A LIVELY ELECTRONIC COMPENDIUM OF RESEARCH, NEWS, RESOURCES, AND OPINION

Astronomy Education Review

2009, AER, 8, 010108-1, 10.3847/AER2009033

Meta-analysis of Planetarium Efficacy Research

Bruce D. Brazell
Navarro College, Corsicana, Texas 75110
Sue Espinoza
Texas A&M University-Commerce, Commerce, Texas 75429
Received: 08/24/09, Published: 09/16/09

Abstract

In this study, the instructional effectiveness of the planetarium in astronomy education was explored through a meta-analysis of 19 studies. This analysis resulted in a heterogeneous distribution of 24 effect sizes with a mean of +0.28, $p < .05$. The variability in this distribution was not fully explained under a fixed effect model. As a result, a random effects model was applied. However, a large random effect variance component indicated that study differences were indeed systematic. The findings of this meta-analysis showed that the planetarium has been an effective astronomical teaching tool.

1. INTRODUCTION

Since the 1960s, a number of studies have compared the effectiveness of planetarium instruction to classroom instruction. Several of these studies produced positive results seemingly supporting the efficacy of planetarium instruction (Dean and Lauck 1972; Wright 1968), while others determined that the traditional classroom was the ideal setting for achieving astronomy objectives (Reed 1970a; Rosemergy 1968; Smith 1966).

The literature review revealed a conflicting picture regarding the instructional efficacy of the planetarium. According to Glass, McGaw, and Smith (1981), meta-analysis is the most appropriate approach to follow in order to resolve the conflicting nature revealed in the research literature on a particular topic. Meta-analysis is an effective way to integrate the findings of multiple studies in order to identify trends and reach generalizable conclusions (Glass, McGaw, and Smith 1981). According to Lipsey and Wilson (2001), "One valuable by-product of meta-analysis is the summary description it provides of the nature of the body of research under inspection" (p. 163). In this article, we present the findings of our meta-analysis of planetarium instructional efficacy studies.

2. A BRIEF DESCRIPTION OF META-ANALYSIS

The research aggregation method known as meta-analysis was pioneered by Gene Glass in 1976. According to Glass, McGaw, and Smith (1981), the main goal of any meta-analytic project is to extract information from the literature in such a way as to reach a general consensus concerning a particular research problem. Lipsey and Wilson (2001) described meta-analysis as "a form of survey research in which research reports, rather than people, are surveyed" (p. 1).

Meta-analytic techniques make use of effect size statistics in order to make comparisons between the outcomes of different studies. In essence, this comparison is possible because the effect size statistic standardizes the findings of a particular study, permitting comparisons between similar studies (Lipsey and Wilson 2001). The effect size statistic represents the magnitude and direction (number of standard deviations above or below the mean) of the treatment interaction between experimental and control group (Glass, McGaw, and Smith 1981). Because effect sizes do not rely on sample size, they are considered to be reliable representations of research outcomes (Thalheimer and Cook 2002). Also, effect sizes can be pooled across similar studies by carefully weighing each study's individual contribution to the aggregation (Lipsey and Wilson 2001).

The goal of any quantitative study is to develop a generalized statement regarding a particular research question (Glass, McGaw, and Smith 1981). Since the outcomes from a single study seldom provide definitive answers, meta-analysis provides a way to find those answers by combining and comparing the outcomes from many different studies on a particular subject.

3. METHODOLOGY

The following steps were followed in carrying out this meta-analytic study:

1. Identify and collect studies.
2. Code features in each study.
3. Quantify the findings on a common scale (effect size).
4. Analyze the data.

3.1. Identifying and Collecting the Studies

Research studies related to planetarium efficacy were found utilizing a variety of techniques. Initially, planetarium studies and related articles were located using online computer databases, accessed via library resources such as Academic Search Complete, ERIC, and UMI ProQuest Digital Dissertations. Second, a comprehensive search through the proceedings and magazines published by the International Planetarium Society was conducted. The references cited in the studies and articles isolated in the initial literature search were analyzed for additional applicable research studies. These procedures resulted in the isolation of 46 unique studies related to planetarium instructional efficacy. Following criterion testing, 19 of these studies were selected for inclusion in the meta-analysis. A brief description of each of these studies can be found in the following section. The remaining 27 studies did not satisfy inclusion requirements and were rejected for one or more of the following reasons:

1. The study was not conducted in a formal educational setting with the participation of students engaged in an astronomy learning unit.
2. The study did not employ the use of a comparison group.
3. The study did not employ an experimental or quasiexperimental design.
4. The study was a comparison between the relative effectiveness of two or more planetarium instructional modes.
5. The study failed to report sufficient statistical information that could be used in the calculation of effect sizes.

