JESEH

VOLUME
3
ISSUE
1

YEAR
2017

Journal of Education in Science, Environment and Health

e-ISSN 2149-214X

JESEH
Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017
e-ISSN:2149-214X

EDITORIAL BOARD

Editors

Section Editors

Editorial Board

Technical Support

Journal of Education in Science, Environment and Health (JESEH)

The Journal of Education in Science, Environment and Health (JESEH) is a peer-reviewed and online free journal. The JESEH is published quarterly in January, April, July and October.The language of the journal is English only. As an open access journal, Journal of Education in Science, Environment and Health (JESEH) does not charge article submission or processing fees. JESEH is a non-profit journal and publication is completely free of charge.

The JESEH welcomes any research papers on education in science, environment and health using techniques from and applications in any technical knowledge domain: original theoretical works, literature reviews, research reports, social issues, psychological issues, curricula, learning environments, book reviews, and review articles. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to the JESEH.

Abstracting/ Indexing

Journal of Education in Science, Environment and Health (JESEH) is indexed by following abstracting and indexing services: SOBIAD, Scientific Indexing Service (SIS), Education Resources Information Center (ERIC).

Submissions

All submissions should be in electronic (.Doc or .Docx) format. Submissions in PDF and other non-editable formats are not acceptable. Manuscripts can be submitted through the journal website. All manuscripts should use the latest APA style. The manuscript template for formatting is available on the journal website.

Contact Info

Journal of Education in Science, Environment and Health (JESEH)
Email: jesehoffice@gmail.com
Web : www.jeseh.net

Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017
e-ISSN:2149-214X

CONTENTS

Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

Relationship of Middle School Student STEM Interest to Career Intent

Rhonda Christensen, Gerald Knezek
University of North Texas

Article Info

Article History

Received:
01 April 2016

Accepted:
16 August 2016

Keywords

STEM Education
STEM interest
Career intent
Gender
Climate change

Abstract

Understanding middle school students' perceptions regarding STEM dispositions, and the role attitudes play in establishing STEM career aspirations, is imperative to preparing the STEM workforce of the future. Data were gathered from more than 800 middle school students participating in a hands-on, real world application curriculum to examine the relationship of the students' interest in STEM and their intentions to pursue a career in a STEM field. Among the middle school students who completed surveys for the MSOSW project, 46.6% expressed a desire to pursue a career in STEM at the time of the post test. Regarding alignment of positive interest in STEM with intent to purse a STEM career, middle school students who have stated that they plan to pursue a career in STEM, also show higher dispositions toward STEM and STEM career measures. Gender differences were also examined, resulting in the finding that middle school males generally have greater intent to pursue a career in STEM, and also show more positive interest in STEM areas. However, females appear to more positively react to the project activities presented in this study than males, so over the course of a project year females tend to "catch up." This is true regarding assessed STEM interest as well as stated intent to pursue a career in STEM. These findings provide additional contributions to the growing base of knowledge about the importance of middle school aspirations for STEM careers.

Introduction

A recent study by the American College Testing service (ACT) found a large gap between student interest in Science, Technology, Engineering and Mathematics (STEM) subjects and the intent to take difficult courses required for a STEM major and obtain a job in the chosen field (ACT, 2015). ACT researchers observed that far too many STEM-interested students are not well prepared to succeed in the rigorous college math and science coursework required of STEM majors. Teachers and curriculum can influence both proficiency and interest in STEM content. Researchers at the Donahue Institute at the University of Massachusetts analyzed effective programs aimed at increasing student proficiency and interest in STEM careers and found that among the common characteristics were challenging, hands-on, real-world learning activities mediated by an engaged, knowledgeable teacher (Bouvier, 2011). Proficiency and interest in STEM must be initiated before students reach secondary school and begin choosing their courses. The authors of the study presented in this paper have been assessing middle school student interest in STEM content and careers for six years. The purpose of this paper is to contrast middle school interest in STEM with intent to pursue a STEM career, in hopes of shedding further light on the types of gaps between interest and intent that exist at the middle school level.

Literature Review

Research on the relationship between student interest in and the pursuit of STEM careers has increased in recent decades. One reason for the lack of the pursuit of STEM careers is that students may lack exposure to the career possibilities in the STEM fields at an early enough age and therefore lack the information they need to consider a career in a STEM field. Various strategies for connecting early interest in and pursuit of STEM careers include project-based and hands-on learning that involve personal and real world relevance (Christensen & Knezek, 2015). One study concluded that students do not see science as being personally relevant to them and need opportunities for career awareness to realize that a career in science is a viable option (Palmer, 1997).

One study seeking to determine a link between elementary students' perceptions and career interests included both student drawings of a scientist as well as student interviews (Buldu, 2006). While the researcher examined the sterotyping of scientists, the conclusion was that student images of a scientist are important indicators of their future plans regarding science (Buldu, 2006). Related research with middle school students found that their occupational preferences and career aspirations are strongly linked to their images of careers (Gottfredson, 1981). Using data from the National Educational Longitudinal Study (NELS), Tai, Liu, Maltese and Fan (2006) found that students who indicated an interest in a career in science when they were in middle school were three more times likely to graduate from college with a science degree.

A study conducted by the Girl Scouts of America compared females interested in STEM fields to those who were not interested in STEM fields. The researchers (Modi, Schoenberg, & Salmond, 2012) found that those who were interested in STEM fields were higher achievers, betters students, had stronger support systems and had exposure to STEM fields. Other factors that have been shown to influence females' perceptions of pursuing a career in STEM are stereotypes regarding performance in mathematics and science areas (Walton & Spencer, 2009; Nguyen & Ryan, 2008) social and cultural cues that discourage girls (Bisland et al., 2011), as well as a lack of confidence in the ability to persevere through difficult material (Dweck, 2006; Halverson, 2011).

This study was guided by the following research questions examined within the context of the pre-post, treatment versus comparison group context of the Going Green! Middle Schoolers Out to Save the World (MSOSW) project:

1. What level of STEM career interest exists among middle school students in the Going Green! MSOSW Project?
2. To what extent is intent to have a career in STEM aligned with interest in STEM for MSOSW students?
3. What differences in level and alignment are attributable to intervening variables such as gender?

Method

Environment for STEM-Focused Intervention

The Going Green! Middle Schoolers Out to Save the World (MSOSW) project has the primary goals of developing middle school students' interest in STEM content areas and raising students' perceptions and attitudes toward STEM careers. MSOSW aims to direct middle school students' enthusiasm for hands-on activities toward long-term interest in STEM while guiding them to solve real-world problems. Students in this project were instructed by their teachers to use energy monitoring equipment to audit standby power consumed by electronic devices in their homes and communities.

Standby power (also called "vampire power") is the electricity consumed by many appliances when they are plugged in but "turned off" (U.S. Department of Energy, 2011). Many appliances consume some electricity while not performing any useful function. Televisions, game consoles, home computers, coffee makers and microwaves are a few of the appliances that commonly consume standby power. The U.S. Department of Energy has estimated that over the lifetime of a typical home appliance, 70% of the power consumed will be when the appliance is turned off (U.S. Department of Energy, 2011). During MSOSW project activities, sixth and seventh grade students learned to measure the vampire power consumption of various appliances in their homes. After measuring standby power, students gathered their data together with classmates in spreadsheet projections to explore energy conservation plans that could lower a family's monthly electric bill and reduce the greenhouse gas emissions that contribute to global climate change. Students shared their results with other participating middle school students from across the United States by contributing to a shared project wiki.

Participants

Data were gathered from treatment and comparison students during the 2013-2014 school year. Pre test data were gathered from 813 students at the beginning of the school year primarily during September – October. When treatment groups completed their unit, these students completed the post test surveys. Comparison students provided post test data in April and May of 2014. Post test data were received from 916 students. Gender distribution was almost equal between males and females with 53.9% males and 46.1% females. A slightly increased percentage of females were among the completed surveys at post test time, compared to pre test time, with 51.9% males and 48.1% females at the time of post test data collection. Male versus female post test data percentages were very close to the U.S. national distribution of 51.2% male and 48.8% female published for the age group 10-14 years in the 2010 U.S. Census (2010).

Instrumentation

The instruments used in this study were the STEM Semantic Survey and the Career Interest Questionnaire (CIQ) and are included in the appendices. The survey battery also included one item to capture students' interests in pursuing specific types of STEM careers. MSOSW participants selected one choice from the following career-oriented question: I plan to have a career in: (1) Science, (2) Technology, (3) Engineering, (4) Mathematics, (5) Other. In addition, demographic items including gender were gathered from the participants.

STEM Semantic Survey

The STEM Semantics Survey was adapted from Knezek and Christensen's (1998) Teacher's Attitudes Toward Information Technology Questionnaire (TAT) derived from earlier Semantic Differential research by Zaichkowsky (1985). The five most consistent adjective pairs of the ten used on the TAT were incorporated as descriptors for target statements reflecting perceptions of Science, Math, Engineering and Technology. A fifth scale representing interest in a career in Science, Technology, Engineering, or Math (STEM) was also created.

Internal consistency reliabilities for the five scales of the STEM Semantics Survey typically range from Alpha = .90 to Alpha = .94 for students such as those participating in this study (Tyler-Wood, Knezek & Christensen, 2010). These reliability estimates fall in the range of "excellent" according to guidelines provided by DeVellis (1991). The five scales had five items each and each item was presented as semantic adjective pairs (fascinating: mundane; exciting: unexciting; and so forth) to describe STEM dispositions and career attitudes.

Career Interest Questionnaire (CIQ)

The Career Interest Questionnaire (CIQ) is a Likert-type (1 = strongly disagree to 5 = strongly agree) instrument composed of 13 items on three scales. This instrument was adapted from a longer instrument developed for a Native Hawaiian Studies project promoting STEM interest in Hawaii (Bowdich, 2009). The subscales of the CIQ document students' perceptions of being in an environment that is supportive of science careers (four items, referred to as Interest), students' intent to pursue educational opportunities that would lead to a science career (five items, referred to as Intent), and the perceived importance of science careers overall (four items, referred to as Importance).

Cronbach's alpha for the CIQ typically ranges from .70 to .93 across subscales (Christensen, Knezek, & Tyler-Wood 2014), and thus falls in the range of "respectable" to "excellent" according to guidelines by DeVellis (1991). Internal consistency reliability for the overall 13-item instrument typically approaches or surpasses .90.

Aspirations for STEM Careers

In addition to the instruments described, one item was also included to capture students' interests in pursuing specific types of STEM careers. MSOSW participants selected one choice from the following career-oriented question: *I plan to have a career in: (1) Science, (2) Technology, (3) Engineering, (4) Mathematics, (5) Other.* For analysis purposes of the paper, data were recoded as STEM and non-STEM interest in a career by coding any of the individual STEM areas as STEM and the "Other" as non-STEM.

Survey Instrument Reliabilities

Cronbach's alpha for the scale reliabilities are listed in Table 1. The reliabilities for pre and post were similar for this group of students with slightly less reliability for post test (end of year). The STEM Semantic scales ranged from .85 to .90 at pretest time. The CIQ parts ranged from .79 to .93 at pretest time.

Table 1. Instrument reliabilities by scale for pretest and posttest

	No. Items	Pretest	Posttest
STEM Semantics Survey			
Science STEM	5	.87	.88
Technology STEM	5	.85	.86
Engineering STEM	5	.90	.90
Mathematics STEM	5	.87	.87
Career STEM	5	.90	.89
Career Interest Questionnaire			
CIQ Part 1	4	.85	.83
CIQ Part 2	5	.93	.92
CIQ Part 3	4	.79	.74
CIQ All			

Results and Discussion

Research Question 1: Level of Interest and Intent

Among the middle school students who completed surveys for the MSOSW project, 44.9% (n = 365 / 813) expressed a desire to pursue a career in STEM at the time of the pretest, while 46.6% (n = 427 / 916) expressed a desire to pursue a career in STEM at the time of the post test. Thus one answer to research question one is that 45-47% of middle school students stated they plan to have a career in science, technology, engineering, or mathematics, when presented with a choice of selecting one of these four STEM areas, or "other".

The corresponding middle school student dispositions toward a career in STEM were captured in their Semantic Differential ratings of "To me, a career in STEM is …" on a 1 (least positive) to 7 (most positive) scale. At the time of the pretest, the mean rating for the 813 respondents was 4.93 (STD = 1.62) while at the post test time the mean rating for the 916 respondents was 5.02 (STD = 1.50). Since the middle (neutral) rating on a 7-point Semantic Differential scale is 4.0, the mean ratings of 4.93 at pretest and 5.02 at post test both represent slightly positive dispositions toward a career in STEM.

If we accept "I plan to have a career in science, technology, math, or engineering" as a middle school student's indication of intent to have a career in STEM, and we accept the Semantic Differential STEM disposition as a middle school student's indication of interest in a career in STEM, then we can formulate expected values (most probable occurences) for any one student for each. In particular, the expected *interest* in STEM for any one student in the MSOSW data sample is slightly positive (upper edge of 4/6 possible intervals = 66th percentile), while the expected *intent* that any one student in the MSOSW data sample plans to have a career in STEM is less than 50%.

Research Question 2: Alignment of Interest with Intent

Research question two addresses the extent to which intent to have a career in STEM is aligned with interest in STEM. This is first addressed through analysis of variance to examine whether the STEM dispositions of middle school students who stated they plan to have a career in STEM, are more positive than the dispositions of those who indicated preference for a career other than STEM.

Overall ANOVA for STEM Career Planned vs. Other

When assessed at the aggregate data level (treatment and comparison, pre and post combined), large differences were found between group means for students who identified themselves as planning to have a career in STEM versus those who did not. These differences were significant ($p < .0005$) on all eight STEM disposition indices gathered from the STEM Semantics and the CIQ. This finding encouraged the research team to examine whether hypothesized differences based on project objectives, such as pre-post changes in STEM dispositions, and treatment versus comparison effects, might be influenced by whether or not the student participants viewed themselves as STEM career candidates. The research team also sought to examine whether participation in project activities influenced student interest in STEM as a Career. These issues will be examined in this section.

ANOVA Comparing STEM Career Planned vs. Other, Pre and Post.

Tables 2 and 3 contains results of analysis of variance of STEM disposition data from all Going Green! MSOSW students at the pretest time, before any of the students had participated in project activities, and also an analysis at post test time. As shown in Tables 2 and 3, at the pretest time, students who indicated they were interested in a career in STEM were significantly ($p < .0005$) more positive on all eight recorded STEM-related measures. At the post test time, all eight measures remained more positive ($p < .001$) for students who indicated they planned to have a career in STEM. This can be viewed as a form of cross-validation of the single-item indicator of career intent, in that high dispositions toward STEM would be expected to accompany intent of interest in STEM as a career. These findings are graphically displayed in Figures 1 and 2.

