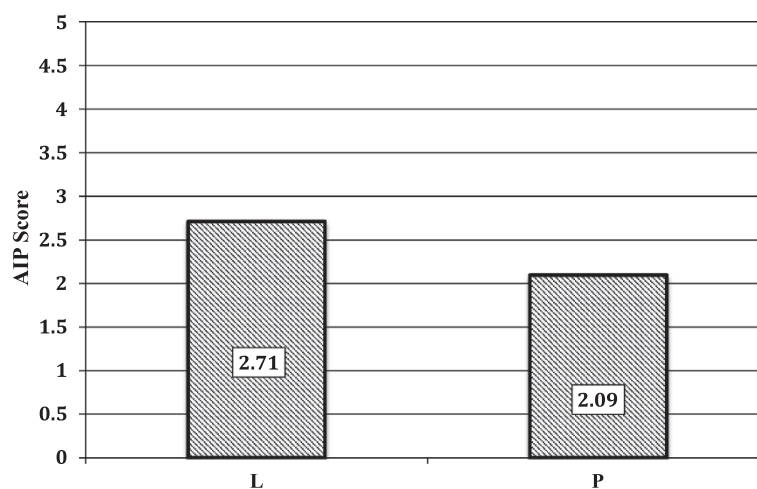


Fig. 6 Assessment Authentic Intellectual Performance. Mean scores for AIP assessments showing a significant difference ($p < 0.05$) between the puzzle-based pedagogy (P) and traditional lecture-based teaching (L). $N = 185$

memorization course. This predominantly negative attitude towards the puzzles contrasts the claims of many other studies that implement puzzle-based learning, and describe it as being “fun”, and possessing an “entertainment factor” [6, 17].

As previously mentioned, the course this study was conducted in was Human Anatomy and Physiology II Lab. Thus, it is important to note that all the participants had completed the Human Anatomy and Physiology I lab without any of the in-class puzzle activities used in this study. This generated a workload expectation that was exceeded by the in-class puzzle activities included in this study, and consequently an inherent motivational resistance to any potential benefit that could be gained through the in-class puzzle activities.

The most striking motivational resistance was apparent in the participants’ overwhelmingly negative attitude toward and effort applied to the AIP assessments. Specifically, the disproportional relation between the degree of perceived effort required to answer the assessment’s question correctly and the credit received for a correct answer. This resulted in many students applying little or no effort to the assessment and in many cases skipping it completely. This was especially true if the participant had applied a self-determined “great deal” of effort on an in-class assignment, and in most instances in which this occurred it was after completing a challenging puzzle activity. This is supported by the percentage of AIP assessments that received a score of “0” (a score of “0” was exclusively used to indicate that the question was left blank). The percentage of AIP assessments with a “0” in the puzzle-based portion of the study was 30.2 %, versus 14.5 % in the traditional lecture-based portion of the study. It is for this reason that we assume there exists a

decrease in the students’ performance on this assessment during the puzzle-based portion of the study. It is possible that this trend in overall decreased performance in AIP scores over time is why other similar studies performed these assessments at only one time point during a semester (*i.e.* once in the spring and once in the fall) [10].

This may suggest that these assessments can only effectively measure a broader conceptual understanding in a limited set of contexts, and not in the context of a Human Anatomy and Physiology II Lab. Alternatively, this may suggest that in order to gain useful information using the AIP assessments that there must be fewer total assessments per individual, and/or that the assessments may need to be conducted such that the degree of perceived effort needed for this assessment is reduced. Furthermore, had this study been conducted in the Human Anatomy and Physiology I lab, it possibly would have eliminated the de-motivational effect of course expectation, as the students would have had less prior experience for comparison, and the benefit of the puzzle-based pedagogy may have been even more dramatically demonstrated.

Conclusion

In keeping with numerous other studies that have highlighted the importance of pedagogical style in various courses, our findings support the critical role of pedagogy in student performance. This study also emphasizes the popular view that a broad contextual framework improves a student’s ability to recall specific details and that the individual characteristics of an instructor play a significant role in utilizing pedagogy to influence students’ performance. Additionally, our findings indicate that there is currently no assessment that can