3.1.1. Brief descriptions of each of the selected studies

In 1966, Smith conducted one of the earliest studies designed to test the efficacy of the planetarium as compared to classroom. One group of sixth grade students received a 40 min planetarium demonstration over selected astronomical topics. The other sixth grade group received a 40 min classroom demonstration covering the same astronomy subject matter. The study used a post-test control group design where only one treatment, planetarium or classroom, was administered to each treatment group. Groups were randomly assigned according to a table of random numbers. Smith (1966) determined that the achievement of the students who experienced the classroom demonstration was significantly better than those who had attended the planetarium, $F(1,698)=37.62$, $p<.05$. This seemed to show that classroom instruction was more effective than planetarium instruction as measured in the area of student achievement. Smith surmised that the classroom fared better than the planetarium because students were more familiar with this type of learning environment.

Tuttle (1966) also conducted a planetarium versus classroom study with the participation of sixth grade students. A standardized aptitude test measuring student spatial abilities was administered as both pre- and post-test. According to Tuttle, the planetarium group exhibited significant increases in mean test scores over the classroom group in terms of improvement in three-dimensional spatial relations ($t=4.0$, $p<.01$) and improvement in two-dimensional spatial relations ($t=2.4$, $p<.02$). Additionally, an instructor created final exam designed to measure knowledge based achievement was administered to each group following their respective treatments. Tuttle noted that the mean of the planetarium group was significantly higher than the mean of the classroom group, $t = 2.02$, $p < .05$. Tuttle felt that the results obtained in this study indicated the planetarium's power to influence student learning. In particular, Tuttle was convinced that the planetarium could help students develop and improve their inherent spatial abilities.

In 1968, Rosemergy conducted a study to determine if students in the sixth grade would gain a greater understanding of astronomical phenomena from instruction inside a planetarium as compared to those students who were not exposed to the planetarium environment. The astronomical phenomena chosen for the lessons included the sky's daily motion, annual motion of the star field, and the phases of the moon (Rosemergy 1968). Rosemergy conducted an analysis of variance (ANOVA) and found that there was no significant difference in achievement between the groups, $F(2, 336) = 2.55$, $p > .05$. As a result, Rosemergy determined that exposure to the planetarium did not significantly contribute to student understanding of the astronomical concepts and that classroom demonstrations were equally as effective.

Also in 1968, Wright conducted a study in astronomy achievement with the participation of eighth grade students divided into four treatment groups. according to the following scheme: (1) Preplanetarium visit, (2) postplanetarium visit, (3) teacher provided planetarium orientation, postplanetarium visit, follow-up activity, and (4) planetarium staff provided planetarium orientation, postplanetarium visit, follow-up activity (Wright, 1968). A one-way ANOVA was used to determine that there were significant differences between the means of the treatment groups, $F(3, 58) = 7.05$, $p < .01$. In particular, Wright (1968) found that there was a significant difference in achievement between group 1 (tested prior to the planetarium visit) and group 2 (tested immediately following the planetarium presentation), $F(3, 58) = 17.87$, $p < .01$. Wright concluded that the planetarium was a beneficial teaching tool for eighth grade students and that even a short program can produce positive results.

In 1970, Reed conducted a study involving the participation of undergraduate students randomly divided into two treatment groups. The classroom group was exposed to a lecture over the daily and annual motion of the stars and planets using chalkboard and celestial globe demonstrations, while the planetarium group was exposed to the same content (Reed 1970a). The experiment employed a post-test only randomized group design. A *t*-test with an accepted 0.05 significance level was employed to determine the difference between the group means. Results indicated a significant difference between the two teaching methods, $t = 2.49$, $p < .05$. Reed determined that the classroom group showed greater gains than the planetarium group—supporting the idea that the classroom was a more effective learning environment.

Reed conducted a follow up study utilizing the participation of college students enrolled during the 1969–1970 school year. These students were assigned to either the planetarium treatment group or the classroom treatment group. Analysis of the data collected in Fall 1969 showed that the classroom group exposed to the chalkboard/celestial globe demonstrations performed significantly better than the group utilizing the planetarium, $t = 2.4$, $p < .02$. The same result was observed with the Spring 1970 groups where the classroom group's performance on the assessment was significantly better than the planetarium group's scores, $t = 4.1$, $p < .001$. Reed (1970b) concluded that the classroom demonstration was more effective than the planetarium presentation in the "attainment of cognitive behavioral objectives" (p. 55). He further implied that the right combination of classroom instruction and planetarium demonstrations would provide the best learning experience when teaching positional astronomy (Reed 1970b).