Table 2. Oneway analysis of variance for non-STEM versus STEM pre and post on STEM semantic measures

		Pretest				**Posttest**			
		N	Mean	SD	Sig	N	Mean	SD	Sig
STEM Science Scale	NonSTEM	447	4.54	1.52		488	4.69	1.44	
	STEM	364	5.28	1.34		426	5.35	1.35	
	Total	811	4.87	1.49	.000	914	4.99	1.43	.000
STEM Math Scale	NonSTEM	445	4.27	1.60		483	4.22	1.59	
	STEM	363	4.72	1.60		426	4.57	1.63	
	Total	808	4.48	1.61	.000	909	4.39	1.62	.001
STEM Engineering Scale	NonSTEM	445	4.38	1.51		480	4.31	1.58	
	STEM	363	5.36	1.49		423	5.19	1.53	
	Total	808	4.82	1.58	.000	903	4.72	1.62	.000
STEM Tech Scale	NonSTEM	443	5.22	1.46		485	5.44	1.40	
	STEM	365	5.76	1.30		425	5.84	1.32	
	Total	808	5.46	1.42	.000	910	5.63	1.38	.000
STEM Career Scale	NonSTEM	448	4.31	1.57		487	4.44	1.42	
	STEM	364	5.70	1.33		427	5.69	1.30	
	Total	812	4.93	1.62	.000	914	5.02	1.50	.000

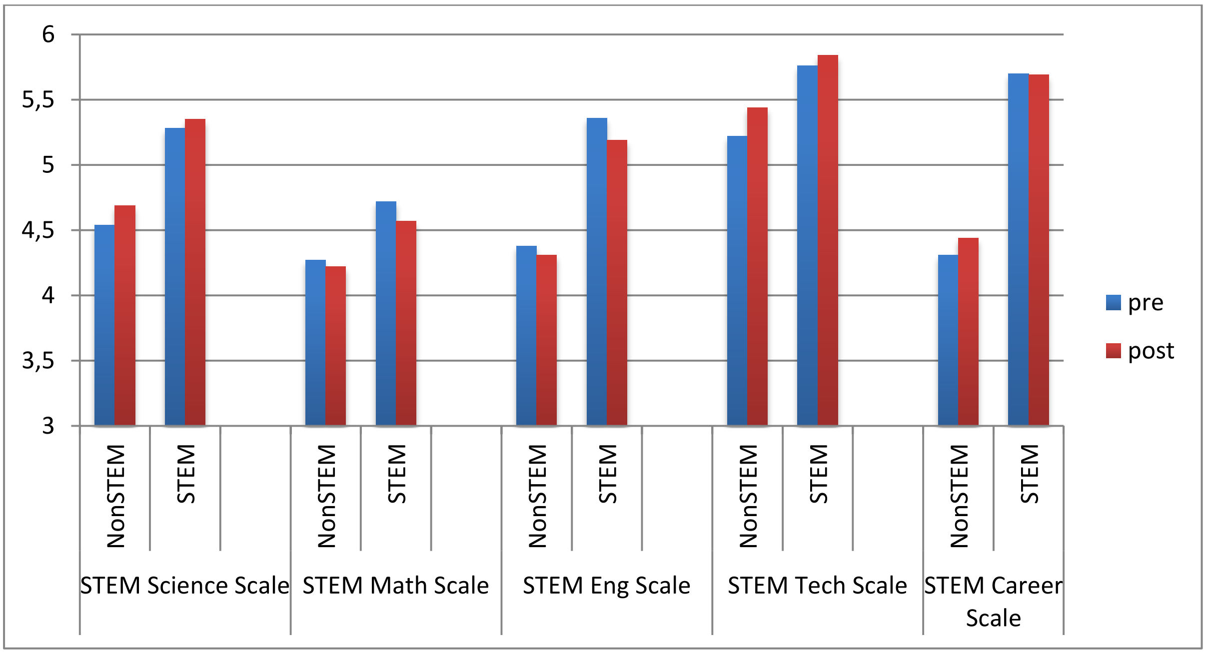


Figure 1. Comparison of STEM semantic ratings for NonSTEM versus STEM career plans, before and after treatment activities

Table 3. Comparison of non-STEM versus STEM student groups on career interest questionnaire measures before and after treatment activities

		Pretest				**Posttest**			
		N	Mean	SD	Sig				
CIQPart1	NonSTEM	446	2.63	.87		487	2.71	.88	
	STEM	363	3.40	.91		426	3.40	.948	
	Total	809	2.98	.97	.000	913	3.03	.970	.000
CIQPart2	NonSTEM	446	2.64	.89		488	2.73	.90	
	STEM	363	3.43	.96		425	3.53	.943	
	Total	809	2.99	1.01	.000	913	3.10	1.00	.000
CIQPart3	NonSTEM	446	3.37	.86		488	3.49	.84	
	STEM	363	3.81	.85		424	3.93	.75	
	Total	809	3.57	.88	.000	912	3.70	.83	.000
CIQ All	NonSTEM	446	2.86	.77		488	2.96	.74	
	STEM	363	3.54	.81		426	3.61	.78	
	Total	809	3.17	.85	.000	914	3.26	.82	.000

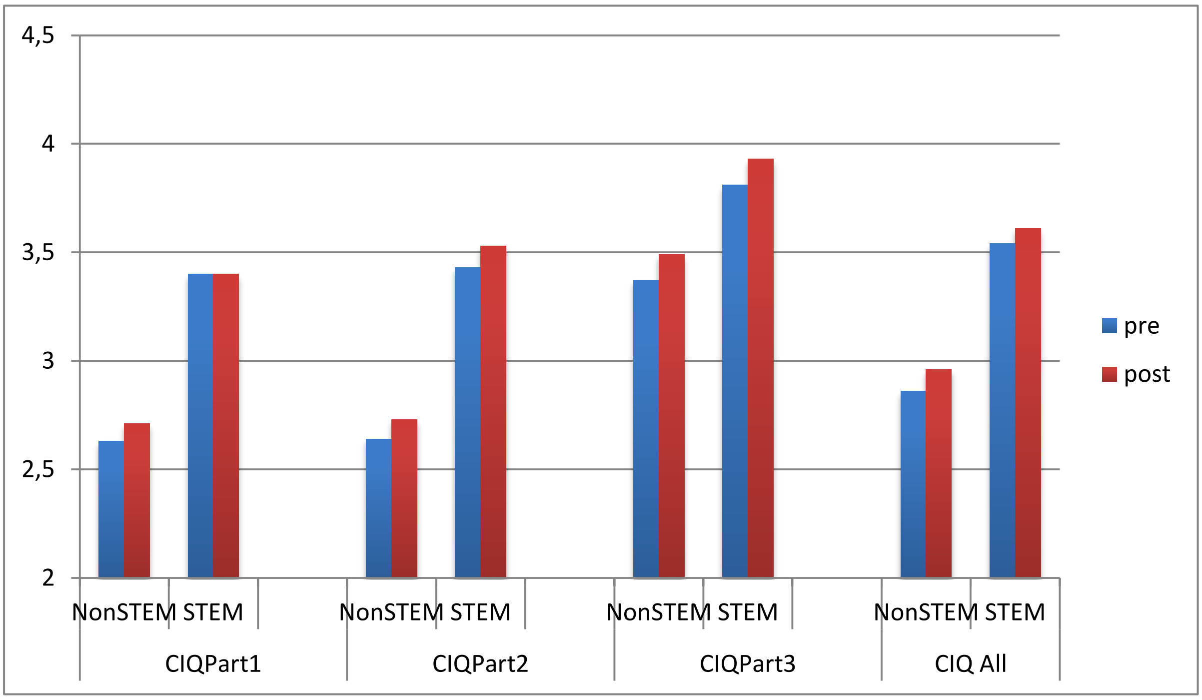


Figure 2. Comparison of non-STEM versus STEM on career interest questionnaire measures, pre and post project activities

Viewing these findings regarding alignment of positive interest in STEM with intent to purse a STEM career, we conclude that with regard to research question 2, middle school students who have stated that they plan to pursue a career in STEM, also show higher dispositions toward STEM and STEM career measures.

Research Question 3: Effects Due to Gender

Male vs. Female Plans to Have a Career in STEM

Table 4 illustrates changes in the students' preference for careers in STEM vs. non-STEM by gender, during the school year. As shown in the non-STEM rows of the table, the proportion of males versus females remained constant from pre to post. Overall, a greater percentage of males indicated preference for STEM careers at both the pre and post test time periods, but the proportion of females planning a STEM career increased from 36% at pretest to 41% at the post test time frame.

Table 4. Male vs. female percentage of students planning STEM career

	Pretest		Post test	
	Male	Female	Male	Female
STEM	64.1% n= 234	35.9% n = 131	58.8%, n = 251	41.2%, n = 176
NonSTEM	45.5% n = 204	54.5% n = 244	45.4%, n = 222	54.6%, n = 267

Treatment vs. Comparison Groups

As shown in Table 5, already at the time of the pretest, comparison group females who indicated interest in a non-STEM career path were much higher in dispositions toward Science than the treatment females who selected the non-STEM career path. The magnitude of the difference was an effect size (Cohen's d) of (4.60-3.90)/1.49 = .47. This is a moderate effect according to guidelines published by Cohen (1988). Reasons for this level of pre-treatment difference are currently unknown.

Positive Effects on STEM Career Intent for Treatment Group Females

Also shown in Table 5, and graphically illustrated in Figure 3, is that for the treatment group of female students, the pre-post gains in dispositions toward Science during the school year of energy monitoring activities were almost identical whether the females began the project year with an intent to have a career in STEM or not. For the males (not illustrated), the effect was very different, in that those who began the year with an intent to have a career in STEM became more positive in their dispositions toward Science during the treatment school year, but the males who began their year not intending to pursue a career in STEM had no significant pre-post change in their dispositions toward Science.

Table 5. STEM semantics science measure for males and females by treatment and comparison

STEM/Non	**Treatment/Comparison**	**Pre/Post**	**Male**	**Female**
STEM	Treatment	Pre	5.29	5.61
		post	4.90	5.81
	Comparison	pre	5.16	5.41
		post	5.31	5.39
NonSTEM	Treatment	Pre	4.85	3.90
		Post	4.75	4.36
	Comparison	Pre	4.57	4.60
		Post	4.68	4.74

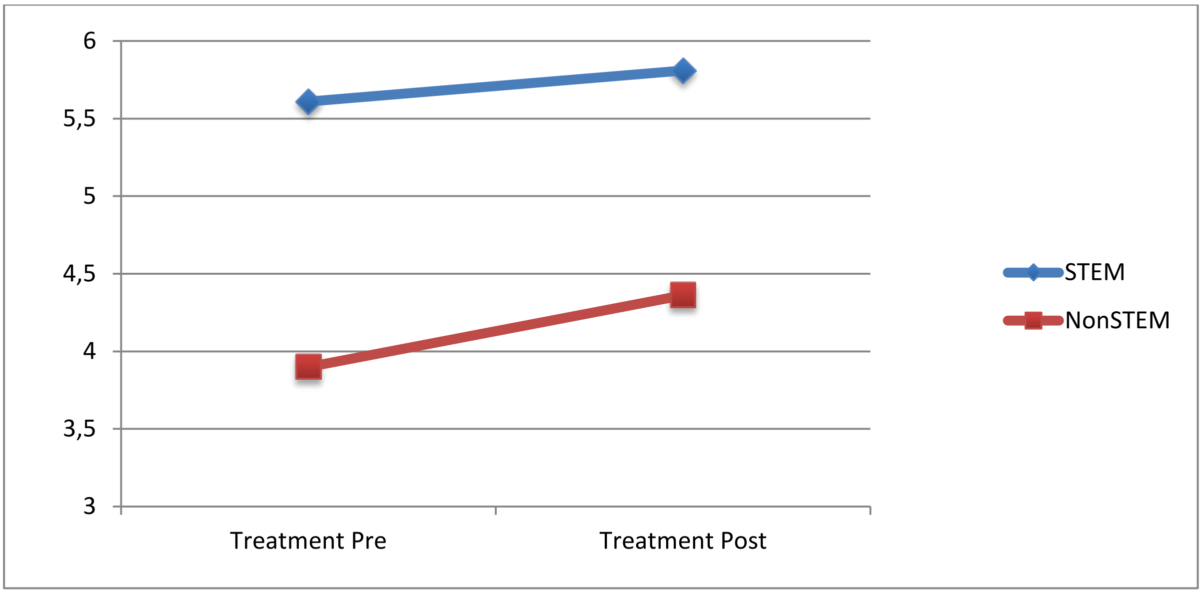


Figure 3. Pre-post changes for females on the STEM science indicator by whether they selected a STEM career or non-STEM career (treatment students)

Positive Effects on STEM Career Interest for Treatment Group Females

As shown in Table 6 and graphically displayed in Figure 4, females who were in the treatment group also gained in their interest in a STEM career from pre to post whether they indicated an interest in STEM as a career path or not. However, females who indicated their interest in a STEM career increased more in their positive dispositions toward STEM as a career than the females who indicated an interest in a non-STEM career. The trends (not illustrated) were not the same for the males, and were not consistent for males with regard to treatment and comparison group pre-post changes in dispositions toward STEM as a career.

