International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Student Moon Observations and Spatial-Scientific Reasoning

Merryn Cole, Jennifer Wilhelm & Hongwei Yang

To cite this article: Merryn Cole, Jennifer Wilhelm & Hongwei Yang (2015) Student Moon Observations and Spatial-Scientific Reasoning, International Journal of Science Education, 37:11, 1815-1833, DOI: 10.1080/09500693.2015.1052861

To link to this article: http://dx.doi.org/10.1080/09500693.2015.1052861

Published online: 11 Jun 2015.

International Journal of Science Education, 2015
Vol. 37, No. 11, 1815–1833, http://dx.doi.org/10.1080/09500693.2015.1052861

Student Moon Observations and Spatial-Scientific Reasoning

Merryn Cole[a*], Jennifer Wilhelm[a] and Hongwei Yang[b]
[a]*Department of Science, Technology, Engineering, and Mathematics Education, University of Kentucky, Lexington, KY, USA;* [b]*Department of Educational Policy Studies and Education, University of Kentucky, Lexington, KY, USA*

Relationships between sixth grade students' moon journaling and students' spatial-scientific reasoning after implementation of an Earth/Space unit were examined. Teachers used the project-based Realistic Explorations in Astronomical Learning curriculum. We used a regression model to analyze the relationship between the students' Lunar Phases Concept Inventory (LPCI) post-test score variables and several predictors, including moon journal score, number of moon journal entries, student gender, teacher experience, and pre-test score. The model shows that students who performed better on moon journals, both in terms of overall score and number of entries, tended to score higher on the LPCI. For every 1 point increase in the overall moon journal score, participants scored 0.18 points (out of 20) or nearly 1% point higher on the LPCI post-test when holding constant the effects of the other two predictors. Similarly, students who increased their scores by 1 point in the overall moon journal score scored approximately 1% higher in the Periodic Patterns (PP) and Geometric Spatial Visualization (GSV) domains of the LPCI. Also, student gender and teacher experience were shown to be significant predictors of post-GSV scores on the LPCI in addition to the pre-test scores, overall moon journal score, and number of entries that were also significant predictors on the LPCI overall score and the PP domain. This study is unique in the purposeful link created between student moon observations and spatial skills. The use of moon journals distinguishes this study further by fostering scientific observation along with skills from across science, technology, engineering, and mathematics disciplines.

Keywords: *Geometry and geometrical and spatial thinking; Middle school education; Lunar phases; Spatial ability; Geometric spatial visualization; Periodic patterns; Astronomy*

Introduction

At some point in our lives, most of us have wriggled our toes through sand. It may have been a warm, sandy beach, a child's sandbox, or another adventure. Most likely, we

*Corresponding author. Department of Science, Technology, Engineering, and Mathematics Education, University of Kentucky, Lexington, KY 40506, USA. Email: merryn.cole@uky.edu

were focused on the other things around us. Some would think nothing of flicking away the sand, but contained in those grains of sand is an opportunity for wonder and investigation. Where did those grains of sand come from? What are they made of? What do they look like? We had some awareness of the sand beneath our feet, but did not pay much attention to it. How often do we experience nature without really seeing it? Unlocking the secrets of nature, much like with the sand, starts with careful observation of the phenomenon. Daily Moon observation journals are a way students can begin to unlock the mystery of the phases and motions of the Moon while practicing making observations. Students can observe where the Moon is in the sky, how it moves, and what it looks like from the comfort of their own yard.

This study examines the link between students' performance on moon journals and the Lunar Phases Concept Inventory (LPCI; Lindell & Olsen, 2002). We suggest students who put more effort into their moon journals (i.e. complete more entries and score better on moon journals overall) during their Earth/Space unit will perform better on the LPCI, both overall as well as on relevant spatial domains, such as the periodic patterns (PP) and geometric spatial visualization (GSV) domains. There is no literature, to our knowledge, comparing these results; we are basing our hypothesis on the general trends of the moon journal and LPCI scores from three schools (Table 1). Our study draws on the science notebook, writing to learn science, and spatial ability literature to frame our work. Students in this study used moon journals, a special type of science notebook focused on a specific phenomenon, to make observations and write about their emerging ideas and questions on the topic. Along the way, they also developed spatial skills to help them make sense of the mental models needed in order to understand the lunar phases.

Background

Understanding many aspects of astronomy relies on an individual's ability to use the positions and motions of celestial objects to explain observed phenomena and to make predictions based on these observations (Plummer, Wasko, & Slagle, 2011). These astronomical explanations also require spatial ability, which has been linked to performance in both mathematics and science (Black, 2005; Lord & Rupert, 1995; Wilhelm, 2009; Wilhelm, Jackson, Sullivan, & Wilhelm, 2013). The LPCI was developed to assess students' mental models of lunar phases (Lindell & Olsen, 2002). These mental models are an important part of making sense of the world around us. Students create mental models to explain phenomena and situations they encounter. A

Table 1. Average moon journal and LPCI scores by school

	Moon journal	LPCI pre (%)	LPCI post (%)	LPCI gain (%)
Juniper	18.31	25.83	34.44	8.43
Butternut	24.47	23.95	47.11	22.84
Willow	22.15	41.92	47.31	5.38

mental model is an individual's unique, internal representation of a fact, a thing, or a phenomenon (Pirnay-Dummer, Ifenthaler, & Seel, 2012). Students use these mental models to interact with and make meaning of the world around them. New experiences are viewed through the lens of the students' existing mental model (Jones, Ross, Lynam, Perez, & Leitch, 2011). The students in this study used their moon observation journals along with their mental models as tools to make sense of the patterns in the appearance and location of the Moon over time, later incorporating a physical three-dimensional model to further refine their understanding of the lunar phases.