Reed and Campbell (1972) reported the results of a different planetarium instructional efficacy study conducted during the Fall 1969 and Spring 1970 semesters with college students as participants. This study employed a post-test only control group design with students randomly assigned to either classroom or planetarium group. During the fall trial, the classroom group performed significantly better than the planetarium group, $t = 2.4$, $p < .02$. Similarly, the classroom group scored significantly better than the planetarium group during the spring trial, $t = 4.1$, $p < .001$. These results supported the conjecture that the classroom was more effective than the planetarium in the teaching of selected astronomical topics as determined by Reed in his previous studies.

Pitluga (1971) conducted a planetarium versus classroom study with fifth grade participants. Students in both the planetarium and the classroom groups were taught a lesson on the phases of the moon. Pitluga found that the classroom group performed significantly better than the planetarium group, $t = 1.649$, $p < .05$. Because this result was similar to that obtained in the studies conducted by Reed, Pitluga was satisfied that his research supported the idea that the classroom was a superior teaching environment than the planetarium.

Also in 1971, Yee, Baer, and Holt conducted a planetarium instructional efficacy study at the high school level. The purpose of this study was to determine if a single visit to the planetarium impacted student learning in terms of achievement and attitudes. The researchers assigned classes randomly to either planetarium or classroom group. This study employed a post-test only control group design with both groups involved in the study taking the post-test. The lesson presented in both the planetarium and classroom consisted of the following astronomical topics: Constellation recognition, basic celestial navigation, stellar magnitude scale, cosmic distance, and a short history of astronomy and recent space exploration (Yee, Baer, and Holt 1971). The researchers found a significant difference between the two groups in favor of the planetarium, $t=4.33$, $p<.001$. In other words, "the planetarium experience had a positive effect on learning and understanding of astronomy in line with its objectives" (Yee, Baer, and Holt 1971, p. 10).

Dean and Lauck (1972) decided to conduct a study similar to the previous classroom versus planetarium studies but differing with respect to the type of assessment instrument used. Accordingly, the researchers employed an oral appraisal, the "Open Sky Test" that was administered outdoors on a clear night individually to each student (Dean and Lauck 1972). This study was conducted at the elementary school level with two different sixth grade classes. The control group was provided three 45 min lessons in the classroom using chalkboard and celestial sphere demonstrations while the experimental group experienced three 45 min lessons in the planetarium over the same astronomical material (Dean and Lauck 1972). The researchers found that the planetarium group scored significantly greater on the assessment than the classroom group, $t=3.446$, $p<.005$. Since the assessment instrument was designed specifically to test for cognitive understanding of positional astronomy, to the researchers, this result verified that the planetarium was an effective teaching tool for this subject (Dean and Lauck 1972).

In 1973, Sunal conducted a planetarium instructional efficacy student at the second grade level. Participants were randomly assigned to one of three groups: (1) Classroom astronomy instruction group, (2) planetarium and classroom astronomy instruction group, and (3) no astronomy instruction in either planetarium or classroom (Sunal 1973). Utilizing an analysis of covariance with a 0.05 level of significance, Sunal determined no significant difference in student performance levels between the classroom group and the classroom/planetarium group. However, a significant difference between the groups exposed to astronomy instruction and the group receiving no astronomy instruction further supports the idea determined in previous studies that the classroom and planetarium are equally effective teaching environments (Sunal 1973). One of the implications of this study was that the goals of the planetarium educator could be achieved when the planetarium was used as a supplement to the regular classroom (Sunal 1973).

Ridky (1975) referred to the influence of the planetarium environment as the "mystique effect" (p. 505), a property of the planetarium environment that is manifested in students who are unfamiliar with the nature and operation of the planetarium. Ridky suggested that an orientation to the planetarium environment conducted prior to instruction would lessen the influence of the mystique effect. In this study, eighth grade students were randomly assigned to two groups with both groups receiving instruction in the planetarium, but only one group would receive an orientation prior to the planetarium lesson. The orientation session provided information over the construction and operation of the planetarium including the purpose of the dome ceiling, functions of major projection devices, and the purpose of the planetarium's lighting system (Ridky 1975). Ridky noted that students who received the orientation performed significantly better than the students who did not. Ridky concluded that the planetarium could be used to provide a positive learning experience if utilized properly.

Hayward (1975) took a different approach toward the problem of planetarium instructional efficacy focusing the subject matter taught in the study on a topic well suited for demonstration in the planetarium—the annual motion of the sun, moon, and planets. Sixth grade students were randomly assigned to one of three groups: (1) Planetarium instruction, (2) classroom instruction, and (3) no astronomy instruction—control group. The group receiving instruction in the planetarium performed well on the objectives based upon observational relationships and in the areas of space/time relationships (Hayward 1975). The classroom group scored well in areas that required recall and classification of astronomical information. In terms of total scores, both classroom and planetarium showed significant gains in achievement when compared to the control group. Although both the classroom and the planetarium group showed gains, Hayward noted that the planetarium simulation seemed better suited to demonstrate observable sky phenomena than the classroom.