Table 6. STEM semantics career measure

STEM/NonSTEM	**Treatment/Comparison**		**Male**	**Female**
STEM	Treatment	Pre	5.87	5.68
		Post	5.34	5.96
	Comparison	Pre	5.67	5.70
		Post	5.64	5.79
NonSTEM	Treatment	Pre	4.24	3.56
		Post	4.25	3.75
	Comparison	Pre	4.32	4.48
		Post	4.56	4.47

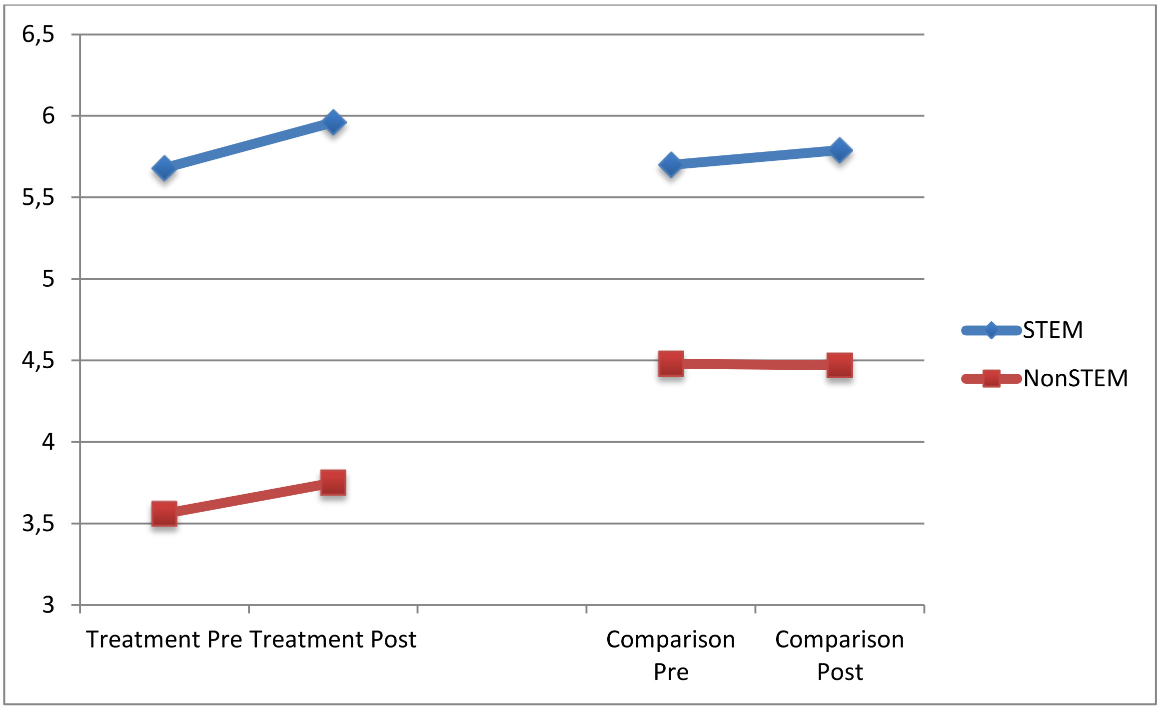


Figure 4. Pre and post comparison on STEM semantics career measure for females who selected STEM as a career plan versus those who chose a non-STEM career plan

Positive Effects on Career Interest Questionnaire Importance Subscale for Treatment Group Females

As shown in Table 7 and graphically displayed in Figure 5, female gains pre to post also became more positive when the chosen indicator is CIQ part 3. This subscale indicator is known as Career Interest Questionnaire Importance (Peterman, Kermish-Allen, Knezek, Christensen, & Tyler-Wood, 2016) and has previously been found to be especially meaningful to female students in that it contains a strong component of "making a difference in the world." (Ceci, Williams, & Barnett, 2009; Christensen, Knezek, & Tyler-Wood, 2015). Trends shown in Figure 5 from the CIQ Importance subscale parallel the major findings shown in Figure 4, from the STEM Semantic Career Interest scale. Female students tend to become more positive in their interests in a STEM career after having participated in the MSOSW hands-on science activities, whether they began the school year with an intent to have a career in STEM or not. For male students the effect is not as consistent and needs more research to identify clear trends.

Table 7. CIQ Part 3 interest in a science career

STEM/NonSTEM	Treatment/Comparison		Male	Female
STEM	Treatment	Pre	3.85	3.97
		Post	3.52	4.21
	Comparison	Pre	3.76	3.85
		Post	3.93	3.97
NonSTEM	Treatment	Pre	3.41	3.49
		Post	3.35	3.65
	Comparison	Pre	3.24	3.46
		Post	3.40	3.56

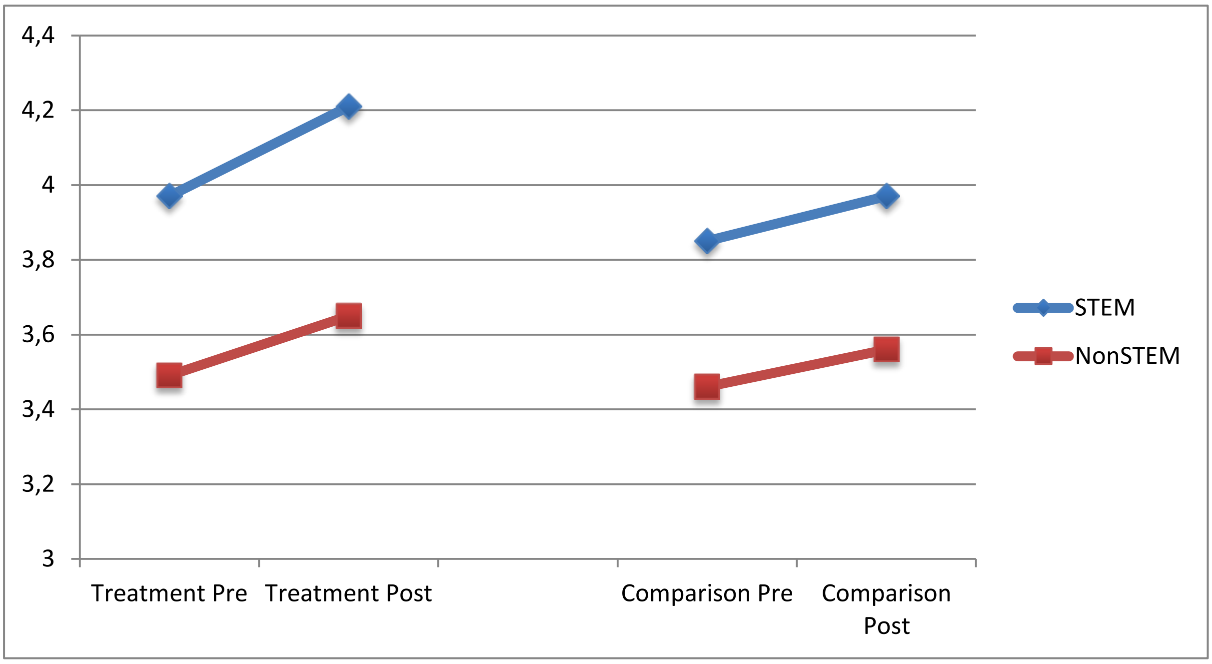


Figure 5. CIQ part 3 pre-post comparisons for females in treatment and comparison by whether they chose STEM as a career plan or another non-STEM career plan

Viewing these findings regarding gender as a whole, we conclude that with regard to research question three, gender differences do exist regarding intent to pursue a STEM career as well as level of interest in STEM as a career. Middle school males generally have greater intent to pursue a career in STEM, and also show more positive interest in STEM areas. However, females appear to more positively react to the MSOSW activities presented in this study than males, so over the course of a project year females tend to “catch up.” This is true regarding assessed STEM interest as well as stated intent to pursue a career in STEM.

The ACT report (2015) pointed out that approximately 34% of students nationwide have a stated and measured intent to pursue a STEM career. In this study, approximately 64% of the students were measured to have a positive interest in STEM (> 4.0 on STEM Semantic scale of STEM as a Career), while 46% reported that they planned to pursue a career in STEM. Three hundred fifty-two (352) of 916 middle school students (38%) in the 2013-2014 MSOSW project reported at the end of the school year that they both planned to have a career in STEM, and also had their interest in STEM as a career measured as positive (> 4.0) on 7-point Semantic Differential Scale. This 38% can be compared with the 34% reported in the ACT (2015) study and reiterates hope that the current group of middle school students across the USA, or at least those with science teachers willing to voluntarily participate in energy consumption and CO_2 production teaching and learning activities, are poised to become a STEM workforce pool of the future larger than previously reported.

Nevertheless, even with the optimistic outcomes reported in portions of this paper, there is still much work to be done. Females in the MSOSW Project were impacted by hands-on project activities that have a real world focus whether or not they expressed an interest in a STEM career. This is an important finding because even if they do not pursue a career in a STEM field, having positive dispositions toward STEM may impact the decisions they make as informed citizens, including but is not limited to those who will be future parents. Further disaggregation of the same data reported in the previous paragraph (not shown), indicated that the percentage of females who had high dispositions (interest) in STEM as a career (47%) was already somewhat lower than for males (53%), and when the added restriction of planning to have a career in STEM was added, the gap widened to only 42% of those with stated and measured intent to pursue STEM as a career being female, while 58% were male. This returns the discussion to the opening statement of this paper, that the ACT (2015) study had identified a gap among young people across the USA, between their measured interest in STEM as a career versus their stated intent to pursue STEM as a career. The findings of the current study concur with the findings of the ACT. The current study findings also identify that the gap is especially large for females, while also indicating that hands-on, real-world, make-the-world-a-better-place activities like those featured in the MSOSW project are especially effective in promoting stated and measured intent to pursue a career in STEM – whether or not the middle school students come to their MSOSW project classrooms with a plan to pursue a career in STEM – if the participants are female.

Conclusion

Attitudes formed during middle school have a large influence on students' academic performance (Liu, Horton, Olmanson, & Toprac, 2011), which in turn affects students' career aspirations (Choi & Chang, 2011). Therefore, understanding middle school students' perceptions regarding STEM dispositions is crucial to preparing our future STEM workforce as well as future citizens. The outcomes of the current study concur with the ACT findings that a gap exists among young people across the US regarding positive interest in STEM as a career versus stated intent to pursue a STEM career. The findings from the current study also provide evidence that progress can be made toward eliminating the existing gender gap in STEM career interest and intent, and, in particular that hands-on-science activities such as those embedded in the MSOSW project are particularly effective in enhancing STEM career interests among middle school girls. This is true for girls whether they begin project activities planning to purse a career in STEM, or not. Additional studies are needed to confirm these findings for different age groups of students and learning activities or school environments..

Acknowledgements

This research was funded in part by the NSF ITEST grant #1312168.

References

Amercian College Testing Service (ACT) (2015). *The Condition of STEM 2015.* Retrieved February 25, 2016 from http://www.act.org/stemcondition/15/.

Bisland, M., Kekelis, L., McCreedy, D., Koster, E., Gupta, P., & Roman, C. (2011). *Science Museum Roundtable*. Discussion at the NCWIT K-12 Alliance Summit, New York.

Bouvier, S. (2011). Increasing student interest in science, technology, engineering, and math (STEM): Massachusetts STEM pipeline fund programs using promising practices. Report prepared for the Massachusetts Department of Higher Education by the University of Massachusetts Donahue Institute. Retrieved from http://www.cssia.org/pdf/20000062-IncreasingStudentInterestinScience,Technology,Engineering,andMath(STEM).pdf

Bowdich, S. (2009). *Analysis of Research Exploring Culturally Responsive Curricula in Hawaii.* Paper presented to the Hawaii Educational Research Association Annual Conference, February 7, 2009.

Buldu, M. (2006). Young children's perceptions of scientists: A preliminary study. *Educational Research, 48*(1), 121-132.

Ceci, S.J., Williams W.M., & Barnett, S.M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin* 135(2), 218-261.

Christensen, R**.** & Knezek, G**.** (2015). Active learning approaches to integrating technology into middle school science classrooms: Reconceptualizing a middle school science curriculum based on 21st century skills. In X.Ge, D. Ifenthaler, & J.M. Spector (Eds.). *Full Steam Ahead: Emerging Technologies for STEAM*. New York: Springer Academic.

Christensen, R., Knezek, G., & Tyler-Wood, T. (2015). Gender differences in high school dispositions toward science, technology, engineering and mathematics careers. *Journal of Computers in Mathematics and Science Teaching, 34*(4), 395-408.

Christensen, R., Knezek, G., & Tyler-Wood, T. (2014). Student perceptions of Science, Technology, Engineering and Mathematics (STEM) content and careers. *Computers in Human Behavior, 34*, 173-186. doi:10.1016/j.chb.2014.01.046.

Choi, N., & Chang, M. (2011). Interplay among school climate, gender, attitude toward mathematics, and mathematics performance of middle school students. *Middle Grades Research Journal, 6*, 14.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

DeVellis, R.F. (1991). *Scale development*. Newbury Park, NJ: Sage Publications.

Dweck, C. (2006). Is math a gift? Beliefs that put females at risk. In *Why Aren't More Women in Science? Top Researchers Debate the Evidence*, ed. S. J. Ceci and W. M. Williams, 47-55. Washington, DC: American Psychological Association.

Gottfredson, L. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, *28*(6), 545-579.

Halverson, H.G. (2011). The trouble with bright girls. *Huffington Post*, March 1. http://www.huffingtonpost.com/heidi-grant-halvorson-phd/girls-confidence_b_828418.html (accessed March 1, 2011).

Liu, M., Horton, L., Olmanson, J., & Toprac, P. (2011). A study of learning and motivation in a new media enriched environment for middle school science. *Educational Technology Research and Development, 59*(2), 249-265.

Modi, K., Schoenberg, J., & Salmond, K. (2012). *Generation STEM: What girls say about science, technology, engineering and math.* The Girl Scout Research Institute.

Nguyen, H.H.H., & Ryan, A.M.M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314-34.

Palmer, D. H. (1997). Investigating students' private perceptions of scientists and their work. *Research in Science & Technological Education, 15*(2), 173-184.

Peterman, K., Kermish-Allen, R., Knezek, G., Christensen, R., & Tyler-Wood, T. (2016, accepted). Measuring student career interest within the context of technology-enhanced STEM projects: A cross-project comparison study based on the Career Interest Questionnaire. *Journal of Science Education and Technology 25*(4), DOI:10.1007/s10956-016-9617-5

Tai, R., Liu, C., Maltese, A., & Fan, X. (2006). Planning early for careers in science. *Science, 312*(5777), 1143-1144.

U.S. Census Bureau (2010). Census 2010 Summary File 1. Retrieved from http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf February 27, 2016.

