Developing the Lunar Phases Concept Inventory

Rebecca S. Lindell and James P. Olsen
Physics, Astronomy and Chemistry Education Research Group
Southern Illinois University Edwardsville

The Lunar Phases Concept Inventory (LPCI) was developed to aid instructors in assessing students' mental models of lunar phases. Based upon an in-depth qualitative investigation of students' initial models of lunar phases, this multiple-choice inventory was designed to take advantage of the innovative model analysis theory [1] to probe the different dimensions of students' mental models of lunar phases. The development of this inventory will be discussed, as well as the processes involved in establishing its reliability and validity.

Introduction

Lunar phases may be one of the most difficult concepts to teach in astronomy. Instructors often believe they have successfully taught the concept of lunar phases, only to find that the majority of their students cannot correctly answer simple questions related to the concept.[2] Unfortunately, the solution to this problem is not simple.

A factor contributing to this lack of success may be the instructor's ignorance of students' prior understanding and the effect it may have on their future understanding. Many students enter the astronomy classroom with a prior understanding of lunar phases developed by prior observation and/ or instruction. This prior understanding may be incomplete, incorrect and/or ineffective in explaining lunar phases.[3] Research on cognitive processes illustrate that an individuals' existing knowledge can be thought of as being structured into mental frameworks or models, and new information must be incorporated within their existing models for learning to occur. Mental models can be either well structured or fragmented, but if they are misaligned with scientifically correct ideas, these models will prove a deterrent to learning the material.[4]

To overcome the difficulty in teaching lunar phases, instructors need to be able to determine if their students have existing mental models of lunar phases and how deeply rooted these mental models are. To accomplish this goal, instructors ideally would conduct one-on-one interviews with each of their students; however, this is not practical for the average astronomy instructor with hundreds of students in their classes. A new alternative to traditional qualitative investigations of individuals' understanding is to combine a research-based multiple-choice instrument with the innovative Model analysis theory.[5] This theory purports that by performing a mathematical analysis of individuals' responses to the different items on the instrument, their mental models, as well as how consistently they use these mental models will be revealed. The Lunar Phases Concept Inventory (LPCI) is a multiple-choice instrument designed to utilize model analysis theory to assess college students' mental models of lunar phases

This paper will discuss the development of the LPCI, as well as results from field-testing the instrument at five institutions across the United States. The primary goal of the field-testing was to obtain data to establish the suitability of each item, as well as to establish the reliability and validity the LPCI.

Development of the LPCI

According to Model analysis theory, an individual's mental models can be modeled as a series of concept dimensions relating to a phenomenon, with each dimension having multiple possible facets. The dimensions represent the separate distinct aspects of the participants' framework, while the facets represent the possible value that each dimension may hold. The set of facets represents the scientifically acceptable understanding, as well as alternative conceptions uncovered by detailed qualitative interviews.

It is for this reason, that a detailed qualitative investigation [6] formed the basis for the development of the Lunar Phases Concept Inventory. This investigation yielded detailed information about the different dimensions and facets of students' mental models of lunar phases. Additional facets were taken from the alternative understandings previously uncovered by the literature. Table I shows the LPCI's concept domain with the correct facet listed as the first response under each dimension.

Table I: Lunar Phases Concept Domain

1) Period of the Moon's orbit around the Earth
   A. One month
   B. Less than one month
   C. More than one month
2) Direction of the Moon's orbit around the Earth as viewed from a point above the north pole
   A. Counter-clockwise
   B. Clockwise
3) Period of the Moon's cycle of phases
   A. One month
   B. Less than one month
   C. More than one month
4) Motion of the Moon
   A. Moves like the Sun (East to West)
   B. Moves opposite the Sun (West to East)
5) Phase and Sun-Earth-Moon positions
   A. Correct relationship
   B. Correct angle/ incorrect side of orbit
   C. Incorrect side of orbit
   D. No relation
6) Phase - location in sky - time of observation relationship
   A. Correct relationship
   B. Incorrect relationship
   C. No relationship
7) Cause of Lunar Phases
   A. Alignment of Earth-Moon-Sun
   B. Obstruction by the Earth's shadow
   C. Obstruction by the Sun's shadow
   D. Obstruction by Object
8) Effect on lunar phase with change in location on Earth
   A. No change in Moon's appearance
   B. Moon appears larger
   C. Moon appears smaller

For a multiple-choice instrument to utilize model analysis theory, it needs to have multiple items for each of the different dimensions, with each item addressing only one dimension and the item distracters correspond to the different facets of the dimension.