Sonntag (1981) conducted a study into the relationship between student spatial abilities and three different astronomy-teaching methods: (1) Planetarium presentation, (2) classroom demonstration, and (3) combination of planetarium and classroom. Sonntag conducted his study with the participation of college seniors enrolled during the fall of 1980. Students were exposed to lessons in positional astronomy grounded on those concepts most appropriate for demonstration in a planetarium: Features of the celestial sphere, annual and diurnal motion, the passage of time, and the causes of the seasons (Sonntag 1981). Sonntag found that approximately two-thirds of the students tended to favor the planetarium or the classroom-planetarium combination treatment over traditional classroom instruction. However, classroom instruction was determined to be significantly superior to instruction in the planetarium as measured by a positional astronomy post-test. When spatial ability was correlated with achievement on the positional astronomy post-test, it was determined to have a dramatic influence. The results of the study seemed to indicate that students with low to medium spatial abilities could benefit from exposure to the planetarium (Sonntag 1981). This study suggested that the power of the planetarium may lie in its capacity to engage student spatial abilities (Sonntag 1981).

A common method used in science education to connect to a student's spatial abilities is to include opportunities to manipulate tactile objects during a learning activity. Edoff (1982) explored the potential of incorporating direct manipulation into the planetarium experience. This study was conducted using a randomly selected sample of fifth and eighth grade students assigned to one of the following groups: Group 1 used a manipulative during the planetarium lesson, group 2 did not use a manipulative during the planetarium lesson, and group 3 did not receive a planetarium lesson (Edoff 1982). Topics included in the lessons were as follows: The celestial sphere and time, seasons, lunar motion, and lunar phases. Utilizing a post-test only control group experimental design, Edoff performed a one-way ANOVA to compare the results of each group's performance. In terms of knowledge retention, Edoff noticed a significant difference between the three group treatments, $F(2,510)=136.02$, $p<.001$, in favor of the planetarium/manipulative group (Edoff 1982). Based upon the study, Edoff felt that using a manipulative in conjunction with the planetarium presentation aided students with knowledge retention. Additionally, Edoff observed that the planetarium groups performed better than the classroom group.

In 1989, a study conducted by Twiest compared students who had experienced an astronomy curriculum with the use of a planetarium to those who experienced the same subject matter only in the classroom. Participating fourth, fifth, and sixth grade students were compared using a pretest–post-test control group experimental design. Astronomical topics covered during the 10 week astronomy unit included the night sky, location and relative positions of the constellations, and motions of celestial objects (Twiest 1989). Twiest found that there was a significant difference in favor of the planetarium treatment at the fourth grade level, $F(1,104)=10.64$, $p<.05$. On the fifth grade level, Twiest determined that there was no significant difference, $F(1,232)=1.10$, $p>.05$, indicating that the planetarium experience did not have an influence on these participants. No significant difference was also observed for the sixth grade participants as well, $F(1,87)=1.04$, $p>.05$. Twiest concluded that the planetarium seemed to have its greatest effect on learning during the student's first encounter at an earlier grade. This study supported the conclusion found in previous studies that the planetarium was an instructional environment that was not more effective than the traditional classroom when used for instruction in the space sciences (Twiest 1989).

In 2000, researchers Baxter and Preece conducted a comparison between the achievement of students using computer planetarium software and those that received instruction in a traditional domed planetarium. This study was conducted in the United Kingdom with the participation of a group of Year 5 and 6 (fourth and fifth grade USA equivalent) and a group of college level preservice science education students. Participants were randomly assigned to computer group or planetarium group during the study. An ANOVA revealed that there was no significant difference between the treatments for the elementary school children participants, $F(1,44)=.184$, $p>.05$. However, Baxter and Preece (2000) did report a significant difference between the treatments when gender was considered in favor of the female participants, $F(1,44)=9.34$, $p<.05$. A similar result was observed when the data from the preservice education students (mostly female) were analyzed. For these participants, Baxter and Preece calculated a large effect size ($d=1.34$) in favor of the planetarium. Based upon this result, Baxter and Preece concluded that the planetarium experiences helped female participants in terms of spatial abilities.

Palmer (2007) conducted a study similar to the one conducted by Pitluga (1971), which focused entirely upon the subject of lunar phases and eclipses. This study employed a quasiexperimental design with the participation of in tact fifth grade classroom groups (Palmer 2007). The classroom group took the post-test immediately after the completion of the fifth grade astronomy unit and immediately prior to their visit to the planetarium. The planetarium group took the post-test immediately following both classroom instruction and the planetarium experience (Palmer 2007). Palmer found that the planetarium presentation significantly influenced student learning, $t(176)=2.657$, $p<.05$. Palmer concluded that the planetarium experience helped students to visualize the concepts of lunar phases and eclipses in the proper perspective thus allowing students to engage their spatial understanding of the concepts. According to Palmer, "It is reasonable to expect that spatial representation demonstrated in a planetarium would help many students construct better mental models of earth, moon, and sun relationships" (p. 58).