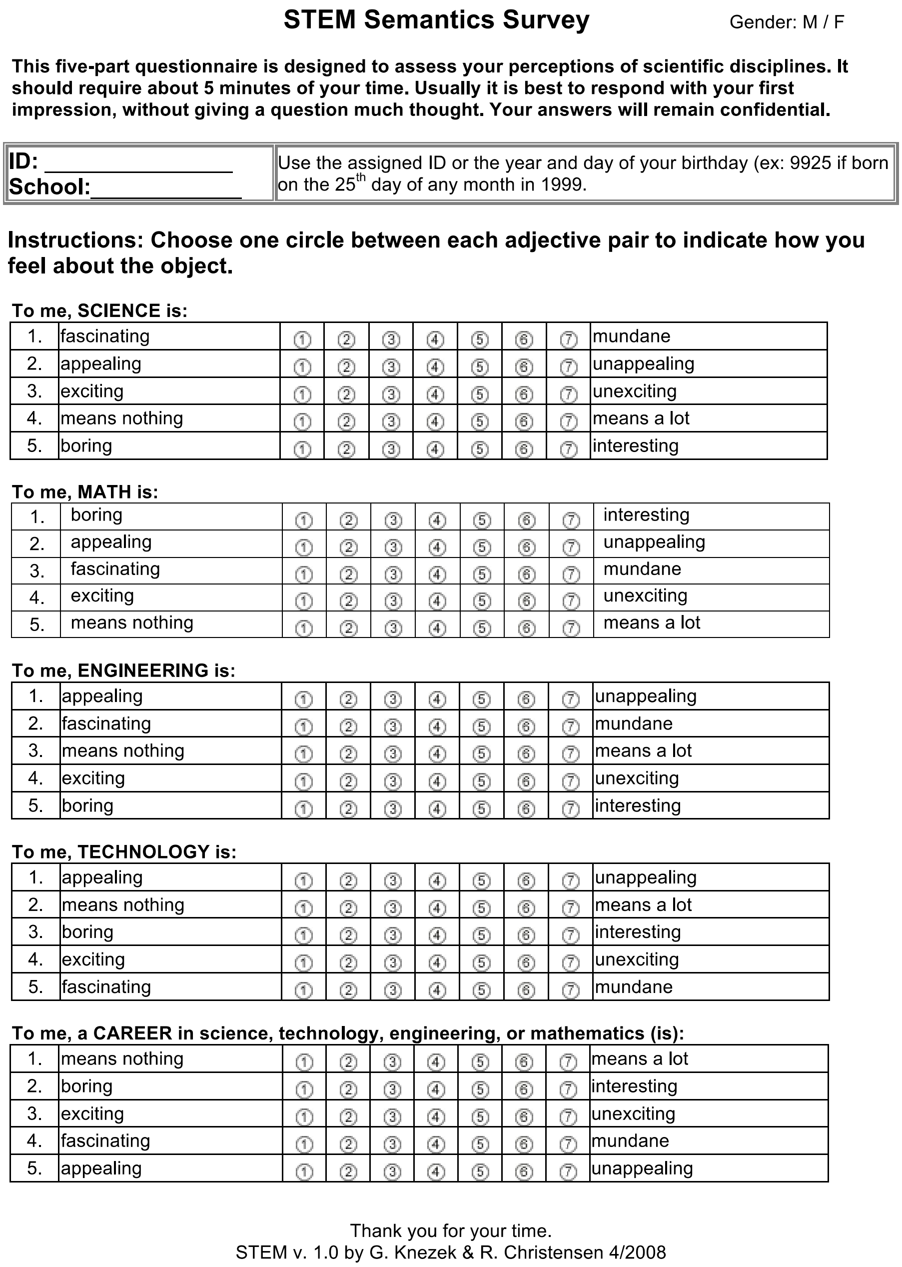
Walton, G.M., & Spencer, S.J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*(9), 1132-39.

Author Information

Rhonda Christensen
University of North Texas
Denton TX 76203, USA
Contact e-mail: *Rhonda.christensen@gmail.com*

Gerald Knezek
University of North Texas
Denton TX 76203, USA

Appendices



STEM Semantics Survey

Gender: M / F

This five-part questionnaire is designed to assess your perceptions of scientific disciplines. It should require about 5 minutes of your time. Usually it is best to respond with your first impression, without giving a question much thought. Your answers will remain confidential.

ID: _________________ **School:**_____________	Use the assigned ID or the year and day of your birthday (ex: 9925 if born on the 25th day of any month in 1999.

Instructions: Choose one circle between each adjective pair to indicate how you feel about the object.

To me, SCIENCE is:

1.	fascinating	①	②	③	④	⑤	⑥	⑦	mundane
2.	appealing	①	②	③	④	⑤	⑥	⑦	unappealing
3.	exciting	①	②	③	④	⑤	⑥	⑦	unexciting
4.	means nothing	①	②	③	④	⑤	⑥	⑦	means a lot
5.	boring	①	②	③	④	⑤	⑥	⑦	interesting

To me, MATH is:

1.	boring	①	②	③	④	⑤	⑥	⑦	interesting
2.	appealing	①	②	③	④	⑤	⑥	⑦	unappealing
3.	fascinating	①	②	③	④	⑤	⑥	⑦	mundane
4.	exciting	①	②	③	④	⑤	⑥	⑦	unexciting
5.	means nothing	①	②	③	④	⑤	⑥	⑦	means a lot

To me, ENGINEERING is:

1.	appealing	①	②	③	④	⑤	⑥	⑦	unappealing
2.	fascinating	①	②	③	④	⑤	⑥	⑦	mundane
3.	means nothing	①	②	③	④	⑤	⑥	⑦	means a lot
4.	exciting	①	②	③	④	⑤	⑥	⑦	unexciting
5.	boring	①	②	③	④	⑤	⑥	⑦	interesting

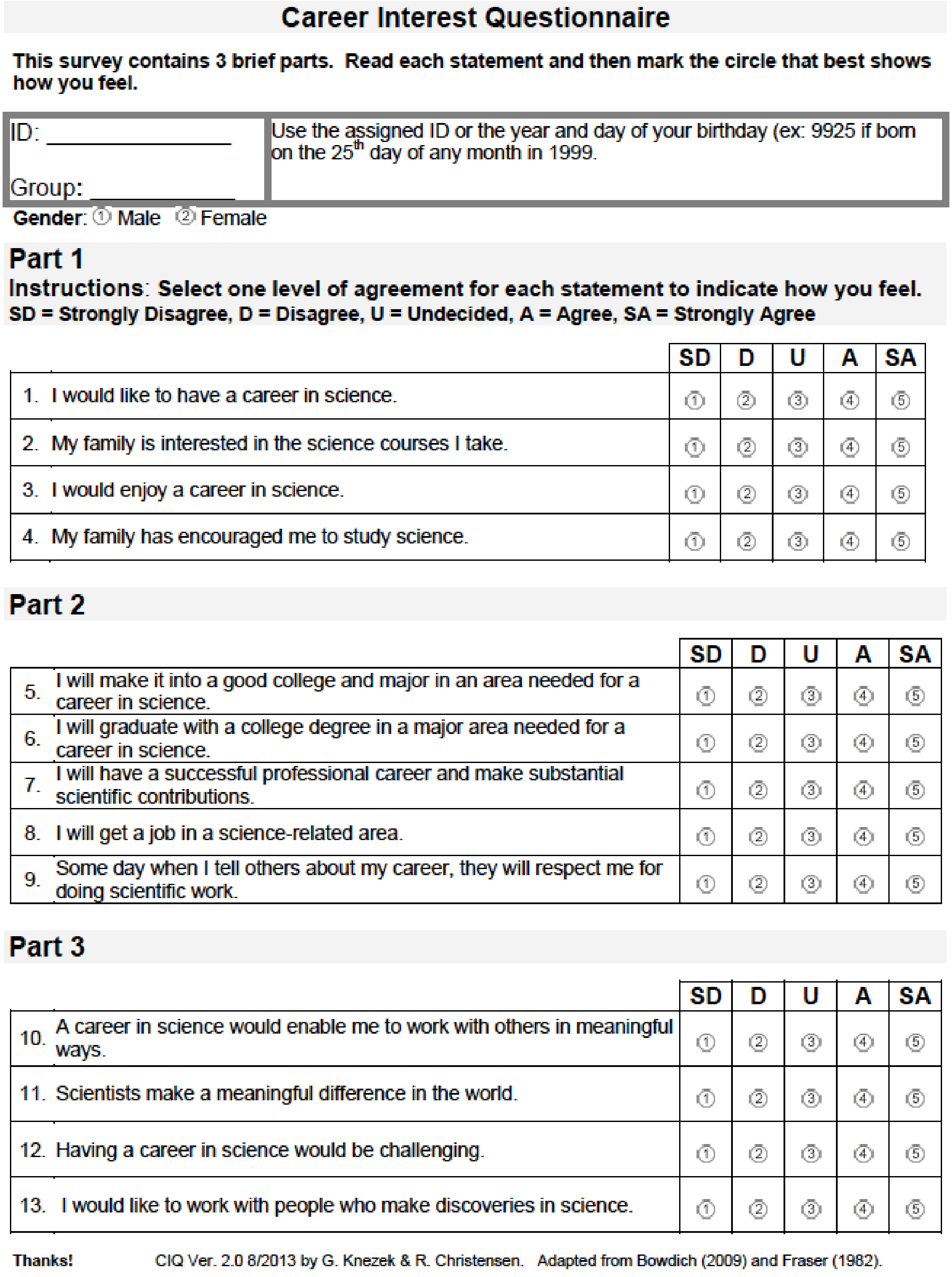
To me, TECHNOLOGY is:

1.	appealing	①	②	③	④	⑤	⑥	⑦	unappealing
2.	means nothing	①	②	③	④	⑤	⑥	⑦	means a lot
3.	boring	①	②	③	④	⑤	⑥	⑦	interesting
4.	exciting	①	②	③	④	⑤	⑥	⑦	unexciting
5.	fascinating	①	②	③	④	⑤	⑥	⑦	mundane

To me, a CAREER in science, technology, engineering, or mathematics (is):

1.	means nothing	①	②	③	④	⑤	⑥	⑦	means a lot
2.	boring	①	②	③	④	⑤	⑥	⑦	interesting
3.	exciting	①	②	③	④	⑤	⑥	⑦	unexciting
4.	fascinating	①	②	③	④	⑤	⑥	⑦	mundane
5.	appealing	①	②	③	④	⑤	⑥	⑦	unappealing

Thank you for your time.
STEM v. 1.0 by G. Knezek & R. Christensen 4/2008



Career Interest Questionnaire

This survey contains 3 brief parts. Read each statement and then mark the circle that best shows how you feel.

ID: ________________ Group: _____________	Use the assigned ID or the year and day of your birthday (ex: 9925 if born on the 25th day of any month in 1999.

Gender: ① Male ② Female

Part 1

Instructions: **Select one level of agreement for each statement to indicate how you feel.**
SD = Strongly Disagree, D = Disagree, U = Undecided, A = Agree, SA = Strongly Agree

		SD	D	U	A	SA
1.	I would like to have a career in science.	①	②	③	④	⑤
2.	My family is interested in the science courses I take.	①	②	③	④	⑤
3.	I would enjoy a career in science.	①	②	③	④	⑤
4.	My family has encouraged me to study science.	①	②	③	④	⑤

Part 2

		SD	D	U	A	SA
5.	I will make it into a good college and major in an area needed for a career in science.	①	②	③	④	⑤
6.	I will graduate with a college degree in a major area needed for a career in science.	①	②	③	④	⑤
7.	I will have a successful professional career and make substantial scientific contributions.	①	②	③	④	⑤
8.	I will get a job in a science-related area.	①	②	③	④	⑤
9.	Some day when I tell others about my career, they will respect me for doing scientific work.	①	②	③	④	⑤

Part 3

		SD	D	U	A	SA
10.	A career in science would enable me to work with others in meaningful ways.	①	②	③	④	⑤
11.	Scientists make a meaningful difference in the world.	①	②	③	④	⑤
12.	Having a career in science would be challenging.	①	②	③	④	⑤
13.	I would like to work with people who make discoveries in science.	①	②	③	④	⑤

Thanks! CIQ Ver. 2.0 8/2013 by G. Knezek & R. Christensen. Adapted from Bowdich (2009) and Fraser (1982).

JESEH
Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

Nature of Science as Portrayed in the Middle School Science and Technology Curriculum: The Case of Turkey

Kemal Izci

Article Info

Article History

Received:
23 June 2016

Accepted:
18 November 2016

Keywords

Nature of science
Science curricula
Textbooks

Abstract

Representation of nature of science (NOS) within curricula including standards, grade level expectations, and textbooks and their alignment with each other to achieve teaching NOS is crucial. Thus, the aims of the study were to a) assess how NOS is portrayed in standards, grade level expectations and a teacher edition of seventh grade textbook of science and technology curriculum of Turkey and demonstrate b) how they aligned with each other to support teaching of NOS. A conceptual analysis was used to achieve the aims of the study by focusing on aspects of NOS , namely: the empirical, tentative, inferential, creative, theory-laden, and social dimensions of NOS; myth of "The Scientific Method"; nature of theories and laws; and social and cultural embeddedness of science. Analyses indicated that the targeted NOS aspects were insufficiently portrayed in these documents in that some important aspects of NOS (e.g., scientific theories and laws) are not included while some others (e.g., inferential and theory-driven) are implicitly represented. The findings also showed that the alignment between the curriculum and the textbook is not adequate to transfer the aims of the curriculum for NOS into classroom practices. Implications for curriculum developers and textbook publishers were also discussed.

Introduction

Accomplishing practical scientific literacy requires people to have an understanding of scientific concepts and knowledge as well as an understanding of the enterprise of science and the nature and function of scientific knowledge (Roberts, 2007; Ryder, 2001). Nature of science (NOS) consists of the social and epistemological part of science and scientific knowledge, and focuses on how scientific knowledge is constructed and how it progresses (Khishfe & Lederman, 2006). Thus, an understanding of the NOS is regarded as a major part of accomplishing scientific literacy for all nations and globally advocated as a general aim of school science (Driver, Leach, Millar, & Scott, 1996; McComas, Clough, & Almazroa, 1998; Millar & Osborne, 1998; National Research Council [NRC], 1996, 2012). Turkey is one of these nations and Turkish science and technology curriculum aims to improve students' scientific and technological literacy by incorporating various aspects of NOS (National Ministry of Education [NME], 2006, 2013). Scientific and technological literacy is emphasized in the Turkish science and technology curriculum as an understanding of the development and nature of scientific knowledge. The interactions among science, technology, society, and environment and using this understanding to solve daily life problems are also highlighted as an important construct in the curriculum (NME, 2006, 2013). The role of reform documents and science education standards are emphasized as crucial in leading educational systems (NRC, 2012). Additionally, the role of teachers and science textbooks are vital to achieve a satisfactory level of scientific literacy (Chiappetta & Fillman, 2007; NRC, 1996). Turkey's National Science Education Standards (TNSES) and science and technology curriculum in the Turkish educational system intend to build on current research findings including research on NOS to improve students' learning and interest in science. Therefore, the science and technology curriculum is one of the main promoters of NOS for Turkey and needs to be aligned with educational research findings about NOS. Doing so may provide better NOS instruction that is research-informed and well-designed.

Additionally, research shows that 90% of school science teachers use science textbooks as a main instructional source for classroom teaching and assignments (Weiss, Banilower, McMahon, & Smith, 2001). Science textbooks include discussions on various aspects of NOS consisting of demonstration of laboratory works and the links among science, technology, and society to help teachers in teaching about aspects of NOS (Chiappetta & Fillman, 2007; McCommas, 1998). Hence, school science textbooks and science standards should reflect current educational research suggestions on NOS to provide better NOS instruction. A small amount of empirical research has assessed how NOS is portrayed in science education standards and textbooks, and to what extent the textbooks respond to the science education reform documents (Abd-El-Khalick et al., 2016; Abd-El-Khalick, Waters, & Le, 2008; Irez, 2009; McComes & Olson, 2002; Niaz, 2010; Vesterinen, Aksela, & Lavonen, 2013).

