

The impact of an interdisciplinary learning community course on pseudoscientific reasoning in first-year science students

Timothy M. Franz¹ and Kris H. Green²

Abstract: This case study examined the development and evaluation of an interdisciplinary first-year learning community designed to stimulate scientific reasoning and critical thinking. Designed to serve the needs of scholarship students majoring in mathematics and natural sciences, the six-credit learning community course was writing-intensive and emphasized general scientific reasoning and critical thinking skills. Success of the course was measured using a pre-test/post-test design that assessed students' paranormal beliefs. Outcomes of the study indicated students' paranormal beliefs were significantly lower at the end of the semester than at the beginning, which was used as a surrogate measure of scientific reasoning that was directly relevant to course content. Supplementary analyses demonstrated that their (a) paranormal beliefs were significantly lower than other students and (b) students self-identified the importance of the scientific reasoning skills learned in the course without being prompted on their teacher-course evaluations. The results of this study can inform the design of interdisciplinary, scientific reasoning courses.

Keywords: Pseudoscientific thinking, critical thinking, learning community, scientific reasoning, first-year students

I. Introduction & Background.

Many college professors attempt to promote scientific reasoning and critical thinking within their courses. According to Shermer (2002, p. 18) scientific reasoning is:

A set of methods designed to describe and interpret observed or inferred phenomena, past or present, and aimed at building a testable body of knowledge open to rejection or confirmation.

Further, Halpern (1997, p.4; see also Halpern, 1998), defines critical thinking as:

thinking that is purposeful, reasoned, and goal directed – the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions when the thinker is using skills that are thoughtful and effective for the particular context and type of thinking task.

Clearly, scientific reasoning and critical thinking skills are an essential foundation of skepticism and evidence-based reasoning that are the foundation for science. Thus, a goal of most first-semester introductory science classes is to acculturate students in reasoning based on evidenced obtained through the scientific method (Druger, 2002). While this has always been a component of science education, more emphasis is being placed on understanding the nature of science and scientific reasoning of late (National Research Council, 1998). Even though many of these early scientific-reasoning skills may be generalized beyond a specific scientific discipline,

¹ Department of Psychology, St. John Fisher College, 3690 East Ave., Rochester, NY 14618, tfranz@sjfc.edu.

² Mathematics Department, St. John Fisher College, 3690 East Ave., Rochester, NY 14618, kgreen@sjfc.edu.

one cannot separate scientific content from scientific reasoning and critical thinking (see for example, Nelson, 1999) although critical thinking can be developed supplementally in such a way that it strongly supports deeper learning of scientific content (Adey & Shayer, 1993). Other studies have examined the impact of learning communities on reasoning and critical thinking (e.g., Browne & Minnick, 2005) in general, focusing on learning communities that pair an introductory science course with a course from another discipline. The present study focuses on whether a specially designed learning community for science majors can influence students' pseudoscientific thinking and thus improve their skeptical, scientific, and critical reasoning.

The Science Scholars Learning Community (LC) was a two-course, team-taught program designed for a specific group of scholarship recipients at St. John Fisher College. The LC program at Fisher was structurally similar to those at other school institutions; first-year students were housed in the residence halls near their LC peers so that the students developed a close-knit cohort that offered academic and social support during the adjustment to college. According to Inkleas, Soldner, Longerbeam, and Leonard (2008, p. 502), this LC fell into the category of "small, limited resourced, primarily residential life emphasis" programs.

The students in the Science Scholars Learning Community came to the college with a sound basic high school understanding of science and were specifically selected for the scholarship program based on excellent high school performance in science and mathematics. The course cluster was designed around comparing pseudoscientific thinking and scientific reasoning and emphasized methods of improving scientific reasoning and critical thinking by learning, discussing, and using scientific methods. Thus, our primary research questions for this study were

1. In what ways can this LC impact students' pseudoscientific thinking?
2. How does the pseudoscientific thinking of the students in the learning community compare with the general student population?

A secondary research questions, assuming a positive answer to the first questions was:

3. Do improvements in pseudoscientific thinking translate into improvements in critical thinking and scientific reasoning?

A. Brief Background about Learning Communities.

The Learning Communities National Resource Center (n.d.) offers several models of learning communities, but in general, learning communities are a collection of courses in which a cohort of students participates. These courses can be either loosely connected or tightly linked with integrated content (Smith, MacGregor, Matthews, & Gabelnick, 2004). In some cases, the same group of students is enrolled in several courses, such as calculus and introductory physics, and a seminar, which helps students make connections between the linked courses. A coordinated studies model, on the other hand, places the cohort in a team-taught block that covers the material of several traditional courses with integration among all topics.

These learning communities, which are planned collaboratively among the faculty members involved combine content knowledge with skills practice (Smith et al., 2004). According to Smith and colleagues, learning communities are one of the solutions to recent calls for educational reform, because students are actively engaged and reflective. The reflection in action (e.g., Schön, 1987) helps students to build metacognitive structures necessary for the critical and creative thinking required about academic content. Because of this, learning communities have been used to improve retention, attendance, and social behaviors for first-year

students (Inkelas, Daver, Vogt, & Leonard, 2007). In fact, they have also been used specifically in science and engineering programs to help students make the transition to college (Smith, et al., 2004) and improve general critical thinking skills (Quitadamo, Brahler, & Crouch, 2009).