3.2. Coding the Studies

Each study in the sample was examined in order to isolate features relevant to the problem of planetarium instructional efficacy. The following independent variables were coded in each study:

1. Publication type (article or dissertation/thesis).
2. Year of publication (1960–1984 or 1985–2008).
3. Grade level (K-12 or college).
4. Random assignment of groups (yes or no).
5. Control of instructor effect (yes or no).
6. Method of test administration (paper or other).
7. Type of assessment instrument used (standard or self-created).
8. Treatment method (lecture or prerecorded).
9. Control method (classroom or computer).
10. Treatment duration (single visit or multiple visits).
11. Astronomical subject matter (observational astronomy or general astronomy information).

Coding procedures also extracted statistical information reported in each study. These outcome measures were the dependent variables and were reported in a variety of statistical forms (means, standard deviations, *t*-scores, and *F*-scores) that were then used to calculate effect sizes.

3.3. Quantifying the Outcomes

The central activity in any meta-analytic study is to compare the outcome measures of each study on a common scale. The common metric used to make this comparison is called the effect size. The effect size is the statistic that makes the comparison of multiple research findings through a meta-analysis possible since it standardizes findings across studies (Wilson 2002). Effect sizes were calculated for each unique comparison represented in the sample of studies. The reader is directed to review *Practical Meta-Analysis* (Sage) by Lipsey and Wilson (2001) for detailed information on calculating effect sizes from various reported statistical information.

3.4. Analyzing the Data

Each study effect size was included in the computation of an overall mean effect size representing the instructional effectiveness of planetariums. Many meta-analysts have used Cohen's rules, which classify mean effect sizes as small (0.2), medium (0.5), or large (0.8). However, Thompson (2002) advises that these rules should only be used as a guideline and that the interpretation of a mean effect size should be made by those more closely aligned with the particular research topic.

A critical question in any meta-analytic study deals with the reason why studies have different effect sizes. "Although the average effect size achieved across all reviewed studies is important, explanatory meta-analyses attempt to explain the variability in obtained effects across studies" (Durlak 1995, p. 321). Variability can introduce a measure of uncertainty in the interpretation of the overall mean effect size. Initially, the distribution of effect sizes was subjected to homogeneity analysis. It was assumed that a homogeneous distribution of effect sizes will differ from the population mean by subject-level sampling error alone (Lipsey and Wilson 2001). However, a heterogeneous distribution exhibits variability that warrants further exploration through the analog to the ANOVA procedures outlined by Lipsey and Wilson (2001).

4. FINDINGS

The 19 studies included in the meta-analysis tested student achievement in a wide variety of astronomical subjects. An effect size was calculated for each independent group using the data extracted from the studies, which resulted in a distribution of 24 effect sizes for student achievement. For the studies conducted by Dean and Lauck (1972), Edoff (1982), Hayward (1975), Palmer (2007), Pitluga (1971), Reed (1970a), Smith (1966), Sonntag (1981), Sunal (1973), Tuttle (1966), Wright (1968), and Yee, Baer, and Holt (1971), only one effect size per study was computed because these studies involved the participation of a single experimental group. However, the studies conducted by Baxter and Preece (2000), Reed (1970b), Reed and Campbell (1972), Ridky (1975), and Twiest (1989) all employed the use of unique experimental groups and thus each contributed more than one effect size to the analysis. The resulting distribution of 24 effect sizes contained 6087 subjects from 19 studies. Because of the small collection of effect sizes, no outliers were identified or eliminated. The effect sizes, inverse variance weights, sample sizes, and mean effect size for these planetarium instructional efficacy studies are shown in Table 1.

Table 1. Effect Sizes for Planetarium Instructional Efficacy

Author	ES	Weight	*N*	
Baxter and Preece 1	+0.01	12.00	48	
Baxter and Preece 2	+1.34	11.95	26	
Dean and Lauck	+0.99	10.68	48	
Edoff	+0.50	81.85	542	
Hayward	+0.38	76.59	312	
Palmer	+0.48	43.18	178	
Pitluga	−0.17	92.56	373	
Reed and Campbell 1	+0.24	98.28	401	
Reed and Campbell 2	+0.42	92.40	378	
Reed 1	−0.44	28.54	117	
Reed 2	−0.72	31.14	133	
Reed	−0.67	10.53	45	
Ridky 1	+0.47	15.07	70	
Ridky 2	−0.02	8.97	36	
Rosemergy	+0.01	53.83	216	
Smith	−0.46	170.33	700	
Sonntag	−0.58	14.11	59	
Sunal	−0.30	102.93	657	
Tuttle	+0.50	14.99	64	
Twiest 1	−0.13	25.95	104	
Twiest 2	+0.10	57.93	232	
Twiest 3	−0.20	21.63	87	
Wright	+1.58	215.99	1187	
Yee, Baer, and Holt	+1.01	16.42	74	
				0.28[a]