Research on analyzing science education standards and textbooks has mainly focused on the representation of scientific literacy and the accuracy and presentation of specific science content. There has also been limited research that focuses on how NOS is portrayed in middle school science textbooks. Furthermore, a few studies conducted in Turkey on middle school curricula (Erdogan & Köseoglu, 2012; Sardag et al., 2014; Ozden & Cavlazoglu, 2015) focused on only the science curriculum or science textbooks and did not investigate their alignment with each other to support teaching and learning of NOS. One of the most important documents leading a nation’s educational system is educational standards. These educational standards determine goals of teachers for their teaching and textbooks, and support teachers to reach these goals in each grade level (Koulaidis & Tsatsaroni, 1996). In terms of science education, science education standards lead what objectives should be targeted and how they need to be taught. Therefore, a nation’s science education standards and science curriculum and textbooks are the main documents that show how this nation considers NOS in its educational system as a goal and supports its teachers to reach this goal. Thus, this study aims to investigate how Turkish science curriculum (NME, 2006) and Turkish middle school science textbooks (NME, 2011) align with each other and to what extent they contain current research suggestions to aid instruction and understanding of NOS.

Review of Literature

Enhancing teachers’ and students’ understanding of NOS to accomplish scientific literacy has been a major part of science education research efforts. These efforts include designing courses and curricula to help teachers and students to improve their views of NOS, assessing students’ and teachers’ understanding of NOS, and providing educational materials and practices to promote transformation of NOS into classroom practices (Abd-ElKhalick & Lederman, 2000; Akerson & Hanuscin, 2007; Khishfe & Abd-El-Khalick, 2002). Although there has been progress to advance students’ and teachers’ understanding of NOS, recent research shows that teachers and students still hold naïve views about some particular aspect of NOS such as the tentative nature of scientific knowledge (Abd-El-Khalick, 2005; Abd-El-Khalick & Akerson, 2004; Bora, Aslan, & Cakiroglu, 2006). Many obstacles such as limited implementation of NOS in pre-service and in-service teacher education programs, the inconsistency between reform documents and textbooks regarding their NOS perspectives, the difficulty of integrating NOS into curriculum, and representation of NOS in science textbooks have been found to be some of the common reasons behind the limited understanding of NOS (Abd-El-Khalick et al., 2016; Driver et al., 1996; McComas, et al., 1998; Rudolph, 2002).

Teaching Nature of Science

Studies in the literature focus on various ways of teaching NOS to promote scientific literacy. The recent research on teaching NOS has shown that the most effective way of teaching NOS is use of an explicit-reflective approach of instruction (Akerson & Hanuscin, 2007; Koksal & Cakiroglu, 2010). In contrast, some researchers think that implicit teaching approaches that address NOS by engaging students in hands-on inquiry and laboratory activities would improve learners’ understanding of NOS. However, research results have shown that the implicit approach for teaching NOS is not as effective as explicit approaches (Abd-El-Khalick & Akerson, 2004; Khishfe & Abd-El-Khalick, 2002). Explicit teaching of NOS refers to intentionally focusing on some aspects of NOS in teaching and designing lessons by incorporating specific examples, activities, questions, and reflection times to facilitate learners’ understanding of intended aspects of NOS (Clough, 2006). Researchers have illustrated that explicit approaches provide activities about aspects of NOS and encourage learners to reflect on NOS activities to help learners develop a better understanding of NOS (Akerson & Hanuscin, 2007; Khishfe & Abd-El-Khalick, 2002; McComas, 2003).

Researchers have also noted the role of explicit integrated versus nonintegrated NOS teaching approaches. In the explicit integrated teaching approach, teachers choose appropriate aspects of NOS and explicitly address them during instruction of science content. On the other hand, in explicit nonintegrated instruction approach, teachers use specific NOS activities such as card sorting, black-box and puzzle solving activities to address specific aspects of NOS without embedding it into science content. Even though research has shown a slight difference between explicit integrated and nonintegrated instruction of NOS on students’ learning (Khishfe, 2008; Khishfe & Lederman, 2006), many researchers have preferred to support the explicit integrated approach for teaching NOS (Akerson & Hanuscin, 2007; Clough, 2006; Matkins & Bell, 2007). Researchers appreciate the contribution of the explicit nonintegrated teaching approach to students' understanding of NOS aspects, but they claim that the explicit integrated approach is more likely to give the sense of doing science to students, improve science content knowledge, and reduce the time anxiety of teachers (Clough, 2006; Clough & Olson, 2001). It is clear that many researchers have indicated explicit NOS instruction is the most effective way of teaching to improve students' understanding of NOS. However, how teachers can be supported to adopt explicit teaching approach to improve students' learning of NOS still needs to be addressed.

Analysis of Materials for Representations of Nature of Science

Currently, there is a lack of research that particularly analyzes the representation of NOS in middle school science textbooks and their alignment with science education reforms (Ozden & Cavlazoglu, 2015). A few studies analyzed middle school science textbooks, but they focused on the content, structure, curriculum, and major concepts rather than focusing on NOS (Chiappetta, Sethna, & Fillman, 1993; Stern & Roseman, 2004). Chiappetta, Sethna, and Fillman (1993) analyzed five middle school life science textbooks in the United States to document how these textbooks provide a balance of scientific literacy themes, which requires equally highlighting scientific, technological, and societal sides of scientific knowledge and its development. They reported that these textbooks highlighted the teaching of content and did not provide a balanced view of scientific literacy. Abd-El-Khalick (2002) analyzed middle school-level science trade books published in the United States to see their representation of the NOS images. He used some aspects of NOS such as empiricism, imagination, and creativity and found that these science trade books did not explicitly represent any aspects of NOS.

Some other studies analyzed diverse aspects of NOS in precollege science textbooks, but these studies mainly focused on high school science textbooks. However, research suggests that the aspects of NOS should be incorporated into all curricula from elementary to college level education to provide a successful continuum of NOS education (Abd-El-Khalick, 2011; Abd-El-Khalick et al., 2016). For example, McComas (2003) analyzed 15 high school biology textbooks in the United States to evaluate the accuracy of their inclusion of scientific laws and theories. This study reported that none of the 15 high school biology textbooks provided an accurate view of scientific laws and theories. The study documented that although textbooks used the terms of law and theory, they did not provide a clear definition of these terms. In another study, Chiappetta and Fillman (2007) analyzed five recent high school biology textbooks used in the United States for the inclusion of NOS. In their analysis, they used four different NOS themes: science as a body of knowledge, science as a way of investigating, science as a way of thinking, and science and its interactions with technology and society. They reported that the textbooks incorporated and used the four aspects of NOS more than the textbooks they analyzed 15 years ago and these textbooks more likely considered the science education reform documents. Similarly, Irez (2009) analyzed five tenth grade high school biology textbooks in Turkey by using the cognitive map data analysis method. His analyses showed that three of the five textbooks defined science as a body of accumulated knowledge while objectivity in science was naively presented in all these textbooks. Furthermore, he reported that all analyzed textbooks provided naïve views and unacceptable definitions about scientific theories and laws, presenting a hierarchy between the theory and law. Abd-El-Khalick and his colleagues (2008) highlighted the importance of including NOS in high school chemistry textbooks and analyzed 14 high school chemistry textbooks in the United States including five 'series' spanning one to four decades. In their analysis, they considered aspects of NOS: science as empirical, tentative, inferential, creative, and theory-driven, along with the myth of the scientific method, the nature of scientific theories and laws, and the social and cultural embeddedness of science. They reported that the analyzed 14 chemistry textbooks poorly portrayed the aspects of NOS. They also reported that representation of these aspects of NOS in textbooks stayed constant or decreased over the past four decades and concluded "These trends are incommensurate with the discourse in national and international science education reform documents …" (p. 1). Vesterinen et al. (2013) analyzed five high school chemistry textbooks used in Finnish and Swedish school system for their representations of NOS. During their analyses, they focused on knowledge of science, investigating nature of science, science as a way of thinking, and interaction of science, technology and society as their themes for analyzing the textbooks. Their findings showed that the textbooks had little emphasis on science as a way of thinking. Also, in terms of tentativeness dimension of NOS, Swedish textbooks were found to be more successful than Finnish textbooks. Recently, Abd-El-Khalick et al. (2016) analyzed 34 textbooks, consisting of 16 biology and 18 physics books, published in the United States. Based on their purpose, they chose textbooks that had at least five series and had been published for at least 3 decades. Their findings showed that (a) the content domain did not make any difference in terms of representation of NOS; (b) the textbooks, over the years, did not significantly improve in representation of NOS; (c) the emphasis of NOS did not align with how NOS was highlighted within reform documents. They also indicated that rather than the reform efforts, the authors of the textbook had a greater role in terms of representing aspects of NOS.

There is also a lack of research analyzing the representation of NOS in different reform documents and science education standards to show how they align with recent research findings on NOS. McComes and Olson (1998) analyzed how different countries and states directed science education standards of NOS in their K-12 learning environments. In their study, McComes and Olson (1998) analyzed the following reform documents in the United States for their inclusion of NOS: Benchmarks for Science Literacy (American Association for the Advancement of Science [AAAS], 1993); Science Framework for California Public Schools (California Department of Education, 1990); National Science Education Standards (NRC, 1996); The Liberal Art of Science (AAAS, 1990). They also looked at reform documents such as: A Statement on Science (Curriculum Corporation, 1994) in Australia; Science in the National Curriculum (Department of Education, 1995) in England/Wales; Science in the New Zealand Curriculum (Ministry of Education, 1993) in New Zealand; and Common Framework (Council of Ministers of Education, 1996) in Canada to analyze how these reform documents from five different countries portrayed NOS. The results of their analysis showed that all documents emphasized the role of NOS in science education standards. However, these documents did not include "the notion of paradigm, the objectivity aspect of science and the idea that science has inherent limitations," and their importance on student understanding of scientific enterprise (McComes & Olson, 1998, p. 49). The results also showed that these documents generally failed to provide definitions of terms such as scientific law and theory. Erduran and Dagher (2014) analyzed the draft of Irish middle school curriculum in terms of its depiction of NOS. While they found most of the aspects of NOS were represented within the curriculum, they suggested that some of the other aspects, such as a nuanced view of NOS, should be more developed and integrated within the curriculum.

Additionally there are a few studies conducted on Turkish science curriculum (e.g., Erdogan & Köseoglu, 2012; Sardag et al., 2014; Ozden & Cavlazoglu, 2015). Erdogan and Koseoglu (2012) analyzed some curricula documents published in 2008 including TNSES and grade level expectation (GLEs) for high school biology, physics and chemistry. Their findings showed that while science as accumulation of knowledge aspect was highlighted in the chemistry curriculum, the aspect of science as inquiry was mostly emphasized in biology and physics curricula. However, their findings also showed that none of the three curricula emphasized science as a way of thinking. Similarly, Sardag et al. (2014) analyzed the GLEs in high school biology, physics and chemistry curricula published in 2013 by Turkey's NME. Their findings also showed that the representations of aspects of NOS were not adequate in these documents; moreover, none of these documents represented imagination and creativity in science aspects. Lastly, Ozden and Cavlazoglu (2015) analyzed middle school science curricula published in 2005 and 2013 to compare their inclusions of NOS aspects and approaches to teaching NOS. Their findings showed that both 2005 and 2013 middle school science curricula did not support explicit approaches of NOS. They also found that although the experimentation in science, scientific method, and socio-cultural embeddedness of science aspects were included within the standards of each curricula, they were not included within the GLEs. Interestingly, they indicated that the 2005 curriculum was more successful than 2013 curriculum in terms of NOS because the standards in 2005 curriculum explicitly provided detailed knowledge about NOS while the 2013 curriculum did not provide any knowledge within its standards. All the three studies (i.e., Erdogan & Köseoglu, 2012; Sardag et al., 2014; Ozden & Cavlazoglu, 2015) merely utilized the TNSES and GLEs within the science curricula. However, textbooks are still the main sources that teachers use to shape their instruction (Weiss et al., 2001). Thus, a more comprehensive analysis including the TNSES, the GLEs, and science textbooks need to be taken into account to illustrate how the Turkish science curriculum and science textbooks support teachers and students in teaching and learning NOS.

Conceptual Framework

Marsh and Willis (2007) mention three levels of curriculum, which are planned, enacted, and experienced curriculum. *Planned curriculum* refers to the most valued knowledge, important learning goals, and guidelines that are addressed in the standards and GLEs. *Enacted curriculum* includes the professional and pedagogical knowledge that teachers need to use to implement and evaluate a curriculum. *Experienced curriculum* is the most important one and consists of real learning environments and actual classrooms in which teachers and students interact and produce learning. McNeil (2003) highlights the experienced curriculum as a live curriculum and states that it is meaningful when teachers and students engage in classroom activities. To accomplish a curricular goal such as improving NOS instruction, these three types of curriculum should be aligned (Marsh & Willis, 2007). The alignment can be accomplished by professional development and textbooks (Bakah, Voogt, & Pieters, 2012). Therefore, to accomplish effective NOS instruction, first, appropriate professional development activities (in the enacted curriculum) should be offered. Second, textbooks should provide appropriate learning activities and guidelines that explicitly address NOS aspects and support teachers' explicit NOS instruction (Akerson & Hanuscin, 2007). Otherwise, one cannot expect teachers to use explicit instruction of NOS by simply emphasizing explicit NOS instruction in the standards (planned curriculum) while textbooks (experienced curriculum) do not include and discuss NOS explicitly (Akerson et al., 2010).

Conceptual frameworks are important as they guide and provide maps to decide how to analyze any form of information (Marsh & Willis, 2007). Researchers use various conceptual frameworks to analyze textbooks to test their usage and representation of some specific content, concept, image, and their epistemological orientation to learning and teaching (Chiappetta & Fillman, 2007). In conceptual analysis, a specific concept or content is chosen for assessment, and the conceptual analysis process looks for quantification and the presence of the specific concept or content. The conceptual framework helps researchers identify and limit the examined concepts to reach a meaningful result in a practical and structured way to facilitate document analysis. We explain the identified and examined aspects below.

Aspects of NOS

In the literature, researchers use various conceptual frameworks to analyze the existence of NOS in science textbooks. While there is not an agreed upon definition of NOS, "there is a strong consensus about characteristics of the scientific enterprise that should be understood by an educated citizen" (NRC, 2012, p. 78). Based on this consensus, Abd-El-Khalick et al. (2008) employed ten aspects of NOS which researchers and various reform documents agreed upon to analyze 14 high school chemistry textbooks for their depiction of NOS. These aspects of NOS, as mentioned earlier, include empirical, inferential, creative, theory-driven, and tentative aspects of science, the myth of the scientific method, scientific theories, scientific laws, social dimension of science, and social and cultural embeddedness of science. These aspects of NOS are also supported by various reform documents (AAAS, 1993; NRC, 1996, 2012) and other national and international researchers (Abd-El-Khalick & Lederman, 2000; McComas, 1998; McComas & Olson, 1998; Osborne et al., 2003). This study also employs the ten aspects of NOS to analyze the science and technology curriculum and related textbooks.