A similar example of a science-based LC involving the coordinated studies model is provided by Morgan, Carter, Lemons, Grumbling, and Saboski (1995). In Fisher's cohort-based cluster learning community model, courses in the LC are linked by a common theme, and the students are housed together to provide a community of support outside of the classroom. Such living-learning communities tend to promote peer-support networks and enhance campus involvement (Dabney, Green, & Topalli, 2006). At Fisher, each LC fits into one of two formats: (a) a standard first-year composition course paired with one content-area, general education course or (b) a cluster of two content courses taught in a writing-intensive format. The Science Scholars LC is an example of the second model. In either model, all courses in a cluster revolve around a common theme, such as the Vietnam War, the nature of self, the economics of sports, or, as in this case, the nature of scientific thinking. LCs have been a part of the Fisher first-year experience for over a decade. Similar to the findings at other schools, the LCs have resulted in higher first-year retention rates (e.g., Dabney et al., 2006; Dodge & Kendall, 2004), cross-disciplinary communication and collaboration among the faculty, and tighter-knit student cohorts.

One example of a science-based LC involving the coordinated studies model is provided by Morgan and colleagues (1995). The course of study involves a year-long experience combining introductory courses in biology with a literature course and an environmental science course in addition to a one-credit seminar. Their findings showed that students in the LC improved significantly on intellectual development, suggesting that the LC had improved their reasoning and critical thinking skills (see also work by Browne & Minnick, 2005) and students' ability to apply these to decision-making and value judgments.

B. Critical Thinking, Scientific Reasoning, and Pseudoscientific Beliefs.

Halpern's (1997) definition of critical thinking, cited above, is a broad attempt to capture the multitude of thinking skills that might be involved in "critical thinking" by generalizing the concept as dealing with "thinking that is purposeful, reasoned, and goal directed." It seems, though, this view of critical thinking as a generic skill that can be acquired independent of specific content may be flawed. Recent studies have found that many academics describe critical thinking differently (Moore, 2013) using such a variety of descriptors that one is tempted to believe that there is no single elephant being described by all of the blind academics. However, deeper analysis of the concept (Davies, 2013) yields a commonality to all of these definitions that is instantiated differently in different disciplines. Thus, there is a common core of learning to think in a way that prioritizes logic and evidence over instinct, and this core can be approached and learned in a way that allows it to easily transfer across disciplines. At the same time, each discipline has its own criteria for what constitute logic and evidence. In the sciences, critical thinking is instantiated by the term "scientific reasoning" which combines aspects of analysis with specific skills related to experimental design (National Research Council's Committee on Undergraduate Science Education, 1999).

Generically then, critical thinking includes the judgment and a skeptical stance toward evidence that is presented. This allows one to then test to what degree a given assertion is supported by the evidence and reasoning presented, and suggests a tradeoff between learning to

think critically and accepting paranormal beliefs. This tradeoff is partially supported by the literature. For example, past research has examined the relationship between critical thinking and paranormal beliefs, showing in general (cf., Roe, 1999) that critical thinking is negatively related to paranormal beliefs (Aarnio & Lindeman, 2005; Cheung, Rudowicz, Kwan, & Yue, 2002; Messer & Griggs, 1989) or reported past paranormal experiences (Royalty, 1995). Further, past research has demonstrated a negative correlation between paranormal beliefs and reasoning ability (Hergovich & Arendasy, 2005) and that paranormal beliefs are unrelated to age (Aarnio & Lindeman, 2005). Other studies, however, have found that higher levels of education do not necessarily translate into lower paranormal beliefs (Farha & Steward, 2006). This same study also examined paranormal beliefs across the disciplines, finding that students in the sciences fell somewhere in the middle in terms of paranormal beliefs (social sciences had the highest percentage of believers) and that there were virtually no differences attributed to gender. Thus it seems that even further training within a scientific discipline, which one expects to focus on scientific reasoning, is not sufficient to eliminate a student's beliefs in paranormal (or more broadly, pseudoscientific) phenomena, suggesting that individuals possess the capacity to engage in different modes of thinking selectively even though the selection may not be made a conscious level. This also suggests that without explicitly applying the generic skill of critical thinking to the specific contents of pseudoscience and the paranormal, students may develop strong scientific reasoning and critical thinking skills while still holding uncritical beliefs in phenomena such as spontaneous human combustion and fortune telling.

C. College and Student Profiles.

St. John Fisher is a private liberal arts college in the Catholic tradition located between the city of Rochester, New York, and the eastern suburbs. Admission is competitive, and the majority of students are drawn from a 100-mile radius of the campus, resulting in a student population mostly from upstate New York. Many students come from the outlying rural school districts, and many are drawn to Fisher because of the Catholic heritage of the school. Most of our Science Scholar Learning Community students fall close to national and state averages in academic achievement in high school (e.g., mean national high school GPA of 88% versus 94% for the science scholars). At the time of this study, students in the top two tiers of scholarships with an interest in biology, chemistry, computer science, mathematics or physics are also offered the opportunity to apply to the Science Scholars Program. Most students have completed several Advanced Placement courses in mathematics and science or have participated and received credit for other college-credit courses before high school graduation. Almost all of the Science Scholars have completed four years of mathematics and science in high school and are academically motivated students with a strong interest in the sciences. Biology and chemistry are the most common major choices followed by mathematics. The remaining minority of students are equally split between computer science and physics.