Science Notebooks

From the middle-ages into the eighteenth century, alchemists diligently worked in their labs hoping to discover the philosopher's stone. Many of their findings were written as analogies or illustrated in paintings (Crosland, 2004). Only other alchemists would be able to interpret and understand the codes left in these paintings and writings. Even then, it was easy to fail to see the meaning in an alchemical text or to read meaning into an analogy where no alchemical meaning was intended (Crosland, 2004). As alchemy evolved into science, scientists still recorded their findings and musings, but wrote them in a way others could understand and repeat their experiments. Since then, lab notebooks have evolved from places to log data into science notebooks used by students and scientists alike to record and interpret not only lab data, but also thoughts and ideas for understanding science.

Throughout history, scientists have used science notebooks to record their thoughts and experiments. These notebooks describe much about how the author is thinking about science (Tweney, 1991). Some scientists, such as Faraday, were mistrustful of their own memory; so they took extensive notes (Tweney, 1991). Other scientists perhaps were either more trusting of their memory, or simply took less detailed notes for another reason. Regardless of the extent of the notes or the format, keeping a science notebook is something modern scientists have in common with those throughout history. As research notebooks of past scientists have been unearthed and interpreted, scholars have learned more about how the science was accomplished (Holmes, Renn, & Rheinberger, 2003). This knowledge goes beyond the polished pieces that end up in published works, and sheds light onto the thoughts and scientific processes used to actually do science.

Science notebooks have become tools for novices developing an understanding of science in the classroom as well as for scientists conducting research. Similar to scientists, students use scientific notebooks in the classroom to record problems they tried to solve, procedures for laboratory activities, observations made, questions that arose, conclusions, or reflections on what they have done or learned. They are a snapshot of the instructional activities in a classroom, including both what teachers focus on and what students do (Baxter, Bass, & Glaser, 2000). Many of these activities mirror the way scientists use their own notebooks for research. Science notebooks can be used for a variety of scientific endeavors, whether these include a broad number of topics or are focused on a specific topic of study. If the focus is on a particular phenomenon, such as the appearance and location of the Moon over time as in this study, we may give the science notebook a more specific name; in this case, we used the term moon journal for these focused science notebooks.

Moon Journals: Science notebooks connecting students to nature

Science involves investigation of both new and known phenomena, often without a known outcome. In order to understand science, children need to be given an opportunity to explore the world around them through hands-on, open-ended activities. When students act like scientists, they 'learn about their works by observing, describing, questioning, and searching for answers' (Doris, 1991, p. 5). Moon journals give children a place to write their observations as well as their emerging thoughts what they are observing. They also provide a place to write to develop students' early understandings and to ask questions for further investigation. In this way, moon journals are a specific form of a science notebook. Like science notebooks, moon journals provide a place for students to record observations, questions, and ideas while studying science. But, they are a focused science notebook where students are studying a particular phenomenon: the appearance and location of the Moon over time.

Chancer and Rester-Zodrow (1997) asked their students to observe the Moon for four weeks in a row. In doing so, they were asking their students to become 'active learners, questioning scientists … and detailed writers' (p. xvii). They showed their students nature notebooks and science notebooks of scientists past, and encouraged their students to act as scientists and observe and take in the world around them. By acting like scientists, students made and recorded their observations and wrote about what they observed in order to make sense of it. Students are drawn into the project through their freedom to express themselves using their own art and writing in their entries. As the journals progress over time, students are able to connect the content of their journals to their lives as well as pose questions they could investigate (Chancer & Rester-Zodrow, 1997).

As Duckworth (1996) suggested, 'thoughts are our way of connecting things up for ourselves … we can only understand them to the extent that we do the work of making these connections ourselves' (p. 26). Writing in a moon journal is one way to record these thoughts and begin making the connections necessary for understanding the motions and changes in appearance of the Moon. The moon journal allows students to record their observations, thoughts, and emerging ideas about what is happening. Newman, Morrison, and Torzs (1993) used the expression 'scientific sense-making' to describe the task of explaining observed phenomena based on data and theory. In order to make sense of science, students not only need the opportunity to make the connections as Duckworth suggested, but also have the time to put their thoughts into words; these two purposeful activities together lead to scientific sense-making for students.

Writing in Science, Writing to Learn

Baxter, Bass, and Glaser (2000) asserted that conducting scientific inquiry is key to developing students' understanding of science. Baxter et al. (2000) go on to say that the most important part is the student's written record, where they can 'formulate questions, shape investigations, monitor progress toward goals, and generate conclusions' (p. 2). Like scientists, students are also dependent on taking notes to guide their learning and should be encouraged to write to learn science just as scientists throughout history have written about their thoughts, processes, and results.

Newell (1983) posited that writing can be a powerful tool for learning, when the writer integrates new information with existing understandings. Howard (1988) proposed that the 'act of writing is father to thought itself', suggesting that writing is more than a communication tool; it is also a tool for developing understanding (p. 88).

Most classroom writing is informational and used as a tool for the teacher to gauge understanding or record notes about what students need to know (Needham & Hill, 1987). Rarely students encounter meaningful, authentic learning tasks where they are asked to write for a specific purpose and audience (Anson, 1988; Applebee, 1981, 1984; Tighe, 1991). When using science notebooks in the classroom, time must be given for both recording data, a traditional use of science notebooks, and also for writing about what was learned (Fulwiler, 2007). Both the science and the writing skills will suffer if sufficient time is not provided for students to synthesize what they have learned and to write about it. Science notebooks can be used as tools to allow students to create understanding through writing, and also as a place for authentic writing tasks where students can act as scientists (Needham & Hill, 1987). Moon journals provide authentic opportunities for students to make observations of the natural world, as well as to reflect on the patterns they observe, their learning, and the questions that still remain (Comstock, 1939; McMillan & Wilhelm, 2007). When students compare their journal observations with others, new understandings can be formed. For example, Wilhelm, Sherrod, and Walters (2008) reported how preservice teachers in the USA corresponded via moon journaling with Australian preservice teachers; this correspondence resulted in reciprocal understandings of Moon patterns visualized from both a northern hemisphere and a southern hemisphere perspective.