Method

In this study, we employed the purposeful sampling method (Meriam, 2009) to choose specific documents to analyze how NOS is portrayed in these documents. We used 2006 middle school science and technology curriculum for grade 6th, 7th and 8th including TNSES, GLEs, and the teacher edition of the middle school science and technology textbook for the 7th grade. All selected documents and the textbook were published by Turkey's NME. We purposefully chose NME's science and technology curriculum published in 2006 and the 7th grade teacher guidebook published in 2011. Although NME published a new science curriculum in 2013, the curriculum did not explicitly address NOS in its standards (Ozden & Cavlazoglu, 2015). In addition, a teacher's guidebook has not been published yet for the 2013 curriculum to guide teachers about how to address NOS in their teaching. Thus, we purposefully chose to analyze the 2006 science curriculum and the 7th grade teacher's guidebook. Our analyses focused on the curriculum for grade 6, 7, and 8, which form the middle school years within the school system in Turkey. We conducted a conceptual analysis method by considering the aspects of NOS provided by Abd-El-Khalick et al. (2008). As indicated in Abd-El-Khalick et al.'s (2008) conceptual analysis method, we utilized explicit versus implicit and integrated versus nonintegrated teaching approaches of NOS as a guiding framework. Overarching research question guiding our study was; How is NOS portrayed in the Turkish middle school science and technology curriculum and the 7th grade teacher's edition science and technology textbook? More specific questions guiding this study were;

a) What aspects of NOS are portrayed in the middle school science and technology curriculum, and the 7th grade teacher's edition science and technology textbook?
b) To what extent does the 7th grade teacher's edition science and technology textbook align with the science and technology curriculum in terms of representing NOS?
c) To what extent does the representation of NOS in the science and technology curriculum and the 7th grade teacher's edition science and technology textbook reflect research findings on NOS to support science teachers' adoption of NOS?

Data Sources

In 2006, NME in Turkey changed and revised middle school science curriculum by incorporating new scientific and technological findings and considering international science education reform documents and standards. NME named the program *Science and Technology* and modified the curriculum and textbooks to achieve nationwide implementation of this program. Turkey has a national curriculum and all textbooks are published or approved by NME and distributed to teachers and students for free. Recently, NME provided an option for teachers to use textbooks from private publishers if NME approved the publishers' textbooks, but all textbooks still need to be aligned with the national curriculum. Thus, all middle school science teachers are expected to use the national science and technology curriculum and textbooks in their classrooms. The purpose of this paper was to analyze how NOS is portrayed in 2006 science and technology curriculum including textbook for the 7th grade. Therefore, the science and technology curriculum including TNSES (NME, 2006), GLEs (NME, 2006), and the 7th grade teacher's edition textbook (NME, 2011) constituted the primary data sources for this study. We did not include students' edition of the textbook because the teacher's edition incorporates the students' edition in it with additional information about curriculum expectations and instructional supports.

Selection of Sections for Analysis

The published 2006 science and technology curriculum includes two parts. The first part contains the TNSES for this program and the second part consists of GLEs. For the first part, the essentials of TNSES and the organizational sections of the TNSES, namely, (a) the role of knowledge, (b) science, technology, society, and environment, (c) science process skills, and (d) attitudes and values were selected for analyzing aspects of NOS. On the other hand, GLEs include seven sections and were analyzed for the purpose of the study. Three of the general expectations include science process skills; science, technology, society, and environment; and attitudes and values embedded in content units. The other four sections consist of content unit expectations and consist of 'physical phenomena,' 'matter and change,' 'living beings and life,' and 'earth and universe' content units.

We chose to analyze the 7th grade teacher edition of the science and technology textbook because it includes the same general instructional suggestions for sixth and eighth grade textbooks for teachers. Moreover, the textbook contains the content units related with astronomy, biology, chemistry, environment, and physics; which we thought presented a more general representation of science domains. In the teacher's edition of the 7th grade science and technology textbook, various suggestions were provided to facilitate teachers' implementation. One of the sections within general suggestions focuses on NOS and science, technology, society, and environment relations which we thought appropriate to analyze for the aim of this study. Furthermore, in each content unit, there are two sections namely "Starting the unit" and "Where is it in our lives?" intended for target gains about scientific process skills and science, technology, society, and environment. We also included these two sections in our analysis to determine the representation of NOS in science content units.

Analysis

A conceptual analysis method using phrases as units of analysis was employed to determine how TNSES and GLEs for middle school science and technology curriculum, and the 7th grade teacher's edition of science and technology textbook portrayed NOS. Two different stages of analysis were carried out. Firstly, we identified the emphasized aspects of NOS by using the ten aspects of NOS, which we adopted from Abd-El-Khalick et al. (2008) and described it in Table 1. Then, we analyzed these documents for which teaching approaches containing explicit versus implicit and integrated versus nonintegrated instruction of NOS these documents chose to use (Akerson & Hanuscin, 2007; Clough, 2006; Matkins & Bell, 2007). Furthermore, we used a scoring rubric which includes four different classifications (see Table 2). If an analyzed document did not present a specific aspect of NOS, we classified it as not represented. If the document inaccurately represented a specific aspect or implicitly causes a misunderstanding of a specific aspect of NOS, we classified it as misrepresented. If the document implicitly presented or did not provide any explanation about the function of a specific aspect of NOS throughout its discussions, we classified it as implicitly represented. If the document explicitly presented a specific aspect of NOS and explained its role and function in the development process of scientific knowledge, we classified it as explicitly represented.

In the second phase of our analysis, we compared how the 7th grade teacher's editions of the science and technology textbook align, in terms of depiction of NOS, with TNSES and GLEs to encourage teachers to integrate NOS into their teaching. We compared promoted aspects of NOS and teaching approaches in TNSES and GLEs (theoretical realm) with the 7th grade teacher's edition textbook (practical realm) to show how science and technology teachers were encouraged to teach NOS in their classrooms.

All the coding and comparison schemas were performed independently by two researchers to provide interrater reliability. Both researchers independently analyzed and scored a sample of selected documents and textbook materials by considering the ten aspects of NOS and the teaching approaches promoted in these materials to teach NOS. Once independent coding was finalized, the coders compare their codes for coding reliability. Inter-rater reliability coefficient was calculated as % 91 (Miles & Huberman, 1994). Confusions in the coding process were resolved through discussion and negotiation to come up with a consensus.

Table 1. The NOS aspects targeted in the analysis of the curriculum

NOS Aspects	Explanations of NOS Aspects
Empirical	Scientific claims are derived from observations of natural phenomena. However, these observations are almost limited by human limitations, such as having limited or no access to direct observations. Prior assumptions also impact these observations.
Inferential	There is a certain distinction between observations and inferences. While observations are descriptions of natural phenomena that are accessible to our senses, inferences are statements about natural phenomena that are not directly accessible to our senses and require scientists to consider cause-and-effect relationships to produce these statements.
Creative	Creativity is an essential part of NOS because science is not always a systemic and rational activity, and it requires scientists to use their creativity to make inferences about natural phenomena.
Theory-driven	Scientists' prior knowledge, theoretical and disciplinary promises, and training influence their observations and interpretations of natural phenomena. Furthermore, these effects impact scientists' selections of problems, observational and interpretational methods, and investigative styles.
Tentative	Scientific knowledge is subject to change while it is reliable and durable because we cannot be certain about any type of scientific knowledge. Scientific claims are changed when new findings and advances are available in the scientific and technological world.
Myth of the scientific method	There is not a specific scientific method that is used by all scientists to produce scientific knowledge. Scientists use scientific process skills such as observation, interpretation, and hypothesis but there is not a certain stepwise way to guarantee the accuracy of produced knowledge.
Scientific theories	Theories are well established and consistent systems of explanations for natural phenomena. Theories use assumptions, axioms and indirect evidence to explain existence and behavior of non-observable things.
Scientific laws	Laws are descriptive statements about observable natural phenomena. Theories and laws are different kinds of knowledge in which theories focus on an explanation of non-observable entities while laws focus on a description of observable phenomena. Thus, there is not a hierarchical relationship between theory and law and theories cannot become laws when enough supporting evidence is found.
Social dimensions of science	Scientific knowledge is socially constructed and includes communication and criticism to enhance its objectivity. Communication plays a critical role in the development of scientific knowledge.
Social and cultural embeddedness of science	Science is a human endeavor and it develops in a cultural context. Science also affects and is affected by cultural variables such as religious and political and economical factors.

Note. Adopted from Abd-El-Khalick et al. (2008).

Findings

Our analyses of the middle school science curriculum including TNSES and GLEs, and teacher edition of the 7th grade science and technology textbook show that while most of the targeted NOS aspects were portrayed in the TNSES and GLEs, the textbook did not represent them or implicitly represent a few of them (see Table 2). The results of our analyses also demonstrate that the alignment among the TNSES, GLEs and the textbook is not in a desired level to support teaching and learning of targeted NOS aspects (see Table 2). On the other hand, our analysis of the textbook illustrates that the textbook was not successful in embedding the addressed aspects of NOS into content units as it implicitly addressed a few of the NOS aspects and did not integrate most of them into the content units (see Table 3). Besides, our analyses show that the curriculum uses implicit and non-reflective approach of NOS teaching and aligning with that the textbook also supports this approach by implicitly integrating a few of the aspects within content units without providing special activates for teachers to support NOS instruction (see Table 3 and Table 4).

Our first research question focused on what aspects of NOS were portrayed in the middle school science and technology curriculum and the 7th grade teacher's edition science and technology textbook. Table 2 shows that five aspects of NOS including empirical, creative and tentative aspects of science, social dimensions of science, and social and cultural embeddedness of science are explicitly stated in TNSES. We found that two aspects of NOS including inferential and theory-driven aspects were implicitly presented in TNSES. We also found that one aspect of NOS, which is the myth of the scientific method, was misrepresented. Furthermore, explanations of scientific theories and laws, two important aspects of NOS, were not represented in TNSES.

Furthermore, our analysis of the GLEs showed (see Table 2) five crucial aspects of NOS (i.e., creativity, theory-driven, myth of the scientific method, scientific theories, and scientific laws) were not represented in the GLEs. While two aspects of NOS including empirical and inferential were implicitly stated, three aspects of NOS consisting of the tentativeness of science, social dimensions of science, and the social and cultural embeddedness of science were explicitly represented in the GLEs.

Table 2. Representations and alignments of the aspects of NOS in the TNSES, GLEs, and the textbook

Aspects of NOS	TNSES	GLEs	Textbook
Empirical	✓	✓	○
Inferential	○	○	○
Creative	✓	×	✓
Theory-driven	○	×	✓
Tentative	✓	✓	○
Myth of the scientific method	●	×	●
Scientific theories	×	×	○
Scientific laws	×	×	●
Social dimensions of science	✓	✓	✓
Social and cultural embeddedness of science	✓	✓	○

✓ =explicitly presented ○ =implicitly presented ● =misrepresented × =not represented

Table 3. Representation of NOS in the 7th grade textbook

Aspects of NOS	Information for teachers	Unit-1	Unit-2	Unit-3	Unit-4	Unit-5	Unit-6	Unit-7
Empirical	○							○
Inferential	○							
Creative	✓							
Theory-driven	○							
Tentative	✓				○			○
Myth of the scientific method	●							
Scientific theories	○							
Scientific laws	●							
Social dimensions of science	✓							
Social and cultural embeddedness of science								○

✓ =explicitly presented ○ =implicitly presented ● =misrepresented × =not represented

The findings of our analysis of the 7th grade teacher's edition of the science and technology textbook show that, as seen in Table 2, three aspects of NOS (creative, theory-driven, and social dimension of science) are explicitly represented to promote teachers' instruction of NOS. Five aspects of NOS (empirical, inferential, tentative, scientific theories, and social and cultural embeddedness of science) are implicitly represented while two central aspects of NOS (the myth of the scientific method and scientific laws) are misrepresented in this textbook. Even though the ten aspects of NOS are represented in the 7th grade teacher edition of the science and technology textbook, our analysis showed that, as seen in Table 3, nine aspects of NOS are presented in the first part of this textbook that includes suggestions for teachers and just three aspects of NOS (empirical, tentative, and social and cultural embeddedness of science) are implicitly integrated into content units to promote teaching of NOS. If an aspect of NOS was explicitly provided within the information for teachers and this aspect of NOS was not provided or implicitly provided in any of the content units, we took into consideration the representation in content units because teachers employed these units in their instruction.

Our second research question was concerned with the alignment between the science and technology curriculum and the science and technology textbook in terms of depiction of NOS. The alignment of the aspects of NOS among the TNSES, GLEs, and science and technology textbook are shown in Table 2. GLEs are built based on the TNSES and science and technology textbooks are constructed on both TNSES and GLEs. However, our analysis of these documents shows that there are many inconsistences among these three documents in terms of representation of the aspects of NOS. For instance, TNSES emphasizes creativity, theory-driven scientific knowledge, and the myth of the scientific method aspects of NOS. While GLEs do not include these aspects of NOS, the teacher's edition of the textbook does.

Furthermore, although empirical, tentative, and social and cultural embeddedness of science aspects of NOS are explicitly represented in TNSES and GLEs, the science and technology textbook only implicitly represents these aspects of NOS. The following quote shows how the social and cultural embeddedness of science aspect of NOS is implicitly represented in the textbook: "The text aims to help students to conceptualize how the notion of element was constructed, the historical development of element concept, and how social occasions influence scientific development during the historical time" (NME, 2011, p. 133). Moreover, the following quote shows how the tentativeness aspect of NOS is implicitly portrayed in the textbook: "Currently, ignoring to use old atom models does not mean the scientists, developers of these models, did not think critically, but it indicates the knowledge in that time was very less than we have now" (NME, 2011, p. 154). Moreover, the textbook misrepresents the two aspects of NOS (scientific laws and the myth of the scientific method) while GLEs do not include these two aspects and TNSES includes one of them. For example, the following quote illustrates how the textbook misrepresents the relationship between theory and law: "If a theory, after a long process, is universally accepted and becomes a scientific fact without getting any criticisms, it will turn into a law" (NME, 2011, p. 9).

In our analysis of the science and technology textbook, we focused on two specific parts in each unit that are "Starting the unit" and "Where is it in our lives?" because these two parts are intended by the textbook authors to take teachers' attention to the TNSES and GLEs for each specific content unit. Our analyses of the curriculum and the textbook show that while the curriculum aims to embed the aspects of NOS, the textbook does not sufficiently embed the aspects of NOS into specific content units to support teachers in promoting students' understanding of NOS. Moreover, even if the suggestions for teachers in the first part of the textbook consist of the nine aspects of NOS as seen in Table 3, just three aspects of NOS (empirical, tentative, and social and cultural embeddedness of science) are implicitly embedded into content units throughout the textbook to guide teachers to teach these aspects of NOS. The findings show that the aim of the science and technology curriculum, which is to achieve the incorporation of NOS aspects into content units, is not reflected by and align with the 7th grade teacher's edition textbook.