Based on these test specifications multiple items were developed. Items were field tested under a variety of conditions and situations. After each field test items were reviewed for suitability, and either eliminated or revised if found unsuitable. After many iterations of this process, the reliability and validity of the original 14-item LPCI was established. [7]

To best utilize model analysis theory [8], it was necessary to generate an additional five questions for a total of 19 and perform additional field-testing of the different items. Table II provides the description of the different test items.

Table II: Summary of questions in current LPCI.

1	Where to see waxing crescent Moon at sunset
2	Time to complete one orbit
3	Moon orbits in which direction
4	Cause of new moon
5	Frequency of new moons
6	Phase of Moon for solar eclipse
7	Full Moon in Australia
8	Orbital period vs. phases period
9	Earth-Moon-Sun Positions for new moon
10	Time to observe quarter moon at highest point
11	Direction of Moon rising
12	Shape of a Moon rising at sunset
13	Sun's location in geocentric perspective for particular phase
14	Time until moon appears the same
15	Moon's appearance half-way around the world
16	Alignment between Sun/Earth/ Moon to produce Waxing crescent Moon
17	Time difference between different phases
18	Cause of phases
19	Direction of moon rising

The purpose of this additional field-testing was to sample a different portion of the introductory astronomy population than was sampled in the previous field-tests, as well as to reestablish the reliability and validity of the different items.

Field Testing

The modified version of the LPCI was field tested in the introductory astronomy course at five different institutions: three large state universities, one intermediate state university and one small private college. Unlike the previous field-test sites, which were large Midwestern state universities, these schools represented a variety of different sizes and geographical locations. Data was collected before and/or after instruction, during either the spring semester of 2002, or fall semester 2000 (Kansas State). Table III summarizes the field test sites.

Table III: Summary of Field Test Sites Pre and Post Average Test scores. Maximum score = 19. Differences between schools significant at $\alpha = 0.05$.

Institution		N	Score	SD
Maryland	Pre	132	7.79	3.14
	Post	144	10.44	3.74
Western Maryland	Pre	103	5.89	2.45
	Post	62	7.97	3.00
SIUE	Pre	58	6.51	2.22
	Post	49	10.78	3.69
Kansas State	Pre	54	8.30	3.40
Arizona	Post	91	8.95	3.02

Establishing Item Reliability

Reliability of this instrument was calculated using the *Cronbach's alpha*. This function measures the internal consistency of each item and presents a lower limit on the stability of the students' scores. [9] For this instrument, the coefficient alpha was calculated for 0.55 for the pretest and 0.75 for the posttest. Although the pretest coefficient alpha is low, the post-test coefficient is acceptable. The rather low pretest coefficient may be due in part to student guessing. To further establish the reliability of the LPCI, it will be necessary to use either the test-retest methods or the split-half methods in the future. [10]

Establishing Item Validity

Prior to the field-testing, experts examined the instrument and assessed that it indeed measured the lunar phases content. After the field test, item validity was established using traditional item analysis, [11] as well as concentration item analysis. [12] The traditional item analysis yielded information on the difficulty of each item, as well as how well an item discriminates between high and low scores, while the concentration analysis yielded information on how the items assess students' mental models.

Traditional item analysis defines the item difficulty as the proportion of participants answering an item correctly and the item discrimination as the difference between the proportions of students in the top 27% of the sample answering the item correctly minus the number of participants in the lower 27% of the sample answering the item correctly. [13]

Item difficulties were calculated for each data set individually and then averaged to obtain a value for the overall item difficulty. The item discriminations were calculated from combined scores of all the pre and post-test sites. [14] Table IV shows the item discrimination values for the pre and posttests, as well as the average item difficulties.