Spatial Ability

The act of keeping a moon journal is an interdisciplinary experience; it includes science observations as well as skills used in many other professions. Students write about their observations, practice drawing what they see rather than what they think they see, make measurements between objects in the sky, and reason with the evidence they collect to create explanations of phenomena. Student moon observation journals foster a purposeful link between what students can observe about the world around them and the spatial relationships inherent in these observations. Lord and Rupert (1995) have found that visuo-spatial aptitude is essential for a variety of professions, such as artists, architects, engineers, and scientists. Research has shown that spatial ability can be improved through instruction and is especially effective when the instruction is targeted and discipline specific (Alias, Black, & Gray, 2002; Uttal, Meadow, Tipton, Hand, Alden, Warren, & Newcombe, 2013).

Males generally outperform females on assessments of spatial ability (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). Research has shown that the difference is more pronounced for certain types of spatial ability (e.g. mental rotation, spatial perception) than for spatial visualization (Voyer et al., 1995). However, Voyer et al. (1995) noted a lack of consensus in the literature as to how to categorize spatial abilities; this lead to some difficulty in categorizing what assessments in the papers used in their meta-analysis were measuring. The LPCI questions can be categorized into four spatial mathematics domains: GSV (visualizing the spatial features of a system from above, below, or within the plane of the system), PP (identifying repeated incidences at regular intervals), spatial projection (SP; mentally projecting to another location on an object and visualizing from that perspective), and cardinal direction (CD; distinguishing directions to identify an object's location in space) (Wilhelm, 2009). In this paper, we chose to focus on the GSV and PP domains because they are the most relevant to the keeping of moon journals and toward making meaning from observations by visualizing the Earth/Moon/Sun geometry and noticing lunar patterns (e.g. phase change, rise time, set time). As Black (2005) noted, we need to create our own mental models of the Earth/Moon/Sun system because we are restricted to a 'single vantage point from Earth of the moving bodies in outer space' (p. 403). Creating this mental model of the system requires GSV—visualizing the spatial features of the system as if we were above, below, or within the plane of the system. The PP domain (identifying repeated incidences at regular intervals) was chosen as students were asked to write about patterns in the location and appearance of the Moon in their moon journals.

This study examines the link between student performance on an assessment of knowledge of lunar phases (LPCI) and the keeping of daily Moon observation journals. We explore the association of the moon journals with both the overall performance on the LPCI as well as two spatial domains, GSV and PP, of the LPCI. Specifically, we address the following research questions: (1) In what ways do daily Moon observation journals contribute to students' understanding of Moon phases? (2) How do Moon observation journals relate to students' performance on GSV and PP spatial domains?

Methods

Participants

Research subjects were sixth-grade students from three south-central US middle schools (Juniper, Butternut, and Willow Middle Schools) in urban settings. This data set includes students from two teachers at Juniper Middle School, three from Butternut Middle School, and one from Willow Middle School. The demographic make-up of the schools can be found in Table 2; demographic make-up of each of the classrooms was similar to that of the rest of the school. The total *N* for the study is 333, with 127 students from Juniper, 156 students from Butternut, and 50 from Willow. Of these students, the gender breakdown is 72 male students and 55 female students at Juniper; 65 male students and 91 female students at Butternut; and 30 male students and 20 female students at Willow. The teachers' experience ranged from 0 to 15 years in the classroom, with an average experience of 5.9 years. Pseudonyms were used.

Each of the teachers followed a project-based curriculum called Realistic Explorations in Astronomical Learning (REAL) focusing on Earth/Space lessons. Juniper teachers spent 9 weeks on this unit, while Butternut and Willow teachers implemented their units in approximately 4 ½ weeks. Due to time constraints, teachers at Butternut and Willow Middle Schools implemented only the lunar phases lessons of the REAL curriculum. Juniper teachers implemented these same lessons as well as additional REAL lessons on other astronomical concepts and student projects. As part of the curriculum, students were asked to keep daily moon observation journals (Figure 1). In these journals, students sketched the appearance and recorded the location (azimuth and altitude) of the moon daily. They were also asked to write about what they observed, specifically noting any emerging patterns and to make predictions about the future appearance and location of the moon. Moon journals were kept for an entire lunar cycle.

Measures

This study includes three sources of data: the LPCI, student daily Moon observation journals, and a teacher survey. The LPCI was used to assess content knowledge of the students. The moon journals were scored using a rubric. The teacher survey was intended to determine how teachers implemented and scored the moon journals in their classrooms.

Lunar phases concept inventory. The quantitative test data source was the LPCI, a 20 question multiple-choice test that assessed eight science domains as well as four spatial-mathematics domains. The assessment was given to all students before and immediately after their Earth/Space unit. Hierarchical Linear Modeling (HLM) was used to assess which factors would best predict student post-test scores on the LPCI as well as post-test scores on the PP and GSV domains on the LPCI.