Our third research question focuses on how the representation of NOS in the science and technology curriculum and the textbook reflect the research findings on NOS to support teachers' adoption of NOS. Our analysis of the curriculum shows that the science and technology curriculum promotes implicit and non-reflective instruction of NOS which is not an effective instructional approach. Aligning with the science and technology curriculum, the textbook also uses implicit and non-reflective approaches for teaching NOS by incorporating a few aspects of NOS into science content units without providing any specific activity or reflective questions for students to emphasize a specific aspect of NOS during instruction.

Furthermore, researchers have also discussed the role of explicit integrated and nonintegrated teaching approaches of NOS. Our analysis of the science and technology curriculum and the textbook shows that none of these documents promote explicit separated teaching approaches to NOS and do not provide any specific examples for teachers to integrate them into their instruction. Moreover, while the science and technology curriculum promotes the implicit integration of NOS aspects into science content units, the textbook does not integrate most aspects of NOS into science content units. Even if the textbook incorporates a few standards related with NOS into its "Starting the unit" part in each unit, our analysis shows that these intended standards are not explicitly addressed in any content activities or instruction by incorporating specific activities or structured questions to facilitate teachers' instruction and learners' understanding of NOS.

Table 4. Representative quotes taken from the TNSES, GLEs, and the textbook

NOS aspects	Representative quotes
✓ Empirical	Science depends on explanation of data driven from observations and experiments. Therefore, the explanations which are not driven from experimental evidences, observations, and scientific theories are not part of science (NME, 2011, p. 9-2, textbook).
○ Inferential	Science is also a way of learning that includes curiosity, imagination, intuition, creativity, investigation, observation, experiment, interpretation, and discussion on data and interpretations (NME, 2006, p. 61, TNSES).
✓ Creative	Scientists use their creativities in their research when they investigate phenomena. Usually, beliefs, curiosities, intuition, and imagination of scientists guide them to investigate phenomena. (NME, 2011, p.9-3, textbook)
○ Theory-driven	Therefore, the explanations which are not driven from experimental evidences and scientific theories are not part of science (NME, 2011, p. 9-2, textbook).
✓ Tentative	Scientific theories are always examined and when different evidences are available, these theories are modified and expanded to explain new and old information (NME, 2006, p. 61, TNSES) By using examples, explain the limited and changeable knowledge about the universe because of the vast space of the universe. (GLEs for the Solar System and Beyond unit-4)
● Myth of the scientific method	Therefore, science and technology curriculum does not aim to transfer the already accumulated knowledge to students but it aims to raise individuals who can investigate, question, search, and relate science concepts to his/her daily life, use *the scientific method* to solve daily life issues, and see the world trough scientists' view. This curriculum uses scientific process skills as an essential part to instruct the way and method of *the scientific method*. (NME, 2006, p. 61, TNSES)
○ Scientific theories	A hypothesis and ideas related to are tested by experiments. If the experiments support the hypothesis, the validity and reliability of the hypothesis increase. If other hypotheses also support this hypothesis, this hypothesis becomes theory (NME, 2011, p.9-4, textbook).
● Scientific laws	If a theory, after a long process, becomes a universal and a scientific fact without getting any critics, it turns into a law (NME, 2011, p.9-4, textbook).
✓ Social dimensions of science	The acceptance of new observations and hypothesis that conflict with the old observations requires approval of a significant part of the scientific community. This is a long, multi-faceted, and complex process. This process includes detailed examination of concepts in academic discussions and reciprocal dialogue and persuasion processes. In these academic discussions, theories are offered, experiments are done, and these academic discussions are influenced by social, cultural, economic, and religious factors, as well as individual or social biases (NME, 2006, p. 62, TNSES). Conceptualize the historical development of atom models and realize the electron cloud model as the more real conceptualization (GLEs for the Structure and Properties of Matter unit-7)
✓ Social and cultural embeddedness of science	Science is a human endeavor and occurs in a social context. The historical background of science shows that the asked questions and used methods in science are influenced by cultural and mental traditions while science impacts thoughts (NME, 2006, p. 62, TNSES).

Note. ✓ =explicitly presented ○ =implicitly presented ● =misrepresented × =not represented

Discussion

By using document analysis method, the study shows the success in depiction of NOS within the Turkish science and technology curriculum including TNSES, GLEs and the 7th grade teacher's edition science and technology textbook. The findings show that the portrayal of targeted NOS aspects is insufficient. Some important aspects of NOS (e.g., scientific theories and laws) are not included while some other significant aspects (e.g., inferential and theory-driven) are implicitly portrayed in these documents. On the other hand, the findings also show that the alignment between the curriculum and the textbook is not adequate in order to transfer the aims of the curriculum for NOS into classroom practices.

Recent research has shown that teachers and students in Turkey and other countries still hold naïve and uninformed views about some certain aspects of NOS while there have been many efforts and progresses made to advance students and teachers' understanding of NOS (Abd-El-Khalick, 2005; Abd-El-Khalick & Akerson, 2004; Bora et al., 2006; Dogan & Abd-El-Khalick, 2008; Irez, 2006; Koksal & Cakiroglu, 2010). One part of these efforts has focused on designing courses and professional development activities for pre-service and in-service teachers to improve teachers' understanding of NOS (Abd-El-Khalick, 2005; Abd-El-Khalick & Akerson, 2004). Though researchers have shown the benefits of such courses and professional development activities (enacted curriculum), these efforts have been insufficient to transform teachers' understanding of NOS into classroom practices because of insufficient emphasis on NOS in curriculums (planned curriculum) and textbooks (experienced curriculum) (Abd-El-Khalick et al., 2016; Dogan & Abd-El-Khalick, 2008). The results of the study also evidenced the lack of alignment between the planned curriculum and experienced curriculum for Turkish middle school science and technology curriculum. Because of the high reliance on textbooks to teach science, it stands to reason that inclusion of NOS in the textbooks is a necessary step to align planned and experienced curriculum to scaffold teachers' instruction and improve students' learning of NOS.

Many internal factors such as teachers' confidence in their understandings about NOS, teachers' perceptions and prioritization of NOS in their instruction, and their students' interests and abilities to engage in NOS instruction have been discussed as factors that impact teachers' use of NOS aspects in their instruction (Abd-El- Khalick & Lederman, 2000; Akerson & Hanuscin, 2007; Khishfe, 2008; Khishfe & Lederman, 2006). Moreover, researchers have also talked about external factors such as reform documents, special educational materials, and textbooks that are thought to mediate and support the transformation of teachers' understanding of NOS into classroom practices to advance students' conceptualization of NOS (Driver et al., 1996; McComas et al., 1998; Rudolph, 2002). In this study, we examined the Turkish science and technology curriculum as a reform document and science and technology textbook to show how these external factors portray NOS to support and mediate Turkish teachers' instruction and students's learning of NOS. As earlier studies indicated the critical role of reform documents in guiding school districts, teachers and textbook publishers to prioritize and give more attention to some instructional units such as NOS (Abd-El-Khalick et al., 2016; Dogan & Abd-El-Khalick, 2008), this study also showed the role of reform document in the Turkish educational system in guiding and motivating teachers, students and textbooks to prioritize NOS as a main aim of science education to produce scientifically literate citizens. Unfortunately, the results of the study showed that the Turkish science curriculum including TNSES, GLEs and teacher's edition textbook is not capable at a high level of guiding and motivating teachers and students in terms of NOS. The results are also in parallel with what Erdogan and Koseoglu (2012), Ozden and Cavlazoglu (2015) and Sardag et al. (2014) found after their analysis of the middle school and high school science curricula of Turkey.

Our analysis of the science and technology curriculum including TNSES and GLEs showed that NOS is insufficiently portrayed in these documents in that some important aspects of NOS (scientific theories and laws) are not included or misrepresented (the scientific method) while some other significant aspects (inferential and theory-driven) are implicitly represented. These implicit and misrepresentation of NOS aspects promote teachers and textbook publishers to hold naïve views about NOS and give less emphasis and space for instruction of NOS. Additionally, these documents do not incorporate research findings such as providing explicit-reflective teaching approaches of NOS to promote teachers' conceptualization and instruction of NOS. Theoretically, while these documents promote implicit and embedded approaches for instruction of NOS that are unproductive and conflicting with scholars' findings, in practice, the textbook does not utilize implicitly embedded NOS instruction to promote teachers' and students' understanding of NOS. This misalignment in teaching approaches is an important factor and hindrance that shows why the standards about NOS in reform documents are not transferred into classrooms (Abd-El-Khalick, 2013, 2014; Akerson et al., 2010; Chiappetta & Fillman, 2007; Wahbeh & Abd-El-Khalick, 2014). Therefore, an alignment should be provided between planned and experienced curriculum which include reform documents and textbooks, to improve NOS instruction and reduce the gap between theory and practice.

The role of textbooks on teaching and learning has been well documented by researchers. It has been shown that students first experience science during interaction with textbooks and textbooks determine the curriculum and main concerns for science teachers (Chiappetta et al., 2006; Weiss et al., 2001). This study shows that NOS is not a main part of the science and technology textbook and is not represented in an organized and consistent way to promote teachers' adoption and students' conceptualization of NOS. Some aspects of NOS (the myth of the scientific method and scientific theories and laws) are misrepresented in this textbook. These naïve representations can cause teachers and students to have misunderstandings and naïve views about these aspects of NOS (Dogan & Abd-El-Khalick, 2008; Irez, 2006; Koksal & Cakiroglu, 2010). Aspects of NOS that are implicitly presented throughout the textbook cause insufficient understanding of NOS, and are also inconsistent with current research findings. This is also supported by Ozden and Cavlazoglu's (2015) findings as they found that both the 2005 and 2013 curricula did not explicitly address NOS aspects.

The inconsistency between reform documents and textbooks is also mentioned in the literature as an inhibiting factor that limits teachers' ability to transform their conceptualization of NOS into classroom practices (Driver et al., 1996; McComas et al., 1998). Our analysis of the science and technology curriculum including TNSES and the GLEs and the textbook show that there is not a strong alignment between the science and technology curriculum (planned curriculum) and the textbook (experienced curriculum) in terms of portrayal of NOS aspects. The findings of this study show that the attention given to NOS in the curriculum and the textbook is not adequate for teachers to prioritize NOS in their instruction because of the attention given to content knowledge. Moreover, the inconsistency between the planned and experienced curriculum limits teachers' adoption of NOS to promote students' conceptualization of NOS and causes students to have naïve ideas of NOS and scientific knowledge (Abd-El-Khalick et al., 2016; Dogan & Abd-El-Khalick, 2008). The findings of this study provide information which shows a possible contributing factor to Turkish teachers' and students' naïve ideas about some aspects of NOS and why they do not have a good understanding of NOS, which were found by researchers in Turkey (Dogan & Abd-El-Khalick, 2008; Irez, 2006; Koksal & Cakiroglu, 2010).

The alignment between reform documents and textbooks on the promoted instructional approach is also important in reaching the aims of reform documents. In terms of NOS, explicit-reflective teaching approach has been highlighted as the most effective way of NOS instruction (Abd-El-Khalick & Akerson, 2004; Khishfe & Abd-El-Khalick, 2002). However, the results of our study show that the science and technology curriculum aims to promote implicit-non reflective teaching approach for NOS, which conflicts with researchers' findings. Moreover, they encourage teachers and textbook publishers to embed the aspects of NOS rather than a separate unit of instruction that addresses NOS. Encouraging teachers and textbook publishers to embed NOS into content units may be a good way to improve learners' understanding of NOS, but it should be explicit and reflective (Akerson & Hanuscin, 2007; Clough, 2006; Matkins & Bell, 2007). Otherwise, it will result in neglecting instruction of NOS by giving more emphasis and time to the instruction of content knowledge. Aligning with this, the findings of our study show that even though the curriculum aimed to embed NOS aspects into content units, the textbook mostly emphasized content knowledge and implicitly integrated a few aspects of NOS into content units. The lack of alignment between the curriculum, textbook, and research findings in terms of instructional approaches of NOS limits students' learning of NOS, teachers' NOS instruction, causes deviation from the goals of curriculum, and results in unproductive instruction of NOS (Akerson et al., 2010; Bakah et al., 2012; Marsh & Willis, 2007).

Limitations and Implications

The findings are limited by some factors as the study only focuses on the science and technology curriculum and the related 7th grade textbook. Other textbooks may provide a more comprehensive way for representation of NOS aspects and teaching of them. Furthermore, the study analyzed the 2006 science and technology curriculum because of its appropriateness for our aims. The results are also connected with Turkey's context and limited by the developers of the curriculum as well as the authors of the textbook.

This study showed that there is a lack of consistency between how NOS is conceptualized in the science education literature and how it is considered in the Turkish science and technology curriculum and textbooks. The critical question is how the transformation required for curriculum writers and textbook publishers can be mediated to portray more accurate NOS treatment to promote scientific literacy which is aimed to be reached in the curriculum we evaluated. To achieve that, planned, enacted, and experienced curriculum should focus on NOS and be aligned to support each other to scaffold teachers' instruction. At the planned curriculum level, first, decision makers of the national standards and textbooks should be informed about the importance of NOS instruction for preparing scientifically literate citizens, which is the ultimate goal of science instruction (Driver et al., 1996; McComas et al., 1998; Millar & Osborne, 1998; MNE, 2013; NRC, 1996, 2012). Decision makers should consider scientific findings rather than political factors to publish and choose appropriate standards and textbooks to support scientific literacy including NOS. Second, the curriculum and textbook writers need to be informed about effective instructional ways for teaching NOS.

In terms of enacted curriculum, the agreed upon aspects of NOS should be explicitly stated in standards and GLEs by incorporating required explanations that facilitate teachers' understandings and practices of these aspects of NOS. Furthermore, curricula and textbooks should promote the embedded or separated explicit-reflective instruction of NOS by providing appropriate educational materials for teachers to produce similar materials for their instructions as favored based on research findings.

For the experienced curriculum, textbooks should provide explicit embedded or integrated approach of NOS instruction by incorporating the common aspects of NOS. To facilitate teachers' instruction of NOS aspects, textbooks should provide appropriate teaching materials and strategies to improve instruction and learning. Moreover, appropriate professional development activities including proper teaching activities and examples need to be provided for teachers to facilitate their practices of curriculum including NOS instruction.

Lastly, a group of researchers could use same criteria used by successful reform documents to assess the curricula and textbooks for their alignment with standards. This assessment requires a consistent framework for NOS, which needs to be produced by science education researchers based on recent research findings to promote a successful integration of NOS into science curriculum.

Acknowledgment

I would like thank Suleyman Cite, a PhD candidate of Science Education for his contribution as a peer reviewer, and Deborah Hanuscin, Associate Professor of Science Education, for her critical review and suggestions on the whole manuscript.