D. Course Background and Design.

The Science Scholars Learning Community was a writing-intensive, two course (6-semester hour) experience. The explicit connection between writing and science helps improve students' scientific writing (Kokkala & Gessell, 2002) and provides them with valuable practice applying the tools of the course independently. Rather than give two separate grades for the two courses in

which the students were officially registered, students received the same grade for both courses, and both instructors participated in reading all major papers. In addition, a substantial component of the course grade was based on in-class and small writing assignments, as well as a group research project that culminated in a large team-run debate.

Unlike many of the LCs at Fisher, the Science Scholar Learning Community was not composed of existing first-year courses like introductory psychology or chemistry 101. Thus, there were no departmental expectations to provide the foundation for future courses in a discipline. Instead, two new courses were created to allow the Science Scholar Learning Community to explore the nature of science independent of a particular discipline. One course was listed in the interdisciplinary studies program while the other was listed under the mathematics, science, and technology integration program. While neither course counted for credit in any major, both courses provided students with credit towards meeting the college core requirements. Together, these courses explored the nature of science by comparing scientific fields and scientific thinking with pseudoscience and science fiction topics. One of the faculty members in this LC normally taught in the Department of Psychology, and the other in the Mathematical and Computing Sciences department. Thus, by design, students were exposed to different perspectives on scientific and critical thinking. The faculty members met regularly to discuss the course and the students and to ensure that each class period connected to the previous and subsequent periods and maintained the structure of the LC.

This also meant that each of the freshman-level writings completed during the semester was directly tied to all LC content rather than being unrelated at times. For example, when students were introduced to the idea of critically analyzing evidence, we also discussed how to summarize sources and properly cite them, with the student practice being tied to their particular paper. During the semester, both instructors worked on the writing topics with the students, as appropriate to the flow of the LC and student needs. Both instructors used writing assignments to help students practice the skills and thinking that were the goals of the LC.

The course had two primary texts. The first provided an outline of skepticism, scientific reasoning, and critical thinking, in contexts ranging from alien abduction scenarios and witch hunts to psychic powers and holocaust deniers. Shermer's (2002) *Why People Believe Weird Things* provided an outline of skepticism, scientific reasoning, and critical thinking in a wide range of contexts. The second text, *Taking Sides: Clashing Views on Controversial Topics in Science and Technology* (Easton, 2002, 2004), was a collection of pro and con essays on a variety of science and technology issues. These two books were supplemented with newspaper articles, magazine articles, Internet sources, videos, and other media as appropriate. Throughout the experience, students were expected to locate and evaluate additional resources, using the information literacy skills emphasized in the LC program at Fisher. In addition, we used two separate writing references (Aaron, 2003; Lester & Lester, 2003).

E. Course Structure and Implementation.

The students in the Science Scholar Learning Community met for approximately three hours on Tuesday and three hours on Thursday afternoons. Each class consisted of a variety of pedagogical modalities, including short lectures, class discussions, student presentations and debates, in-class assignments, and/or watching films or film clips that supported course content. During the term, students explored and wrote four papers, each dealing with different scientific topics. We began the course each semester by using a demonstration of psychic powers and tied

this demonstration to a unit of introductory scientific reasoning. The other instructor served as a confidant in these tricks, encouraging the students to take on evidence presented and think critically about it, and having them list various things they could change about the demonstration. The goal was for the students to think critically while the “psychic” tried to adapt as needed to overcome their suggestions and tests. In each year of the LC, the students eventually found a controlled change in the demonstration which negated the abilities for the instructors to successfully complete the trick.

All of this served to accomplish one main learning goal: it established the course content as learning how to think about the world critically. Students often reported that this one activity, on the first day of the semester, made a considerable difference for them in the ability to begin to think critically about information. Throughout the course, the instructors provided experiences to support the development of critical thinking in a structure similar to that in cognitive acceleration (Adey & Shayer, 1993) through stages of concrete preparation, cognitive conflict, metacognition and bridging. Concrete preparation was provided by connections to real situations and experiences, such as the introductory psychic demonstration. Cognitive conflict was generated through readings, discussions and various media presentations exposing students to multiple components and perspectives on the issues. The writing component of the LC provided opportunities for metacognition to be manifested, and parallels among the various situations and the common ways in which thinking goes wrong served as a bridge throughout the course.

The course structure used active learning principles to help to improve learning (Yoder & Hochevar, 2005). For example, students were frequently engaged in small group discussions and short presentations to the class during class time. Almost every class period involved some form of informal or formal writing, usually a short paragraph or so related to the current discussion, either to prime the students before the discussion or to summarize their ideas after. Some of these focused specifically on improving their writing skills, such as revising a particular paragraph or sentence of their work, or re-writing a paragraph using a different voice.