Moon journals. Moon journals were scored using a five-point scale in each of six categories: number of entries, entries with sketches, entries with sentences, entries with

Table 2. Demographic make-up of participating schools

	Juniper (%)	Butternut (%)	Willow (%)
White	84	74	45
Black	7	10	30
Hispanic	3	8	15
Asian	3	5	7
Other	3	3	3
Free/reduced lunch	30	26	56

Figure 1. Sample moon journal entry. ‘I think that the moon is waxing because the moon’s lit side is enlarging’

azimuth and altitude angles, accuracy of the terminator line (the boundary between the lit and unlit portions of the Moon), and sentence quality. The sentence quality score was based on the number of quality sentences present; each student was asked to write at least two sentences per observation. A student received full credit in this category if two quality sentences were present for each entry. Partial credit was also given to entries with one quality sentence. Quality sentences were those that were scientifically accurate and relevant to the Moon or sky. For instance, a sentence that described recently observed changes in position of the Moon from day to day would be a quality sentence, while a statement of ‘the Moon rose in the western sky’ (scientifically inaccurate) or ‘there was a dog barking while I was outside’ (irrelevant) would not be counted as quality sentences. To avoid penalizing multiple times, the scores for the last five categories were determined as a percentage of the number of entries present. For instance, the score for accuracy of the terminator line was based on a percentage correct out of the number of sketches present, not on the number of entries. Category scores were then summed to give the final score. To establish interrater reliability, both scorers initially coded the same 16 moon journals separately. The scorers agreed on 80.2% of the items. Upon discussion of the scores, the interrater reliability rose to 97.7%. The remaining moon journals were then scored using the same rubric.

Teacher survey. Teachers were asked to complete a questionnaire about how moon journals were used in their classrooms. The survey was conducted via email and included six questions. The teachers were asked to describe how they implemented and scored the moon journals and to describe their students’ level of interest in the moon journals. Questions also addressed the frequency of student-led and teacher-led discussions about the moon journals.

Statistical Analysis

The LPCI post-test data were analyzed using SAS 9.3 (SAS, Inc., 2007). The data consist of a total of 333 students with each one measured on 10 analysis variables. Out of the 10 analysis variables from the students, 3 are dependent variables (DV): (1) LPCI post-test total (PostTotal) score, (2) LPCI post-test periodic pattern (PostPP) domain score, and (3) LPCI post-test geometric spatial visualization (PostGSV) domain score. The other seven variables were used as predictors. A description of each variable along with descriptive statistics for each is provided in Table 3.

We built a regression model that analyzed the relationship between each of the students' LPCI post-test score variables (PostTotal, PostPP, and PostGSV) and several predictors outlined in Table 3. To analyze the LPCI post-test data, we chose a linear mixed model approach over a linear model approach due to the hierarchical structure of the data: students were nested within instructors. The HLM model search process failed to support the use of HLM for any of the three DV's. Given that the extensive search within the multilevel model space fails to justify the addition of a higher-level structure for any of the three DV's, we moved to the single-level model framework.

Results

LPCI and Moon Journals by School

Table 1 shows the average moon journal and LPCI scores by school. Juniper and Butternut students scored similarly on the LPCI pre-test, but showed differences in the

Table 3. Descriptive statistics of all variables

Variable	*N*	Mean	Std. dev.	Min.	Max.	Note
PostTotal	333	8.535	3.412	1	19	DV: Post-test LPCI Total Score
PostPP	333	0.511	0.285	0	1	DV: LPCI PP Domain Score
PostGSV	333	0.505	0.250	0	1	DV: LPCI GSV Domain Score
PreTotal	333	−0.324	2.516	−5	8	Pre-test LPCI score overall: Centered by 6
PrePP	333	0.049	0.258	−0.3	0.7	Pre-test PP Domain Score: Centered by 0.3
PreGSV	333	−0.002	0.183	−0.3	0.557	Pre-GSV: Centered by 0.3
JournalScore	333	−0.006	4.602	−12	9	Overall score on Moon Journal; Centered by 21
Entries	333	0.267	11.693	−19	28	Number of Moon Journal entries; Centered by 21
Gender	333	0.502	0.501	0	1	0 = girls, 1 = boys
Experience	333	0.012	5.590	−7	8	Teaching Experience, Centered by 7

Note: DV, dependent (outcome) variable; other listed variables are independent (predictor) variables.

pre- to post-test gains (8.43% vs. 22.84%, respectively). They also showed differences in their moon journal scores. Willow students scored considerably higher on the LPCI pre-test than students in the other two schools and showed a lower pre- to post-test gain (5.38%) than the other schools. These differences in LPCI scores were observed prior to the analysis of the moon journals, prompting an exploration of the moon journal scores as a possible contributing factor for the differences. The Willow students' moon journal scores fell in between the scores from Butternut and Juniper.

Moon Journals

Evaluation of the student moon journals shows a wide range in the content and quality of student entries. Of particular interest is what the students wrote in addition to their sketches. Some students identified patterns in illumination of the Moon (Figure 2(a) and 2(b)) while others noted the change in position in the sky (Figure 3). These types of entries emphasize the spatial patterns students begin to make sense of through the moon journals. Along with these observations, students also attempted to make sense of the technical terms that go along with describing the phases, such as when the moon is waxing vs. waning. They also often discussed the location of the Moon in relation to other objects in the sky, such as planets and constellations. It is clear that some students were surprised by their observations, and commented on observations they thought did not match up to their expectations. Others seem to enjoy seeing their predictions coming true. They note changes in azimuth angle (cardinal direction) and altitude angle (height in the sky) from day to day. Regardless of their focus, writing about their observations allowed students to consider patterns of lunar movement both throughout the day and from day to day, illumination of the Moon, and location of the Moon.

The moon journal scores used in the analysis were limited to the number of entries and the overall score. The moon journals included from 2 to 49 entries, with an average of 21 entries per journal. This was compared to the expected 30 entries from each student, which would cover an entire lunar cycle. The overall score ranged from a low of 9 to a high of 30; 30 was the highest possible score that could be obtained using the rubric. These two scores, the raw number of entries and the overall total moon journal score, were used as predictor variables because the authors felt they represent the amount of effort put in and the quality of the moon journals overall. Most of the entries (96%) included a sketch of the Moon. Of the entries with sketches, 69% showed a correct terminator line. Two-thirds (66%) of the entries included sentences, 85% of which were considered quality sentences. Quality sentences were those that described the Moon or sky as opposed to unrelated or scientifically inaccurate explanations.