[a]$p < .05$

A mean effect size of +0.28 indicated that the planetarium produced a small positive effect on student learning. The 95% confidence interval for this mean effect size ranged from +0.23 to +0.34. Because the confidence interval did not contain zero, the mean effect size of +0.28 was statistically significant, $p < .05$. Additionally, the z-test value for the mean effect size was calculated to be 10.31, which exceeds the critical value of 1.96 at $p = .05$ further indicating statistical significance.

The effect size distribution was also subjected to homogeneity analysis to determine if there were differences between the effect sizes that could be attributed to more than just subject-level sampling error. The Q-value (a test for homogeneity) determined in the analysis was 631.81, which was greater than the critical value of 41.64 obtained from a chi-squared distribution with 23 degrees of freedom and $p = .01$. As a result, the effect size distribution was determined to be heterogeneous, indicating that there was variability in the distribution greater than could be assigned to sampling error alone. Therefore, the distribution was subjected to an analog to the ANOVA procedure, assuming a fixed effects model, in order to isolate the source of this variability. Below is a summary of the more notable results of this analysis.

1. K-12 students experienced a significant positive effect (+0.34) due the planetarium treatment while college students did not respond significantly to the planetarium treatment (+0.09).
2. Students exposed to observational astronomy lessons (+0.51) responded significantly better than students who were exposed to general astronomy information lessons (−0.24) when utilizing the planetarium.
3. In the studies explored, quasiexperimental groups (+0.68) performed significantly better than experimental groups (+0.26) in response to the planetarium treatment.
4. Groups that participated in single visits (+0.32) to the planetarium responded better than those groups that participated in multiple visits (+0.06).

5. Studies published in journal articles (+0.38) reported greater results in response to the planetarium than those published in theses or dissertations (+0.26).
6. Studies published between 1965 and 1986 (+0.30) reported greater results in favor of the planetarium when compared to results published between 1987 and 2008 (+0.20).

In this analysis, it was determined that each of the categorical variables was unable to fully explain the variability in the effect size distribution beyond that of subject-level sampling error. The reason was because the within groups' homogeneity proved to be significant for all categorical subgroups indicating that any residual variability was heterogeneous. According to Lipsey and Wilson (2001), in order to fully account for all variability, a categorical variable must have a significant between group homogeneity and a nonsignificant within group homogeneity, which was not the case for any of the categorical variables subjected to the analog to the ANOVA analysis in the current study.

As a result, the analog to the ANOVA procedure did not fully isolate any of the sources of the original variability in the effect sizes.

It has been suggested by Lipsey and Wilson that in these situations, where there is an amount of variability present in the distribution that cannot be attributed to systematic sources or that has not been isolated in the coding stage of the analysis, the variability in the effect sizes is due to a random source. Hence, a random effect model analysis was performed. Using the SPSS meta-analysis macroprovided by Lipsey and Wilson (2001), this approach resulted in the calculation of a random effects variance component with a value of $v = 0.51$. It was assumed that this value was an additional source of variability and was added to the variance attributed to subject level sampling error already present in the effect size distribution. The random effects variance component was utilized to recalculate the mean effect size and confidence interval of the achievement effect size distribution, resulting in a mean effect size of +0.18 with a 95% confidence interval ranging from −0.12 to +0.47, $p = .2363$. Since this interval contains zero, it is considered to be a nonsignificant result. Additionally, the random effect variance component $v = 0.51$ is much larger than the sampling error estimate (based upon the square of the standard error) of 0.022. According to Lipsey and Wilson, when this occurs, it suggests that the differences between studies are indeed systematic and were not isolated in the coding procedures.

5. DISCUSSION AND CONCLUSIONS

The results of this meta-analysis were in agreement with Sunal's (1976) summative narrative of planetarium efficacy research. After a comprehensive literature review, Sunal reached the conclusion that the "planetarium experience can change student performance in most, if not all, of the goal areas of planetarium education" (p. 348). The present meta-analytic study revealed that planetariums could be used to produce positive effects in student learning especially for grades K–12. As such, K–12 educational institutions should incorporate planetarium instruction into the science curriculum.