Concentration analysis [15] allows one to calculate a concentration factor for each item and provides a method for classifying the level of the concentration factors. The concentration factor yields information on the distribution of student responses to each item, ranging from 0% for equal concentration to 100% for all concentrated on one item. The classification scheme classifies concentration factors above 50% as high (H) concentrations and below 20% as low (L) concentrations. Utilizing this method the concentration factors were calculated and classified for each item, as shown in Table V.

For an item to be considered valid for inclusion on the LPCI it needed to have pre and post discrimination values above 30%, as well as needed to have pre and post concentration factors above 20%. [16] Items failing all four criteria would need serious revision, while items

Table IV: Average Item difficulty (percent correct), standard deviation and discriminations before and after instruction. All values represented in percentages.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Pre Diff.	31	47	41	15	81	23	50	52	24	8	66	13	21	66	34	27	20	18	55
Stdev	2	13	8	5	8	8	5	8	7	2	9	5	5	7	7	5	5	7	7
Discr.	17	63	24	33	44	45	41	48	45	10	38	8	23	52	54	38	21	31	40
Post Diff.	34	73	66	39	87	51	56	65	59	14	72	24	27	71	46	45	31	31	65
Stdev	9	3	13	10	7	14	5	14	12	3	14	14	8	4	5	11	9	10	5
Discr.	31	47	26	68	32	68	62	56	72	8	33	31	23	41	62	61	33	68	46

Table V: Item concentrations and classifications before and after instruction. All values represented in percentages.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Pre	25	39	13	26	66	22	28	23	27	4	45	36	7	48	22	12	34	34	30
Class.	M	M	L	M	H	M	M	M	M	L	M	M	L	M	M	L	M	M	M
Post	24	63	35	27	77	35	38	43	41	06	55	20	11	57	28	23	30	18	46
Class.	M	H	M	M	H	M	M	M	M	L	H	M	L	H	M	M	M	L	M

failing less of the criteria might need slight revision.

Based on these criteria, item 13 was removed completely from the instrument, while items 1, 10 12, 13, 16, 17 and 18 were revised. Additional field-testing is needed to determine if these revisions are adequate or if these items should be thrown out completely.

Acknowledgements

We would like to thank Grace Deming, Jeff Marx and Ed Prather for collecting data for this study. In addition we would like to thank the rest of the PACER group for insightful comments and suggestions.

In addition, we would like to thank the thoughtful comments of the referees.

---

1 L. Bao, *Dynamics of student modeling: A theory, algorithms and application to quantum mechanics*. Unpublished dissertation. University of Maryland (1999).

2 A. Lightman & P. Sadler, "Teacher predictions versus actual student gains" *The Physics Teacher*, 31, 162-167, (1993).

3 See for example: M. Zeilik & W. Bisard, "Conceptual change in introductory-level astronomy courses: Tracking misconceptions to reveal which-and how much-concepts change", *Journal of College Science Teaching*, 29(4), 229-232, (2000);

4 Bransford, J.D., Brown, A. L. and Cocking, R. R., *How people learn: Brain, mind experience and school.* National Academy Press (1999).

5 See ref. 1.

6 R. Lindell, *Enhancing college students' understanding of lunar phases*. Unpublished Dissertation. University of Nebraska-Lincoln (2001).

7 See ref. 1, Chapter 5.

8 See ref. 1.

9 L. Crocker & J. Algina, *Introduction to classical and modern test theory*. New York: Holt, Reinhart and Winston, Inc (1986).

10 See ref. 9.

11 See ref. 6.

12 L. Bao & E. F. Redish, "Concentration analysis: A quantitative assessment of student states" *Physics Education Research Supplement to American Journal of Physics*, 69(7) supplement 1, S46-S53, (2001).

13 See ref. 9.

14 The item discrimination measures how well an item discriminates between low scores and high scores. For this reason, it is appropriate to combine the different pre and post-test sites to calculate the discrimination values.

15 See ref 12.

16 See ref 9 and 12 for guidelines.