II. Method.

Clearly, the class was designed to stimulate critical thinking and scientific reasoning. During two of the four years that we co-taught the course, we assessed whether students’ level of reasoning improved. To do so, we conducted a pre-test/post-test nonexperimental design using the Paranormal Beliefs Scale (Tobacyk & Milford, 1983) to assess their level of reasoning about paranormal phenomena. We also compared students’ scores on the Paranormal Beliefs Scale to a) a control group of means from students in the original Tobacyk and Milford (1983) study, and b) a control group of students at the college. Finally, we conducted a qualitative analysis to examine students’ unprompted reports of learning about critical thinking and scientific reasoning through the comments on our teacher-course evaluations.

A. Design and Participants.

Our design met the criteria of a quasi-experiment (see Cook & Campbell, 1979). The key participant group included the students from two cohorts of the LC. To learn about content-specific paranormal and pseudoscientific critical thinking, we collected information from the first-year Science Scholars students in the class during two different years. In the first year of the study (2003, the second year we taught the class), we collected data from 22 of 23 enrolled

students in the class ($n = 14$ women, $n = 9$ men), for a 96% response rate. In the second year (2004, the third year we taught the class), we collected data from all 21 enrolled students ($n = 14$ women, $n = 7$ men), for a 100% response rate.

The main phase of the study utilized a pre-test/post-test design. Specifically, responses to the Paranormal Beliefs Scale (PBS, Tobacyk & Milford, 1983) at the beginning of the semester were compared to responses to the PBS at the end of the semester. Data from the key participants were also compared to several other data sets: (a) normative data from Tobacyk and Milford and (b) posttest data from faculty and other students from a second phase of the study (spring 2007).

B. Measures.

The key surrogate measure of reasoning for the study was the 25-item Paranormal Beliefs Scale created by Tobacyk and Milford (1983). We used this scale because our curriculum directly compared pseudoscientific claims (e.g., psychic powers and alchemy) to what is known in their respective scientific fields (e.g., psychology and chemistry). The Paranormal Beliefs Scale provides one complete measure of paranormal beliefs that is based on an average of the items. It also includes seven subscales, including traditional religious beliefs (e.g., *The soul continues to exist though the body may die*), psi (e.g., *Mind reading is not possible*), witchcraft (e.g., *Black magic really exists*), superstition (e.g., *Black cats can bring bad luck*), spiritualism (e.g., *Reincarnation does occur*), extraordinary life forms (e.g., *The Loch Ness monster of Scotland exists*), and precognition (e.g., *Astrology is a way to predict the future*). Most measures had sufficient inter-item reliability; those that did not may be due to the small sample sizes used to calculate the α coefficients (see Table 1).

Table 1. Cronbach’s α coefficients for paranormal beliefs scale and each subscale as a function of sample and pre-test/post-test administration.

	Sample: Science Scholar Students 2006/07		
	Pre-test α ($n = 43$)	Post-test α ($n = 43$)	Students ($n = 228$)
Entire Paranormal Beliefs Scale (25 items)	.85	.89	.92
Traditional Religious Beliefs scale	.90	.88	.85
Psi subscale (4 items)	.81	.70	.84
Witchcraft subscale (4 items)	.69	.80	.84
Superstition subscale (3 items)	.85	.88	.81
Spiritualism subscale (4 items)	.78	.80	.80
Extraordinary Life Forms subscale (3 items)	.88	.95	.90
Precognition subscale (3 items)	.71	.54	.75

In all of the groups, we had acceptably low measures of skewness and kurtosis (all < 1 , well under the generally accepted minimum of 2), and the variances in different groups all had acceptably low F_{\max} test results (all < 2 , well under the generally accepted cutoff of 3). Further, though the students in this design are not fully independent because they were in the same class, it is common for researchers to assume independence in evaluations of classroom behavior because the students are working independently. Thus, the study appropriately met the requirements of normality, heterogeneity of variance, and independence necessary to conduct t -tests and ANOVAs.

C. Procedure.

The Paranormal Beliefs Scale was administered at the beginning of the first class meeting, using a paper-and-pencil questionnaire. The same measure was also administered on the last day of the semester (14 weeks later). To guarantee anonymity, students selected a pseudonym that they used on both the pretest and posttest so that their responses could be matched.

For comparison purposes, we also collected information from other SJFC students. Unlike the previous study, these data were collected using an online survey. All undergraduate students ($n = 2,703$) were sent a link to the online survey and brief explanation of the study via e-mail the week following finals but prior to commencement. They were given 10 days to complete the online survey. A total of 326 participants responded by at least starting the survey.

Participants in these comparison samples first saw a page with consent information, and were required to consent prior to participating. Four did not consent, and their data were removed. Respondents were asked to indicate their status (i.e., undergraduate or graduate students). Those who declined to indicate status ($n = 19$) or indicated they were graduate students ($n = 25$) were immediately routed to the debriefing by the system. To exclude students who were in the classroom sample, respondents were also asked (a) whether they were Science Scholars and (b) what year they began at St. John Fisher College. This resulted in the elimination of 12 more responses. Finally, 38 participants did not fully complete the Paranormal Beliefs Scale, for a total usable sample size of 228 participants.

All phases of this study were reviewed and approved by the St. John Fisher College Institutional Review Board. Each participant in our classes consented to allowing us to use his or her data in the study (those in the online survey who did not were immediately routed to the debriefing). All data were collected anonymously, and participants were debriefed either face-to-face or with a written online paragraph, depending on the type of administration.