A key component of any science curriculum is teaching students how to observe, collect, and record scientific information. These experiences are obtained by exposing students to science exploration in a controlled laboratory environment. According to Tomlinson (1997), "We need to teach these concepts in the classroom (and in the planetarium, if possible) by allowing children to explore, experiment, observe, try, fail, and to use other hands-on approaches while at the same time confronting their misconceptions" (p. 7).

Reed (1973) encouraged the use of the planetarium as a science laboratory setting and suggested that using it "solely as a celestial demonstration chamber" should be avoided. Instead, planetarium instruction should involve active participation of students. The present meta-analysis showed that planetariums have great potential to produce positive effects when utilized for interactive observational astronomy instruction. Through active participation in the observation of the simulated motions of celestial objects, students could gain important practice in the development of science process skills.

The conclusions reached in the meta-analysis were based upon a sample drawn from a relatively small population of planetarium efficacy studies. As such, these conclusions should be taken as only a summary of available planetarium efficacy research literature at the time that this project was conducted. There is much that remains to be done in the area of planetarium instructional efficacy research. The studies in this meta-analysis differed in terms of a number of methodological features. However, the coding process did not isolate all of the factors that influenced the measured efficacy of the planetarium in this study as was apparent in the homogeneity analysis results.

The results of the current meta-analytic study also agreed with Sunal's (1976) conclusion that planetarium instructional efficacy research "must become a far more systematized process with objectives better stated in terms of concrete, observable, and trainable behaviors and should include the establishment of a relationship between research and actual decision making" (p. 348). The methods employed in future planetarium efficacy studies should agree on key issues related to subject matter focus, treatment and control conditions, and experimental design in order to provide for more comparability between studies. Additionally, future studies should report statistics in the form of effect sizes based upon common statistical measures in order to become more readily comparable.

A feature common to all of the studies included in the meta-analysis was the issue of researcher created assessment instruments. Although each researcher subjected their instrument to validity and reliability testing, the specific outcomes measured in each study covered a wide range of astronomical concepts and skills. Therefore, the results of these assessments were not as generalizable as they would have been if these assessments were identical. The planetarium education community should begin the process of developing a standardized assessment instrument that could be utilized to assess student learning. An assessment instrument should be designed to test observational astronomy concepts that are best suited for instruction in the planetarium environment.

The current meta-analysis revealed that the factors that influence learning in the planetarium are many and varied and have not yet been fully isolated. The same situation was realized by Smith (1974) who noted the following.

> Unfortunately, in most comparative studies no single factor is isolated as the critical variable. Thus, comparative studies will usually conclude either in favor of or against the planetarium, but never provide information as to which of the multiple factors operating in the planetarium have assisted or retarded the learning process with respect to some behavioral objective. (p. 2)

The use of a standardized planetarium education assessment instrument will enable researchers to continue the quest to isolate those variables that influence student learning in the planetarium. Furthermore, different studies in planetarium efficacy that utilize the same assessment instrument will be more comparable leading to the potential of developing truly generalizable results concerning the effectiveness of planetariums as instructional tools in astronomy education.

Palmer (2007) found that the planetarium could be an effective tool for astronomy instruction. However, in his literature review, Palmer found "little research into the impact of planetarium experiences on students' cognitive understanding of science concepts" (p. 61). The small number of studies found in support of this meta-analysis indicates that additional research is needed in the area of planetarium instructional efficacy. These studies should be conducted at all levels of instruction and should primarily be comparisons between the planetarium and other methods of instruction. In particular, these studies should focus on specific observational astronomy subjects conducive to the planetarium environment and avoid the multiple topic approach that was prevalent in many of the studies collected for this meta-analysis.

Scott (1985) stated that the planetarium was "an educational resource to be used at all levels of instruction, from elementary to college and adult education" (p. 248). In order to realize this vision, the planetarium must continue to be the subject of educational research. The use of a common methodology, the employment of a standardized instrument, and a focus on observational astronomy topics will aid in the generalizability of future planetarium efficacy studies and assist in the establishment of baselines for the improvement of practices and the encouragement of further research.

References

Baxter, J. H. and Preece, P. F. 2000, "A Comparison of Dome and Computer Planetaria in the Teaching of Astronomy," *Research in Science and Technological Education,* 18, 63.

Dean, N. J. and Lauck, G. M. 1972, "Planetarium Instruction-Using an Open-Sky Test," *The Science Teacher (Washington, D.C.),* 39, 54.

Durlak, J. A. 1995, in *Reading and Understanding Multivariate Statistics*, eds. Grimm, L. G. and Yarnold, P. R., Washington: American Psychological Association, 319–352.

Edoff, J. D. 1982, "An Experimental Study of the Effectiveness of Manipulative Use in Planetarium Astronomy

Lessons for Fifth and Eighth Grade Students," *Dissertation Abstracts International, B: The Sciences and Engineering,* 43, 3496.