Teacher Questionnaire

The emphasis teachers placed on the moon journals varied by classroom. While in general the teachers felt the moon journals are an important part of the curriculum,

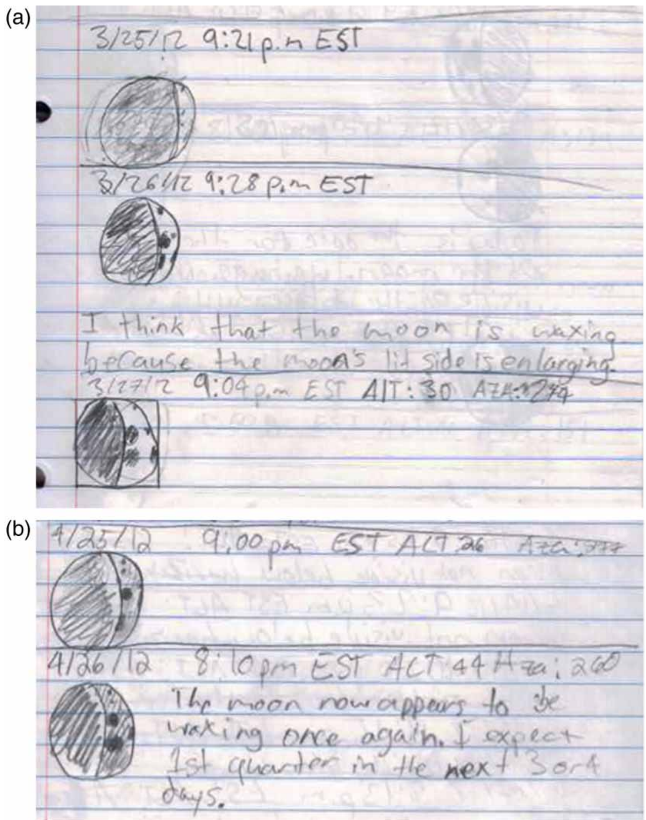

Figure 2. (a) Student entry noting increased illumination. 'The moon now appears to be waxing once again. I expect 1st quarter in the next 3 or 4 days.' (b) Student entry noting increased illumination and making a prediction. 'The moon is getting higher and moving to the East. It is also getting fuller'

they also noted that student interest in keeping the journals declines over time. Some also suggested that it is better for the advanced students to keep moon journals than the average students. There was also a range in how the moon journals were graded in the classrooms. Some teachers gave participation credit for having a moon journal, others graded each required piece (drawing, sentences, location, etc) individually, and another teacher did not grade the moon journal directly at all, but allowed her students to use it on a graded moon quiz. The teachers at Butternut Middle School each took time in their classrooms at least a couple of times each week to have students share their entries with the class and discuss what they had been observing. They also discussed as a class why the moon was moving and how its location was changing over time. The teachers at Juniper Middle School also discussed the observations as a class, but their students completed entries less often, resulting in less frequent discussion of how the Moon was changing. All of the teachers asked students to put their moon journals together in class, usually with time provided for students to decorate their covers. The teachers at Juniper Middle School started the

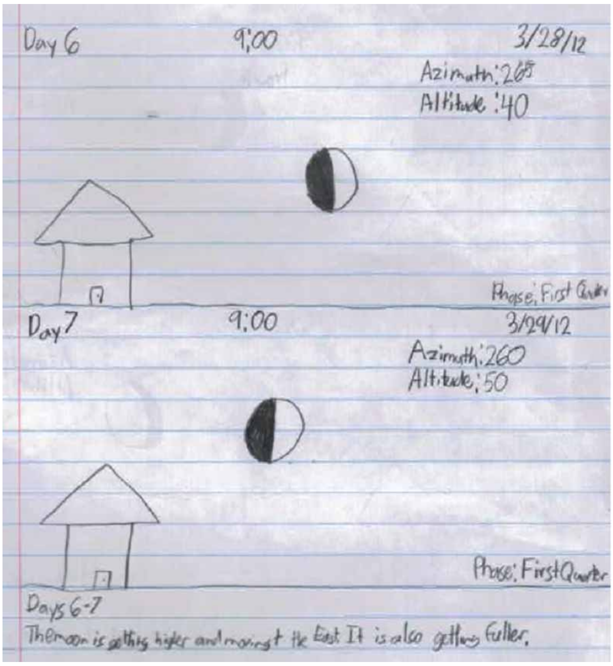

Figure 3. Student entry noting daily change in moon location

journal with some notes about the Moon, while the teachers at Butternut Middle School started with a description of what to observe and record.

Linear Regression Model

All three models share three predictors: pre-test score (corresponding to the DV), number of moon journal entries, and moon journal score. The models for PostTotal and PostPP include only these three predictors. In contrast, the model for PostGSV also includes student gender and teacher experience, indicating that these are also significantly related to student post-test GSV domain scores. Each intercept estimate is interpreted as the average score on the DV (dependent variable) when all selected predictors in the model equal zero. For PostTotal, the estimate for the intercept can be interpreted this way: the average LPCI PostTotal score for all those students who are at the average level on all three predictors (21 entries, 21 score on moon journals, and a score of 6 on PreTotal) is estimated to be 8.68225 out of a possible 20 points. For all students scoring at the average level on the other predictors, the estimated PostPP score would be 48.74%. When students are at the average level on the predictors for PostGSV and are Female, the PostGSV score is estimated to be 47.52%.

Tables 4–6 show the unstandardized estimates for β values, standardized estimates of variables, intercept, p-values, and R^2 values for the final models for the overall LPCI, PP domain, and GSV domain outcome variables. The unstandardized estimates measure the relationship between the dependent variable and each predictor in the model,

Table 4. Final model results for predicting LPCI total scores

	Unstandardized estimate	Standard error	Standardized estimate	*t*-Value	Pr > |*t*|
Intercept	8.68225	0.15412	0	56.33	< .01
Entry	0.05927	0.01717	0.20314	3.45	< .01
Journal	0.18386	0.04309	0.248	4.27	< .01
PreTotal	0.5009	0.06175	0.36931	8.11	< .01

Note: $R^2 = 0.3396$; Total LPCI scores (pre and post) were reported as a raw score (0–20) while PP and GSV domain scores were reported as a percent (0.0–1.0).