D. Analysis Plan.

The main analyses for the students in our classes utilized paired-samples *t*-tests to compare their pretest scores to those from their posttest. We also conducted one-sample *t*-tests comparing the scores from the students in our classes to the averages provided by Tobacyk and Milford (1983). Finally, we compared posttest scores of the students in our classes to other students' scores using a one-way ANOVA. All analyses were conducted using the entire scale average as well as with all seven subscales.

III. Results.

A. Analyses of Research Questions.

The first research question asked whether students in our class had lower paranormal beliefs at the end when compared to their scores at the beginning. We tested this question using paired-samples *t*-tests for the entire Paranormal Beliefs Scale and all of the subscales. All but one post-test mean was statistically lower than the pre-test mean, and of those that were significant, all were in the medium to large range (Table 2). The only subscale that did not decrease significantly was the belief in extraordinary life forms. Thus, students had lower paranormal beliefs at the end of the course than they did at the beginning.

Table 2. Means and standard deviations for paranormal beliefs scale and each subscale from students in our class as a function of pretest-posttest administration (n = 43).

	Pretest <i>M</i>	Posttest <i>M</i>	t-test	<i>p</i>	Cohen's <i>d</i>
Entire Paranormal Beliefs Scale (25 items)	2.71 (.45)	2.34 (.54)	7.10	<.001	.74
Traditional Religious Beliefs scale	3.92 (.88)	3.73 (1.04)	3.73	<.001	.20
Psi subscale (4 items)	2.48 (.73)	1.94 (.67)	5.14	.001	.77
Witchcraft subscale (4 items)	2.41 (.74)	1.91 (.75)	5.56	<.001	.67
Superstition subscale (3 items)	2.00 (.97)	1.75 (.91)	3.04	.004	.27
Spiritualism subscale (4 items)	2.80 (.73)	2.35 (.79)	6.22	<.001	.59
Extraordinary Life Forms subscale (3 items)	1.88 (.88)	1.85 (.97)	.44	.66	.03
Precognition subscale (3 items)	3.22 (.77)	2.68 (.80)	4.69	<.001	.69

Our second research question examined whether students in our class had lower paranormal beliefs at the end of the semester than the students in the original Tobacyk and Milford (1983) study. We tested this question using independent-sample *t*-tests comparing the pre-test and post-test means and standard deviations to the means and standard deviations from Tobacyk and Milford. As Table 3 shows, all of the post-test means were significantly lower than the means from Tobacyk and Milford. Interestingly, many of the pre-test means were also significantly lower. Furthermore, as can be seen in the table, the nonsignificant finding from Research Question 1, which compared pre-test to post-test scores on the extraordinary life forms subscale, may be due to the fact that the Science Scholar students were significantly lower in pre-test means than Tobacyk and Milford's. Further, other research has reported anomalies in this subscale (Aarino & Lindeman, 2005).

Another aspect of the second research question examined whether students in our LC had lower paranormal beliefs at the end of their semester than the typical student at St. John Fisher College. We tested this by comparing post-test scores using independent samples *t*-tests. As depicted in Table 4, the results of these analyses demonstrated that the students in our learning community scored significantly lower on the entire Paranormal Beliefs scale than the typical St. John Fisher College student. Additionally, the students in our learning community scored lower on each of the paranormal beliefs subscales than the typical St. John Fisher College student, although only four of seven of these comparisons were statistically significant.

B. Supplementary Qualitative Analysis on Teacher-Course Evaluation Comments.

To answer the third research question, a supplementary qualitative analysis using the teacher-course evaluations across the two years of the study was conducted. The teacher-course evaluations ask for quantitative assessment and optional written feedback about areas such as course goals and objectives, aspects of the course students found beneficial, overall impressions,

and additional comments. Without being asked or prompted, 14 students wrote that the course improved critical thinking, eleven used the term *skeptic* or *skepticism*, eleven discussed how course content improved their thinking and questioning, eight discussed how it improved their ability to evaluate and/or analyze, and one specifically discussed how it improved scientific reasoning. For example, one student wrote that the course “emphasized the critical thinking portion.” Another wrote that the content “expanded our learning and made us skeptical.” Another comment stated that “many discussions helped me to think scientifically and made me much more articulate.” Finally, one student summed up the course experience by saying, “I learned to think more critically.” The unsolicited comments provide further evidence of meeting the goal of improving scientific reasoning and evidence-based critical thinking.

Table 3. Means and standard deviations for paranormal beliefs scale and each subscale from students in our class (n = 43) compared to means from Tobacyk and Milford (1983) as a function of pretest-posttest administration.

	Tobacyk & Milford		Cohen's d	Pre-test		Post-test	
	M (SD)	M (SD)		M (SD)	Cohen's d	M (SD)	Cohen's d
Entire Paranormal Beliefs Scale (25 items)	3.07 (.48)	2.71* (.45)	.55	2.34* (.54)		1.47	
Traditional Religious Beliefs scale	4.24 (.90)	3.92* (.88)	.36	3.73* (1.04)		.52	
Psi subscale (4 items)	3.19 (.84)	2.48* (.73)	.90	1.94* (.67)		1.65	
Witchcraft subscale (4 items)	2.77 (.85)	2.41* (.74)	.45	1.91* (.75)		1.07	
Superstition subscale (3 items)	2.08 (.82)	2.00 (.97)	.09	1.75* (.91)		.38	
Spiritualism subscale (4 items)	2.64 (.79)	2.80 (.73)	.21	2.35* (.79)		.37	
Extraordinary Life Forms subscale (3 items)	2.82 (.83)	1.88* (.88)	1.10	1.85* (.97)		1.07	
Precognition subscale (3 items)	3.52 (.84)	3.22* (.77)	.37	2.68* (.80)		1.02	

* Means differ from Tobacyk & Milford (1983) at $p < .05$.