Glass, G. V., McGaw, B., and Smith, M. L. 1981, *Meta-Analysis in Social Research*, Beverly Hills, Sage.

Hayward, R. R. 1975, "The Developing and Field Testing of an Instrument Using the Planetarium to Evaluate the Attainment of the Concept of Annual Motion," *Dissertation Abstracts International, B: The Sciences and Engineering,* 36, 4382.

Lipsey, M. and Wilson, D. 2001, *Practical Meta-Analysis. Applied Social Research Methods Series*, Thousand Oaks: Sage, Vol. 49.

Palmer, J. C. 2007, "The Efficacy of Planetarium Experiences to Teach Specific Science Concepts," *Dissertation Abstracts International, B: The Sciences and Engineering,* 68, 939.

Pitluga, L. 1971, "An Experimental Comparison of Planetarium Teaching Programs," M.S. thesis, State University of New York–Oswego.

Reed, G. 1970a, "Is the Planetarium a more Effective Teaching Device than the Combination of the Classroom Chalkboard and Celestial Globe?" *School Science and Mathematics,* 70, 487.

Reed, G. 1970b, "A Comparison of the Effectiveness of the Planetarium and the Classroom Chalkboard and Celestial Globe in the Teaching of Specific Astronomical Concepts," *Dissertation Abstracts International, B: The Sciences and Engineering,* 31, 4580.

Reed, G. and Campbell, J. R. 1972, "A Comparison of the Effectiveness of the Planetarium and the Classroom Chalkboard and Celestial Globe in the Teaching of Specific Astronomical Concepts," *School Science and Mathematics,* 72, 368.

Ridky, R. W. 1975, "The Mystique Effect of the Planetarium," *School Science and Mathematics,* 75, 505.

Rosemergy, J. C. 1968, "An Experimental Study of the Effectiveness of a Planetarium in Teaching Selected Astronomical Phenomena to Sixth-Grade Children," *Dissertation Abstracts International, B: The Sciences and Engineering,* 28, 4959.

Scott, R. L. 1985, "Teaching in the Planetarium," *Journal of College Science Teaching,* 14, 248.

Smith, B. A. 1966, "An Experimental Comparison of Two Techniques (Planetarium Lecture-Demonstration and Classroom Lecture-Demonstration) of Teaching Selected Astronomical Concepts to Sixth-Grade Students," *Dissertation Abstracts International, B: The Sciences and Engineering,* 27, 887.

Smith, T. V. 1974, "The Planetarium in Education: A Review of the Literature," Nova University, ERIC Document Reproduction Service No. ED111658.

Sonntag, M. 1981, "An Experimental Study of Teaching Method, Spatial Orientation Ability, and Achievement in Selected Topics of Positional Astronomy," *Dissertation Abstracts International, B: The Sciences and Engineering,* 42, 4783.

Sunal, D. 1973, "The Planetarium in Education: An Experimental Study of the Attainment of Perceived Goals," *Dissertation Abstracts International, B: The Sciences and Engineering,* 34, 1779.

Sunal, D. 1976, "Analysis of Research on the Educational Uses of a Planetarium," *Journal of Research in Science Teaching,* 13, 345.

Thalheimer, W., and Cook, S. 2002, "How to Calculate Effect Sizes from Published Research: A Simplified Methodology," Retrieved April 14, 2008, from http://work-learning.com/effect_sizes.htm.

Thompson, B. 2002, "Statistical, Practical, and Clinical: How Many Kinds of Significance do Counselors Need to Consider?" *Journal of Counseling and Development,* 80, 64.

Tomlinson, G. 1997, "Grade Appropriate Concepts," *The Planetarian,* 26, 7.

Tuttle, D. 1966, “Effects of the Use of the Planetarium upon the Development of Spatial Concepts Among Sixth Grade Students in Elgin,” M.S. thesis, Northern Illinois University.

Twiest, M. G. 1989, “The Attitudinal and Cognitive Effects of Planetarium Integration in Teaching Selected Astronomical Concepts to Fourth, Fifth, and Sixth-Grade Students,” *Dissertation Abstracts International, B: The Sciences and Engineering,* 51, 473.

Wilson, D. 2002, “Practical Meta-Analysis,” Presentation to the American Evaluation Association, http://mason.gmu.edu/~dwilsonb/ma.html.

Wright, D. L. 1968, “Effectiveness of the Planetarium and Different Methods of its Utilization in Teaching Astronomy,” *Dissertation Abstracts International, B: The Sciences and Engineering,* 3, 507.

Yee, A. H., Baer, J. M., and Holt, K. D. 1971, *An Evaluation of the Effectiveness of School Planetarium Experiences, Educational Technology Research No. 41*, Englewood Cliffs: Educational Technology Publications.

ÆR
010108-1–010108-11