Table 5. Final model results for predicting PP domain scores

	Unstandardized estimate	Standard error	Standardized estimate	*t* value	Pr > |*t*|
Intercept	0.48739	0.01279	0	38.1	< .01
Entry	0.00666	0.0014	0.27291	4.76	< .01
Journal	0.00894	0.00354	0.14416	2.52	< .05
PrePP	0.44932	0.04949	0.4057	9.08	< .01

Note: $R^2 = 0.3610$; Total LPCI scores (pre and post) were reported as a raw score (0–20) while PP and GSV domain scores were reported as a percent (0.0–1.0).

Table 6. Final model results for predicting GSV domain scores

	Unstandardized estimate	Standard error	Standardized estimate	*t* value	Pr > |*t*|
Intercept	0.4752	0.01679	0	28.3	< .01
Gender	0.05697	0.02392	0.11411	2.38	< .05
Entry	0.0058	0.00135	0.27133	4.31	< .01
Journal	0.01151	0.00336	0.21189	3.42	< .01
PreGSV	0.36291	0.06497	0.26555	5.59	< .01
Experience	−0.00672	0.00219	−0.15025	−3.07	< .01

Note: $R^2 = 0.2792$; Total LPCI scores (pre and post) were reported as a raw score (0–20) while PP and GSV domain scores were reported as a percent (0.0–1.0).

whereas standardized estimates evaluate the relative importance of each predictor in predicting the dependent variable. The unstandardized estimates of the predictor variables show that the pre-assessment scores associated with the outcome variable for that model are significant in all cases. Two additional predictors, number of moon journal entries and overall moon journal score, are shared by all three models. This indicates that either of the two is still able to contribute significantly to the prediction of each LPCI post-test score given the contributions from all other predictors in the model. The models for PostTotal and Post PP are simpler, which each model containing only three predictors. In contrast, the model for PostGSV is more complex with five predictor variables. Because gender is included in the model for PostGSV, we can see there is a statistically significant difference between male students and female students in terms of average GSV scores, when we hold constant the values of other predictors. With instructor experience in the model, we can see that instructors' teaching experience is significantly related to students' LPCI GSV post-test score, after adjusting for the other predictors in the model. Unlike the other predictors, however, an increase in teachers' experience is associated with a lower PostGSV score.

The standardized estimates measure the relative importance of each predictor to the dependent variable. Thus, their values can be directly compared across predictors. In two (total LPCI and PP Domain) of the three models, the pre-assessment score was the strongest predictor of the relevant outcome variable, while the number of moon journal entries was strongest for the GSV Domain model. The relative strength of the other variables differed for each of the final models. In the model addressing the overall LPCI score, the standardized estimate for the PreTotal is 0.3691. The next strongest predictor is the overall moon journal score (0.248) followed by the number of moon journal entries (0.20314). From these values, we can say that the PreTotal score is nearly twice as strong of a predictor than the number of moon journal entries, though both (along with the total moon journal score) are statistically significant predictors of the final total LPCI score. Similarly, the relative importance of predictors for the PP Domain score, based on standardized estimates, is PrePP (0.4057), followed by number of moon journal entries (0.27291), and finally overall moon journal score (0.14416). The model for the GSV Domain scores contains more predictors than the models for the total LPCI score or the PP Domain score, but the predictor variables can be similarly ranked. The relative importance of the predictor variables in the model of GSV Domain score are as follows: number of moon journal entries (0.27133) > PreGSV score (0.26555) > moon journal score (0.21189) > gender (0.11411) > teacher experience (−0.15025).

Discussion

The authors claimed that students who made many entries in their moon journals and scored high overall on the journal would also score well on the LPCI post-test. Students could come into the classroom already having memorized the phases of the Moon, but they would need strong spatial skills such as those developed keeping the moon journal and experiencing the REAL curriculum in order to transfer that knowledge to the rest of the LPCI. In addition, knowing the names and shapes for the phases of the Moon is not sufficient for understanding the causes of the lunar phases; doing well on the LPCI requires an understanding of the causes of the phases and the motions of the Moon.

Moon Journals and Student Understanding of Lunar Phases

In addressing research question 1, our data suggest students who put more effort into thinking and writing about what they observed in addition to recording the appearance and location of the sky likely gained the most from the moon journals. The results of the final model show that students who kept a moon journal for at least 3 weeks (21 entries) tended to score significantly better on their post-test than students who made fewer entries. Furthermore, the overall score on the moon journal was an even stronger predictor of a student's post-LPCI score. These two predictors, moon journal score and number of entries, support our assertion that students who kept moon journals tended to perform better on the LPCI than students who put minimal effort into their moon journal.

While the moon journals were significant predictors for the LPCI scores overall, they were used differently by the teachers, possibly explaining why there was such a difference in the gain in overall LPCI score exhibited by the students at Juniper and Butternut Middle Schools. The teachers at Butternut emphasized the daily observations. One of the teachers discussed the student observations as a warm-up activity for the class each day. Others discussed the observations throughout the week, though not daily. The students at Butternut made the observations outside for the most part, using Stellarium, a free planetarium software, to fill in as needed on days when they could not see the Moon. In contrast, the teachers at Juniper only discussed the moon journals as a class a few times throughout the course of the observations. Their students made fewer observations, with one teacher reporting that students regularly used Stellarium to gather their data on the Moon rather than making observations outside. Less is known about how the moon journals were used at Willow; the teacher did not respond to the questionnaire. From looking at the student moon journals, however, we can see that students made daily entries in their moon journals, though they only wrote about their observations every other day. These students completed entries for about two weeks longer than the students at the other schools, allowing them to see more of the patterns repeat. These differences in the emphasis on moon journals in the classroom, combined with the fact that moon journal scores overall and number of entries were significant predictors of LPCI post-test score, may help explain some of the differences in student understandings of lunar phases between the schools.