These self-reported changes were also seen in the ways the students presented and used evidence in their writing and in classroom activities, such as the team debate. For example, during we often witnessed the students challenging each other's evidence and claims during the debate. One side would make an assertion, and students on the other side would immediately begin digging on the Internet to locate the information and explore other evidence related to it. This led to much deeper debates and discussions than one might expect if students came prepared only to work with previously prepared notes, as it allowed for spontaneous exchanges and an analysis of new information as it was presented. It should be noted that students were directed that they be “engaged” in the debate even when they were not speaking, but that the instructors did not specifically require students to conduct on-the-spot searches to challenge the opposing side; the students carried out these activities on their own, powerfully demonstrating

some of the ways that they had internalized the concept of critical thinking. Such displays occurred in all four years of the LC.

Table 4. Means and standard deviations for post-test paranormal beliefs scale and each subscale from students in our class (n = 43) compared to means from the faculty (n = 86) and other students (n = 228).

	Post-test <i>M (SD)</i>	Other SJFC Students <i>M (SD)</i>	Cohen's d
Entire Paranormal Beliefs Scale (25 items)	2.34 (.54)	2.74* (.62)	.69
Traditional Religious Beliefs scale	3.73 (1.04)	3.84 (.91)	.11
Psi subscale (4 items)	1.94 (.67)	2.62* (.88)	.87
Witchcraft subscale (4 items)	1.91 (.75)	2.38* (.87)	.58
Superstition subscale (3 items)	1.75 (.91)	1.94 (.83)	.22
Spiritualism subscale (4 items)	2.35 (.79)	2.77* (.88)	.50
Extraordinary Life Forms subscale (3 items)	1.85 (.97)	2.05 (.95)	.20
Precognition subscale (3 items)	2.68 (.80)	3.32* (.92)	.74

* Means differ from the posttest mean at $p < .05$.

IV. Discussion and Implications for Practice.

Our evaluation demonstrated that, at least as measured using the Paranormal Beliefs Scale, an interdisciplinary LC that emphasizes the scientific method through the use of writing and other active-learning techniques can decrease students' pseudoscientific thinking. Specifically, at the end of the semester students in the LC scored lower on the Paranormal Beliefs Scale than (a) they did in the beginning of the semester, (b) than other students at the college, and (c) than mean scores (used as norms) provided in the original research by Tobacyk and Milford (1983). This decrease was not the primary goal for the LC; indeed, it was largely a product of students' applying the main topics and tools of the course to the content (pseudoscience) we chose as a vehicle for exploring critical thinking. Pseudoscience was selected as the content primarily for two reasons. First, we expected it to be engaging, allowing students to consider a variety of situations and ideas in different contexts that can be entertaining. Second, the LC was designed for students from biology, chemistry, computer science, mathematics, and physics. Pseudoscience allowed the students to bring many of these disciplines into the discussion, providing an interdisciplinary approach that we value.

Our LC examined scientific reasoning and critical thinking through the use of formal writing, informal writing, debates, group projects, in-class group exercises, and problems that compared pseudoscientific concepts to scientific counterparts. Through this comparison, students were compelled to analyze and evaluate claims using a generalized scientific method.

Specifically, our course was structured around Shermer's (2002) classification of where thinking goes wrong, including problems in scientific thinking, problems in pseudoscientific thinking, logical problems in thinking, and psychological problems in thinking. Thus, the LC seems to have served the first-year Science Scholar students well in meeting their needs to understand scientific thinking and improve their writing without being indoctrinated into a specific discipline. Further, the class allowed students to extend critical thinking and scientific reasoning concepts beyond what they typically experience in the classroom and/or a laboratory course.

The students in the LC were drawn from several different majors in the sciences and mathematics. Thus, while the ideas of scientific thinking were also being taught in other courses, the only common experience to all students was this LC. Even when critical thinking processes are made explicit, they tend to be discipline-focused so that biology courses develop thinking like a biologist while physics courses teach how to think like a physicist. Discipline-focused reasoning can limit both the tools one uses in reasoning as well as the domain to which the reasoning is applied. The LC described here, as well as the scale used, lies far from any of these particular science disciplines. Thus, one can reasonably conclude that the writing intensive, interdisciplinary experience of the LC was one of the tools that furthered their ability to analyze such claims and consider evidence.

Other colleges and universities could easily modify this LC and apply it to creating an LC that serves their students because the techniques and evaluation described here are quite portable. The active learning techniques can be modified and used by instructors at most institutions regardless of academic emphasis or size. The writing projects and classroom debates could easily be incorporated into other courses, and the *Taking Sides* text is available in many areas (such as climate change) that would provide a starting point for such activities. Finally, by applying critical thinking skills and scientific reasoning to pseudoscientific topics, students were highly engaged, a necessary first step to learning.