effectively quantify broader conceptual understanding across multiple fields of academia. In conclusion, a puzzle-based pedagogy, when compared to traditional lecture-based teaching, can effectively enhance the performance of students on assessments, even when the assessments only test a limited conceptual understanding of the material.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LS proposed project and was critical for every aspect of project, including sole authorship of submitted manuscript. LS also designed puzzle activities, as well as in-class lectures. LS acted as a teaching assistant in 4 of 7 course sections and in doing so administered all assessments to participants, graded assessments, graded both puzzle activities and prelab in-class assignments, answered questions from participants during in-class activities, and gave lectures. AD graded assessments, as well as puzzle activities and prelab in-class assignments, ran all statistical analyses, had critical input regarding experimental design, provided intellectual contributions during the initial planning of project, and approved of the final draft of the manuscript. JP acted as a teaching assistant in 3 of 7 course sections and in doing so administered all assessments to participants, graded assessments, graded both puzzle activities and prelab in-class assignments, answered questions from participants, and gave lectures. JP provided intellectual contributions during the initial planning of project, and approved of the final draft of the manuscript. CY provided some intellectual contributions during the initial planning of project, and approved of the final draft of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Funding and Facilitation of Project: The University of Akron, The Integrated Biosciences Program, The Department of Biology.

Mentors: Dr. Peter Niewiarowski, Pro. Jeffery Spencer, Dr. Bruce Cushing

Received: 3 March 2014 Accepted: 11 June 2015

Published online: 23 June 2015

References

- Tanner K, Allen D. Approaches to biology teaching and learning: understanding the wrong answers—teaching toward conceptual change. *Cell Biol Educ*. 2005;4:112–7.
- Boaler J. Alternative approaches to teaching, learning and assessing mathematics. *Eval Program Plann*. 1998;21:129–41.
- Rivers DB. Using a course-long theme for inquiry-based laboratories in a comparative physiology course. *Adv Physiol Educ*. 2002;26:317–26.
- Thomas J. A review of research on project-based learning. San Rafael, CA: The Autodesk Foundation; 2000.
- Tretten R, Zachariou P. Learning about project-based learning: assessment of project-based learning in Tinkertech schools. San Rafael, CA: The Autodesk Foundation; 1997.
- Falkner N, Sooriamurthi R, Michalewicz Z. Puzzle-based learning for engineering and computer science. *Computer*. 2010;20–28.
- Fisher KM, Wandersee JH, Moody D. Mapping biology knowledge. Dordrecht, Netherlands: Kluwer; 2001.
- Frey BB, Schmitt VL, Allen JP. Defining Authentic Classroom Assessment. *Practical Assessments Research & Evaluation*. 2012. p.17.
- Newmann FM, Archbald DA. The nature of authentic academic achievement. In: *Toward a new science of educational testing and assessment*. Albany, NY: State University of New York Press; 1992.
- Newmann FM, Wehlage G. Authentic pedagogy and student performance. *Am J Educ*. 1996;104:280–312. Additional documentation: Five Standards of Authentic Instruction. *Educational Leadership*.
- Newmann FM. Authentic assessment in social studies: standards and examples. In: *Handbook of classroom assessment: learning, adjustment and achievement*. San Diego, CA: Academic; 1997.
- Newmann FM. Research news and comment: an exchange of views on “semantics, psychometrics, and assessment reform: a close look at ‘authentic’ assessments”. *Educ Res*. 1998;27(6):19–20.
- Palm T. Performance assessment and authentic assessment. *Practical assessments research & evaluation*. 2008;13.
- Barron BJS, Schwartz DL, Vye NJ, Moore A, Petrosino A, Zech L, et al. Doing with understanding: lessons from research on problem- and project-based learning. *J Learn Sci*. 1998;7:271–311.
- Kraemer EW, Lombardo SV, Lepkowski J. The librarian, the machine, or a little of both; a comparative study of three information literacy pedagogies at Oakland University. *Coll Res Libr*. 2007;68:330–42.
- Merrick KE. An empirical evaluation of puzzle-based learning as an interest approach for teaching introductory computer science. *IEEE Trans Educ*. 2010;53:677–80.
- Michalewicz M. *Puzzle-based learning: an introduction to critical thinking, mathematics, and problem solving*. Melbourne Victoria Australia: Hybrid Publishers; 2008.
- Greening T, Kay J, Kingston JH, Crawford K. Problem-based learning of first year computer science. In: *Proceedings of the 1st Australasian Conference on Computer Science Education*. Association for Computing Machinery. 1996. p. 13–8.
- Wood DF. Problem based learning. *BMJ*. 2008;336:971.
- Rubinstein J, Dhoble A, Ferenchick G. Puzzle based teaching versus traditional instruction in electrocardiogram interpretation for medical students—a pilot study. *BMC Med Educ*. 2009;9:4.
- Parhami B. Motivating computer engineering freshmen through mathematical and logical puzzles. *IEEE Trans on Educ*. 2009;52:360–4.
- Prethuis W, Bygstad B. Business intelligence in college: a teaching case with real life puzzles. *J Inform Technol Educ: Innovations In Practice*. 2012;11.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