Moon Journals and Students' Spatial Understanding

In addressing research question 2, the final models show that students who performed better on their moon journals also tended to score significantly better on the PP and GSV domains of the LPCI post-test. For every 1 point increase in the overall moon journal score, participants scored 0.18 points (out of 20) or nearly 1% point higher on the LPCI post-test when holding constant the effects of the other two predictors, LPCI pre-test score and number of moon journal entries. Similarly, students who increased their scores by 1 point in the overall moon journal score scored approximately 1% higher in the PP and GSV domains of the LPCI. Also, student gender and teacher experience have been shown to be a significant predictor of post-GSV scores on the LPCI in addition to the pre-test scores, overall moon journal score, and number of entries that were also significant predictors on the LPCI overall score and the PP domain.

As with the overall post-test LPCI score, students who included more entries and performed better on the moon journal overall (as evidenced by the overall moon journal score) tended to perform better on their post-test PP domain score. In addition, the GSV domain post-test score tended to be higher for students who scored better on the number of entries and overall moon journal score, but also for male students. Finding that male students tended to score higher than female students on the GSV domain was not surprising. Previous studies (Linn & Petersen, 1985; Voyer et al., 1995; Wilhelm, 2009) have shown a similar result where males tend to outperform females on assessments of spatial ability. Students with more experienced teachers tended to score slightly, but significantly, lower on the GSV post-test score than students with less experienced teachers. This could be due to 'less practiced' teachers recently experiencing innovative, teacher preparation programs. These teacher preparation programs may have provided these beginning teachers with experiences in helping their students model phenomena, use science notebooks in the classroom, or analyzing and interpreting data. Newer teachers may be more familiar with the processes of scientific inquiry and the shifting focus in new science standards from pieces of content knowledge to science practices.

An examination of the quality of entries demonstrates that students who put more effort into their moon journals appeared to notice more patterns in the appearance and location of the moon in the sky. These patterns additionally relate to their development of spatial skills, as they are describing the apparently changing location of celestial objects in relation to their single position on Earth. Since students noticed and wrote about patterns of the appearance of the Moon and movement of the Moon on a daily basis and throughout the day in their moon journals, it is not surprising that students who put more effort into their moon journals also scored well on the PP domain of the post-test. The daily observations of the moon journals allowed students to experience the periodic motions of the Moon, rather than simply reading or hearing about them. They were able to use these first-hand accounts in their moon journals to make sense of the patterns. In addition, students were able to see that the Moon is not only out at night, but also during the day for some phases. This observation could cause students to start thinking about how the Earth, Moon, and Sun would need to be arranged to produce certain phases. The moon journal provides a place for students to write about these emerging ideas, asking questions and beginning to make sense of the lunar phases (Howard, 1988; Newman et al., 1993). In addition, a later lesson in the REAL Curriculum has students create three-dimensional models of the Earth, Moon, and Sun geometries to produce each phase of the Moon, referring students to their moon journals to check that the geometries make sense with their observations.

This combination of modeling and observations allows students to understand the geometry needed for each phase, but also requires GSV, or visualizing a system in space from above/below/within a system's plane. As Black (2005) observed, we must create our own mental models of the system because we are unavoidably restrained to a single vantage point in space. The three-dimensional models allow students to see the arrangement of the Earth, Moon, and Sun, while the moon journals allow the students to check whether their model agrees with what they observed from their fixed position on the Earth.

Students were asked to record the appearance and location of the moon each day in their journals, and also asked to explain any emerging patterns and make predictions for the future appearance and location of the moon. This pattern recognition and prediction-making requires students to develop and use spatial skills. Students need to be able to use spatial skills, such as GSV to create mental models that explain the natural phenomena they are directly observing. The questions on the LPCI can be categorized into four spatial-mathematical domains: PP, GSV, CDs, and SP. Students who note the change in azimuth angle of the moon in their daily entries are noting how the CD of the moon changes over time. Similarly, students who note the change in illumination of the moon are noting the PP of the moon. These daily moon observations helped develop students' spatial skills and mental models as reflected by their results on the LPCI. Results of a related study showed a difference in scores on the PP Spatial-Mathematical domain of the LPCI in favor of the experimental females over the control group females (Wilhelm, Toland, Jackson, Cole, & Wilhelm, 2014). This difference was partly attributed to the daily moon observations, since the experimental group kept moon journals and the control group did not.

Conclusion and Significance

While students need other experiences to fully understand the motions and phases of the Moon, daily Moon observation journals create a link between the classroom experiences and the natural world students experience on a daily basis. The journals also provide an opportunity for students to begin to ask questions about and make sense of what they are seeing in the sky and in the classroom much like the notebooks of early scientists show how they thought about and began to make sense of the natural world around them (Tweney, 1991). When this purposeful connection is made, students are better able to respond to assessment questions that require knowledge of lunar phases as well as those that incorporate spatial ability.

Our data suggest teachers and their students may benefit from the purposeful use of moon journals in their classrooms. The Moon needs to be observed daily and over a long enough period of time (at least 5 weeks) to observe the entire lunar cycle and see the pattern repeat. These observations alone are not enough. In-class discussion of the moon journals also needs to take place, similar to some of our teachers using the moon journals as their daily warm-up activity. These discussions allow for students to compare their observations with others and come to a consensus. This classroom discourse is important. Students need to write in their moon journals to make sense of their observations, but they also need to discuss their observations with peers to learn from each other and rectify differences in their mental models.