V. Limitations and Future Research.

As with any classroom-based research, this study has limitations. For example, our students were a select group of high performers, making it quite easy for us, as novice writing instructors, to work with them on improving their writing. Second, we only evaluated two of the four years when we taught the course. However, we modified and improved it considerably after the first year so the evaluation would be of a considerably different course. Further, we kept the course structure mostly constant among the final three of our four years, so an evaluation during the fourth year would likely yield similar results to that of years two and three. Finally, we used the Paranormal Beliefs Scale (Tobacyk & Milford, 1983) as a surrogate measure of scientific reasoning. Given the focus of our course on contrasting pseudoscientific thinking with scientific reasoning, this measure is likely reasonable at tapping some of the thinking around pseudoscience.

The Paranormal Beliefs Scale by itself may be only a surrogate measure of scientific reasoning and critical thinking, but the results from this study are supported by other, anecdotal data from the course. One such experience relates to the way a group of the students managed to turn our initial psychic demonstration against us. During a later class activity exploring the statistics of ESP-type card guessing, several students in the class achieved a perfect 10/10 ratio guessing the cards with four different card viewers. Random guessing would only explain this event as a once-in-the-history of the universe likelihood. After a discussion of this, the students

admitted to having created a method of “tells” so that a confederate would easily know what kind of card was being viewed simply from the way the viewer held the card. Such actions clearly demonstrate students moving toward deeper critical thinking in the sense of Bloom’s taxonomy (Anderson & Krathwohl, 2001), moving well beyond the lower levels of reasoning and far into the application and creation levels of reasoning.

With respect to future research, critical thinking can be evaluated either by using the literature to identify or develop a content-specific critical thinking scale (Renaud & Murray, 2008) or instead by using a general critical thinking scale (e.g., Cheung, Rudowicz, Kwan, & Yue, 2002). Specific critical thinking scales, such as those to better understand scientific reasoning about biology (McMurray, Beisenherz, & Thompson, 1991), critical thinking about diversity (Pascarella, Palmer, Moye, & Pierson, 2001), or for evaluating paranormal beliefs (Tobacyk & Milford, 1983) can measure content-specific critical thinking. On the other hand, any instructor who modifies our course content could also consider general methods of evaluating critical thinking, including the widely used Watson-Glaser Critical Thinking Appraisal short form (Watson & Glaser, 2008), which has at least some use in assessing general critical thinking (e.g., Loo & Thorpe, 1999).

Another future question that remains is whether a course such as this would work beyond our self-selected scholarship students or with non-science students. As reported, we had a set of scholarship students who were, on average, highly motivated, more prepared for college than the typical Fisher student, and focused on learning science in their first year. Programs like this tend to have a critical thinking emphasis (Inkleas & Weisman, 2003). Regardless of our sample, the active methods used in the course should help to motivate many students (Yoder & Hochevar, 2005). Further, our interdisciplinary focus improved general critical thinking and scientific reasoning within and beyond the classroom and could apply to students who are not science majors. Thus, it is likely that this will work beyond the sample, and future courses should test and evaluate its reliability as an instructional approach. Learning communities are used, in part, to improve retention rates and student satisfaction. While these are admirable goals, LCs can also be used to improve general scientific reasoning and critical thinking (Browne & Minnick, 2005) and also, as this research demonstrates, can improve science-specific reasoning.

Acknowledgments

We thank our students for participating in this evaluation as well as Steve Brauer and Theresa Nicolay for reading and commenting on earlier drafts of the manuscript.

References

- Aarnio, K., & Lindeman, M. (2005). Paranormal beliefs, education, and thinking styles. *Personality and Individual Differences, 39*, 1227-1236.
- Aaron, J. E. (2003). *The Little, Brown essential handbook for writers*. New York: Addison-Wesley Educational Publishers.
- Adey, P., & Shayer, M. (1993). An exploration of long-term far-transfer effects following an extended intervention program in the high school science curriculum. *Cognition and Instruction, 11*(1), 1-29.

Anderson, L., & Krathwohl, D. (Eds) (2001) *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Browne, M. N., & Minnick, K. J. (2005). The unnecessary tension between learning communities and intellectual growth. *College Student Journal*, 39, 775-783.

Cheung, C., Rudowicz, E., Kwan, A. S. F., & Yue, X. D. (2002). Assessing university students' general and specific critical thinking. *College Student Journal*, 36, 504-525.

Committee on Undergraduate Science Education, National Research Council. (1999). *Transforming undergraduate education in science, mathematics, engineering, and technology*. Washington DC: National Academy Press.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton-Mifflin Company.

Dabney, D., Green, L., & Topalli, V. (2006) Freshman learning communities in criminology and criminal justice: An effective tool for enhancing student recruitment and learning outcomes. *Journal of Criminal Justice Education*, 17(1), 44-68.

Davies, M. (2013). Critical thinking and the disciplines reconsidered. *Higher Education Research & Development*, 32(4), 529-544.

Dodge, L., & Kendall, M. (2004). Learning communities. *College Teaching*, 52(4), 150-5.