This study is unique in the purposeful link created between student moon observations and spatial skills. The use of moon journals distinguishes this study further by fostering scientific observation along with skills from across science, technology, engineering, and mathematics fields and other disciplines. In addition, there is no prior reported link, to our knowledge, made between moon observation journals and test scores. We believe that future work will continue to show a strong link between these improved spatial skills and performance in mathematics and science.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

Alias, M., Black, T. R., & Gray, D. E. (2002). Effect of instructions on spatial visualisation ability in civil engineering students. *International Education Journal*, *3*(1), 1–12.

Anson, C. M. (1988). Toward a multidimensional model of writing in the academic disciplines. In D. A. Joliffe (Ed.), *Advances in writing research: Writing in academic disciplines* (Vol. 2, pp. 1–33). Norwood, NJ: Ablex.

Applebee, A. N. (1981). *Writing in the secondary school: English and the content areas* (NCTE Research Report No. 21). Urbana, IL: National Council of Teachers of English.

Applebee, A. N. (Ed.) (1984). *Contexts for learning to write: Studies of secondary school instruction*. Norwood, NJ: Ablex.

Baxter, G. P., Bass, K. M., & Glaser, R. (2000). *An analysis of notebook writing in elementary science classrooms*. (Center for the Study of Evaluation Report 533). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Black, A. (2005). Spatial ability and earth science conceptual understanding. *Journal of Geoscience Education*, *53*(4), 402–414.

Chancer, J., & Rester-Zodrow, G. (1997). *Moon journals: Writing, art and inquiry through focused nature study*. Portsmouth, NH: Heinemann.

Comstock, A. B. (1939). *Handbook of nature study* (24th ed.). Ithaca, NY: Comstock Publishing.

Crosland, M. P. (2004). *Historical studies in the language of chemistry*. Mineola, NY: Courier Dover Publications.

Doris, E. (1991). *Doing what scientists do: Children learn to investigate their world*. Portsmouth, NH: Heinemann.

Duckworth, E. (1996). *The having of wonderful ideas and other essays on teaching and learning*. New York, NY: Teachers College Press.

Fulwiler, B. R. (2007). *Writing in science*. Portsmouth, NH: Heinemann.

Holmes, F. L., Renn, J., & Rheinberger, H. J. (Eds.). (2003). *Reworking the bench: Research notebooks in the history of science* (Vol. 7). Boston: Kluwer Academic Publishers.

Howard, V. A. (1988). Thinking on paper: A philosopher's look at writing. In V. A. Howard (Ed.), *Varieties of thinking: Essays from Harvard's philosophy of education research center* (pp. 84–92). New York, NY: Routledge, Chapman & Hall.

Jones, N., Ross, H., Lynam, T., Perez, P., & Leitch, A. (2011). Mental models: an interdisciplinary synthesis of theory and methods. *Ecology and Society*, *16*(1), 46. Retrieved from http://www.ecologyandsociety.org/vol16/iss1/art46/

Linn, M., & Petersen, A. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*, 1479–1498.

Lindell, R., & Olsen, J. P. (2002). Developing the lunar phases concept inventory. In S. Franklin, J. Marx, & K. Cummings (Eds.), *Physics education research conference proceedings*. New York: PERC Publishing.

Lord, T. R., & Rupert, J. L. (1995). Visual-spatial aptitude in elementary education majors in science and math tracks. *Journal of Elementary Science Education*, *7*(2), 47–58.

McMillan, S., & Wilhelm, J. (2007). Students' stories: Adolescents constructing multiple literacies through nature journaling. *Journal of Adolescent & Adult Literacy*, *50*(5), 370–377.

Needham, R., & Hill, P. (1987). *Teaching strategies for developing understanding in science*. Leeds: University of Leeds.

Newell, G. E. (1983). Learning from writing in two content areas: A case study/protocol analysis (Doctoral diss., Stanford University, 1983). Retrieved from ProQuest Dissertations & Theses Global (Order No. 8314481).

Newman, D., Morrison, D., & Torzs, F. (1993). The conflict between teaching and scientific sense-making: The case of a curriculum on seasonal change. *Interactive Learning Environments*, *3*, 1–16.

Pirnay-Dummer, P., Ifenthaler, D., & Seel, N. M. (2012). Designing model-based learning environments to support mental models for learning. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments* (pp. 66–94). New York, NY: Routledge.

Plummer, J., Wasko, K., & Slagle, C. (2011). Children learning to explain daily celestial motion. *International Journal of Science Education*, *33*(14), 1963–1992.

SAS, Inc. (2007). *All-possible-regressions selection based on PRESS or other statistics* (Version 101) [Computer software]. Cary, NC: SAS, Inc. Retrieved from http://support.sas.com/kb/24/986.html

Tighe, M. A. (1991, March). *Teaching composition across the curriculum in southeastern Alabama and in Suffolk County, England.* Paper presented at the annual meeting of the Conference on College Composition and Communication, Boston, MA.

Tweney, R. D. (1991). Faraday's notebooks: The active organization of creative science. *Physics Education*, *26*(5), 301–306.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352–402.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*(2), 250–270.

Wilhelm, J. (2009). Gender differences in lunar-related scientific and mathematical understandings. *International Journal of Science Education*, *31*(15), 2105–2122.

Wilhelm, J., Jackson, C., Sullivan, A., Wilhelm, R. (2013). Examining differences between preteen groups' spatial-scientific understandings: A quasi-experimental study. *The Journal of Educational Research*, *106*(5), 337–351.

Wilhelm, J., Sherrod, S., & Walters, K. (2008). Project-based learning environments: Challenging preservice teachers to act in the moment. *The Journal of Educational Research*, *101*(4), 220–233.

Wilhelm, J., Toland, M., Jackson, C., Cole, M., & Wilhelm, R. (2014, March 30–April 2). *How instruction, gender, and race affect students' spatial-scientific learning*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, NARST – Pittsburgh, PA.