Druger, M. (2002). More than science. *Journal of College Science Teaching*, 31(5), 340 – 1.

Easton, T. A. (2002). *Taking sides: Clashing and controversial views in science, technology, and society*, 5th ed. Guilford, Ct.: Dushkin/McGraw-Hill.

Easton, T. A. (2004). *Taking sides: Clashing and controversial views in science, technology, and society*, 6th ed. Guilford, Ct.: Dushkin/McGraw-Hill.

Farha, B., & Steward, G. (2006). Paranormal beliefs: An analysis of college students. *The Skeptical Inquirer*, 30(1), 37-40.

Halpern, D.F. (1997). *Critical thinking across the curriculum: A brief edition of thought and knowledge*. Mahwah, NJ: Lawrence Erlbaum Associates.

Halpern, D.F., (1998) Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53, 449-455.

Hergovich, A., & Arendasy, M. (2005). Critical thinking ability and belief in the paranormal. *Personality and Individual Differences*, 38, 1805-1812.

Franz, T. and Green, K.

Kokkala, I., & Gessell, D. A. (2002). Writing science effectively: Placing biology and english students in an author-editor relationship. *Journal of College Science Teaching*, 32, 252-57.

Inkelas, K.K., & Weisman, J.L. (2003). Different by design: An examination of student outcomes among participants in three types of living-learning programs. *Journal of College Student Development*, 44, 335-368.

Inkelas, K.K., Soldner, M., Longerbeam, S.D., & Leonard, J.B. (2008). Differences in student outcomes by types of living-learning programs: The development of an empirical typology. *Research in Higher Education*, 49, 495-512.

Inkelas, K.K., Daver, Z.E., Vogt, K.E., & Leonard, J.B. (2007). Living-learning programs and first-generation college students' academic and social transition to college. *Research in Higher Education*, 48, 403-434.

Learning Communities National Resource Center. (n.d.). Washington Center for Improving the Quality of Undergraduate Education at Evergreen State College Retrieved on January 30, 2013 from <http://www3.evergreen.edu/wacenter/pResources.asp?pid=73>

Lester, J. D., & Lester, J. D., Jr. (2003). *Principles of writing research papers*. New York: Pearson Education.

Loo, R., & Thorpe, K. (1999). A psychometric investigation of scores on the Watson-Glaser Critical Thinking Appraisal new form S. *Educational and Psychological Measurement*, 59, 995-1003.

McMurray, M. A., Beisenherz, P., & Thompson, B. (1991). Reliability and concurrent validity of a measure of critical thinking skills in biology. *Journal of Research in Science Teaching*, 28, 183-191.

Messer, W.S., & Griggs, R.A. (1989). Student belief and involvement in the paranormal and performance in introductory psychology. *Teaching of Psychology*, 16, 187-191.

Moore, T. (2013). Critical thinking: Seven definitions in search of a concept. *Studies in higher education*, 38(4), 506-522.

Morgan, P., Carter, J., Lemons, J., Grumbling, O., & Saboski, E. (1995). Preliminary assessment of a science learning community for first-year students. *Journal of College Science Teaching*, 25(2), 102-109.

National Research Council. (1998). *National science education standards*, fifth printing. Washington, D.C.: National Academy Press.

Nelson, C. E. (1999). On the persistence of unicorns: The trade-off between content and critical thinking revisited. In R. Aminzade and B. Pescosolido (eds.), *The social worlds of higher education: A handbook for teaching in a new century* (168-184). Thousand Oaks, CA: Pine Forge Press.

Pascarella, E. T., Palmer, B., Moye, M., & Pierson, C. T. (2001). Do diversity experiences influence the development of critical thinking? *Journal of College Student Development*, 42, 257-271.

Quitadamo, I, Brahler, C.J., & Crouch, G.J. (2009). Peer led team learning: A prospective method for increasing critical thinking in undergraduate science courses. *Science Educator*, 18, 29-39.

Renaud, R. D., & Murray, H.G. (2008). A comparison of a subject-specific and a general measure of critical thinking. *Thinking Skills and Creativity*, 3(2), 85-93.

Roe, C.A., (1999). Critical thinking and belief in the paranormal: A re-evaluation. *British Journal of Psychology*, 90, 85-98.

Royalty, J., (1995). The generalizability of critical thinking: Paranormal beliefs versus statistical reasoning. *The Journal of Genetic Psychology*, 156, 477-488.

Schön, D. (1987). *Educating the reflective practitioner*. San Francisco: Jossey-Bass.

Shermer, M. (2002). *Why people believe weird things: Pseudo-science, superstition, and bogus notions of our time*. New York, NY: Henry Holt and Company, LLC.

Smith, B.L., MacGregor, J., Matthews, R.S., & Gabelnick, F. (2004). *Learning communities: Reforming undergraduate education*. San Francisco John Wiley & Sons, Inc.

Tobacyk, J., & Milford, G. (1983). Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology*, 44, 1029-1037.

Watson, G.B., & Glaser, E. M. (2008). *Watson-Glaser critical thinking appraisal short form manual*. San Antonio, TX: Pearson Education.

Yoder, J. D., & Hochevar, C.M. (2005). Encouraging active learning can improve students' performance on examinations. *Teaching of Psychology*, 32, 91-95.