References

Abd-El-Khalick, F. (2005). Developing deeper understandings of nature of science: The impact of a philosophy of science course on preservice science teachers' views and instructional planning. *International Journal of Science Education, 27*(1), 15–42.

Abd-El-Khalick, F., & Akerson, V.L. (2004). Learning about nature of science as conceptual change: Factors that mediate the development of preservice elementary teachers' views of nature of science. *Science Education, 88*(5), 785–810.

Abd-El-Khalick, F., & Lederman, N.G. (2000). Improving science teachers' conceptions of the nature of science: A critical review of the literature. *International Journal of Science Education, 22*(7), 665–701.

Abd-El-Khalick, F., Waters, M., & Le, A. P. (2008). Representations of nature of science in high school chemistry textbooks over the past four decades. *Journal of Research in Science Teaching, 45*(7), 835-855.

Abd-El-Khalick, F. (2013). Teaching with and about nature of science, and science teacher knowledge domains. *Science & Education, 22*(9), 2087–2107.

Abd-El-Khalick, F. (2014). The evolving landscape related to assessment of nature of science. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education* (2nd ed.) (pp. 621–650). Mahwah, NJ: Lawrence Erlbaum.

Abd-El-Khalick, F., Myers, J. Y., Summers, R, Brunner, J., Waight, N., Wahbeh, N., Zeineddin, A. A. and Belarmino, J. (2016). A longitudinal analysis of the extent and manner of representations of nature of science in U.S. high school biology and physics textbooks. *Journal of Research in Science Teaching*. doi. 10.1002/tea.21339

Akerson, V. L., Cullen, T. A., & Hanson, D. L. (2010). Experienced teachers' strategies for assessing nature of science conceptions in elementary classrooms. *Journal of Science Teacher Education, 21*, 723–745.

Akerson, V. & Hanuscin, D.L. (2007). Teaching nature of science through inquiry: Results of a 3-year professional development program. *Journal of Research in Science Teaching, 44*(5), 653-680.

American Association for the Advancement of Science. (1993). *Benchmarks for scientific literacy*. New York: Oxford University Press.

Bakah, M. A. B., Voogt, J. M. and Pieters, J. M. (2012), Curriculum reform and teachers' training needs: the case of higher education in Ghana. *International Journal of Training and Development, 16*(1), 67-76.

Bora, N. D., Aslan, O. & Cakiroglu, J. (2006). *Investigating science teachers' and high school students' views on the nature of science in Turkey*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Chiappetta, E. L., Fillman, D. A., & Sethna, G. H. (1993). Do middle school life science textbooks provide a balance of scientific literacy themes? *Journal of Research in Science Teaching, 30,* 787-797.

Chiappetta, E. L., & Fillman, D. A. (2007) Analysis of five high school biology textbooks used in the United States for inclusion of the nature of science. *International Journal of Science Education, 29*(15), 1847-1868.

Clough, M.P. (2006). Learners' Responses to the Demands of Conceptual Change: Considerations for Effective Nature of Science Instruction. *Science & Education*, *15*(5), 463-494.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham, England: Open University Press.

Erdogan, M.N. & Koseoglu, F. (2012). Ortaogretim fizik, kimya ve biyoloji dersi ogretim programlarının bilimsel okuryazarlık temaları yonunden analizi. *Kuram ve Uygulamada Egitim Bilimleri, 12*(4), 2889-2904.

Erduran, S., & Dagher, Z.R. (2014) Regaining focus in Irish junior cycle science: Potential new directions for curriculum and assessment on Nature of Science. *Irish Educational Studies, 33*(4), 335-350.doi: 10.1080/03323315.2014.984386.

Irez, S. (2006). Are we prepared? An assessment of preservice science teacher educators' beliefs about nature of science. *Science Education, 90*(6), 1113–1143.

Irez, S. (2009). Nature of science as depicted in Turkish biology textbooks. *Science Education, 9(3),* 422- 447. doi: 10.1002/sce.20305

Khisfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective views versus implicit 'inquiry orientated' instruction on sixth graders views of the nature of science. *Journal of Research in Science Teaching, 39*(7), 551–578.

Kishfe, R., & Lederman, N. (2006). Teaching nature of science within a controversial topic: Integrated versus nonintegrated. *Journal of Research in Science Teaching*, *43*(4), 395–418.

Koksal, M. S. (2010). *The effect of explicit embedded reflective instruction on nature of science understandings, scientific literacy levels and achievement on cell unit*. (Unpublished doctoral dissertation). Middle East Technical University, Ankara, Turkey.

Koksal, M. S. & Cakiroglu, J. (2010). Examining science teacher's understandings of the NOS aspects through the use of knowledge test and open-ended questions. *Science Education International, 21*(3), 197-211.

Koulaidis, V., & Tsatsaroni, A. (1996). A pedagogical analysis of science textbooks: How can we proceed? *Research in Science Education, 26,* 55-71.

Marsh, C. J. & Willis, G. (2007). Curriculum: Alternative Approaches, Ongoing Issues. (4th ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Matkins, J. J., & Bell, R.L. (2007). Awakening the scientist inside: Global climate change and the nature of science in an elementary science methods course. *Journal of Science Teacher Education, 18,* 137-163.

McComas, W. F. (1998).*The nature of science in science education: Rationales and strategies*. Dordrecht, the Netherlands: Kluwer.

McComas, M.R. (2003) A textbook case of the nature of science: Laws and theories in the science of biology. *International Journal of Science and Mathematics Education, 1*(2), 141–155.

McComas, W. F., & Olson, J. K. (1998). The nature of science in international science education standards documents. In W. F. McComas (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 41 – 52). Dordrecht, the Netherlands: Kluwer.

McNeil, J. (2003). *Curriculum: The teacher's initiative* (3rd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks: Sage Publications.

Millar R. & Osborne J.F.(1998). *Beyond 2000: Science education for the future*. London: King's College

London. National Ministry of Education. (2006). *Elementary education science and technology curriculum (Grade 6, 7, and 8).* Ankara; Talim Terbiye Kurulu Baskanlıgı.

National Ministry of Education (2011). *Elementary school science and technology: Teacher guide*. Ankara, Turkey: Devlet Kitaplari.

National Ministry of Education (2013). *Primary school (elementary and middle school) science curriculum (for grade 3, 4, 5, 6, 7 and 8).* Ankara: Talim Terbiye Kurulu Baskanlıgı.

National Science Teachers Association. (2000). NSTA position statement: The nature of science. Retrieved from http://www.nsta.org/about/positions/natureofscience.aspx.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academic Press.

National Research Council (2012). *A framework for K-12 science Education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Niaz, M. (2010). Science curriculum and teacher education: The role of presuppositions, contradictions, controversies and speculations vs Kuhn's 'Normal Science'. *Teaching and Teacher Education, 26*(4), 891-899.

Roberts, D. (2007). Scientific literacy/science literacy. In S. Abell & N. Lederman (Eds.), *Handbook of research on science education* (pp. 729–780). Mahwah, NJ: Lawrence Erlbaum Associates

Rudolph, J. L. (2002). Portraying epistemology: School science in historical context. *Sci Ed, 87***,** 64-79.

Ryder, J. (2001). Identifying science understanding for functional scientific literacy. *Studies in Science Education, 3*6, 1-44.

Sardag M., Aydın S., Kalender N., Tortumlu, N., Ciftci, M., Perihanoglu S., (2014). The integration of nature of science in the new secondary physics, chemistry, and biology curricula. *Educ. Sci., 174*(39) 233–248, DOI: 10.15390/EB.2014.3069.

Stern, L., & Roseman, J.E. (2004). Can middle-school science textbooks help students learn important ideas? Findings from project 2061's curriculum evaluation study: Life science. *Journal of Research in Science Teaching, 41*(6), 538–568.

OECD (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What 'ideas-about-science' should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching, 40*(7), 692 – 720.

Ozden, M. & Cavlazoglu, B. (2015). Ilkogretim fen dersi ogretim programlarında bilimin dogası: 2005 ve 2013 programlarının incelenmesi. *Egitimde Nitel Arastırmalar Dergisi-Journal of Qualitative Research in Education, 3*(2), 40-65.

Vesterinen, V.-M., Akesla, M., & Lavonen, J. (2013). Quantitative analysis of representations of nature of science in Nordic upper secondary school textbooks using framework of analysis based on philosophy of chemistry. *Science & Education, 22,* 1839–1855.

Wahbeh, N.,&Abd-El-Khalick, F. (2014). Revisiting the translation of nature of science understandings into instructional practice: Teachers' nature of science pedagogical context knowledge. *International Journal of Science Education, 36*(3), 425–466.

Weiss, I.R., Banilower, E.R., McMahon, K.C.,&Smith, P.S. (2001*). Report of the 2000 national survey of science and mathematics education*. Chapel Hill, NC: Horizon Research.

Author Information

Kemal Izci
Necmettin Erbakan University
Eregli College of Education, Eregli/Konya
Contact e-mail: *kemalizci@gmail.com*

Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

Examining the Relationship between Middle School Students' Critical Reading Skills, Science Literacy Skills and Attitudes: A Structural Equation Modeling

Ersin Karademir, Ufuk Ulucinar

Article Info

Article History

Received:
25 May 2016

Accepted:
09 November 2016

Keywords

Middle school students
Critical thinking
Critical reading skill
Science literacy skill
Attitudes toward science literacy

Abstract

The purpose of this study is to verify the causal relationship between middle school students' critical reading skills, science literacy skills and attitudes towards science literacy with research data according to the default model. Through the structural equation modeling, path analysis has been applied in the study which was designed in correlational model. The sample of the study consists of 1170 students enrolled in 18 schools determined based on socio-economic status. However, the results showed that validity criteria was satisfied with the 1130 participants. As data collection tool, the scale of critical reading skills, science literacy skills test, and the attitude scale toward science literacy were used. Descriptive statistics, Pearson product moment correlation analysis technique and path analysis were employed for the data analysis. The findings of the research show that there are significant and positive relations between critical reading skills, science literacy skills and attitudes towards science. In addition, critical reading skills have a statistically positive predictive effect on science literacy skill. It has also been founded that science literacy skills predict the attitudes towards scientific literacy in a meaningful way. Moreover, critical reading skills directly predict the scientific literacy in a meaningful, positive direction. It has been revealed that science literacy skill, which exists in the model, acts as partial mediating variable. As a result, as put forward in theory, critical thinking, critical reading skill, and science literacy skills and attitudes have been found to be closely related.

Introduction

In today's science education, gaining 21st century skills such as digital age literacy, exploratory thinking, effective communication and high productivity has become an increasingly popular topic in the international debate within the science curriculum attainments (McGregor & Kearton, 2010). As one of digital age literacies, scientific literacy is an output of targetting the knowledge and approach related to scientific concepts and processes to ensure personal decision making, participation in social and cultural relations and economic productivity. Since intertwined with science and technology, it has been accepted as important in our modern society (Turin man, Omar, & Osman Daud, 2012). Smith, Loughran, Berry and Dimitrakopoulos (2012) argue that because of the complexity of the concept of science literacy and different learning expectations about desired learning outcomes from students, what it means exactly in terms of in-class applications and students' learnings is a controversial issue. However, it is now seen as an objective of educational program and an essential component of conscious citizenship in many countries around the world. Not only in formal education, but also in science and art centers, museums, written and visual media, scientific journals, politics, medicine, film and drama, improving science literacy are investigated (Jenkins, 2010). Because, in terms of helping students to see how scientists think and reach scientific conclusions, it is both a fundamental component of science education and a fundamental objective of science literacy (Lawson, 2010).Within the goals of the educational program in question, as well as providing students to gain scientific concepts, content and conceptual knowledge; scientific classifications, time, and developing cognitive concepts about causality of both physical and psychological events take part as well (Westby & Torres-Velásquez, 2000). In terms of literacy skills, science learning takes place by learning how to use the language of science as well as learning the facts and definitions or experimental process of science. As a result, learning a language requires opportunities to use it and writing science in standard forms (Wellington and Osborne, 2001). Thus, Yore, Hand, & Florence (2004) defined that a literacy skill includes language-based actions like reading, writing, listening, speaking and monitoring. Although each action takes place in the context of all disciplines, it is claimed that language is the basis of learning science and therefore science literacy.

Act of reading, as well as having an important role in the acquisition of science is one of the main attainments of the many native language programs. Most of the time, reading is seen as an action that students take only to find out the information in the text. In this reading process, students just look for some facts and assume that the resource is true. Since this reading process takes place in a period of time as short as possible, it limits students' use of mental skills (Wheeler, 2007). Most of the teachers at schools today, unfortunately, can not prepare their students as critical thinkers. For this reason, many of the students have stated that they have been unable to develop adequate critical reading skills (Kadir, Subki, Jamal, and Ismail, 2014). In terms of educational view, in order for a student to be a critical reader, he/she should spare time to examine the argument in the text in logical, theoretical, historical, ethical, social, and personal aspects because critical reading involves a thinking process that questions the results and the accuracy of these results in the text the reader deals with.Thus, critical thinking, which is a kind of high-level thinking skill, requires a reader to participate in reading process in an active and constructivist way. In order to prepare students for a complex world as more scientifically and technologically, the best education applications are required. Starting from pre-school, children must learn to be able to think critically, to integrate information correctly and to solve problems in new situations (Center for Science, Mathematics, and Engineering Education, 1998). In this research, by putting forward the causal relationship between critical reading skill as high-level thinking action, science literacy out of the 21st century skills and attitudes towards science literacy, it has been aimed to make recommendations about the goals and objectives of the educational program.

Critical Reading

Accepted as an education ideal, higher order thinking (Siegel, 1988; Lipman, 1995), is a skill that all educators agree on importance of its achievement (Shahrokh, 1998). As in all disciplines, in reading education, too, the development of higher order thinking skills is emphasized **(**Collins, 1993; Paul & Elder, 2008; Zabihi & Pordel, 2011; Barnett & Bedau, 2011). Reading, comprehension, analysing and integration of a text is the basis of reading. For this reason, reading is a process which lays out a higher order thinking potential. Higher reading comes true effectively when the reader can relate the known with new information in order to answer some questions (Collins, 1993). According to Criscuolo (1965), critical reading is a critical process in which higher order mental process is used for interpreting and evaluating the information read. In this process, while developing his critical reading skill, a student must compare what they read with the previous information. In the context of the relationship between critical thinking and critical reading, critical thinking means critical reading. According to researchers explaining critical reading with critical thinking skills, critical reading is students' technique of analyzing what they read, synthesizing and evaluating (Kadir, Subki, Jamal and Ismail, 2014; Akin, Koray and Tavukcu, 2015), discovering the information and evaluating (Zabihi and Prodel, 2011). Paul and Elder (2008), stated that critical reading is the art and science of evaluating and analyzing a text using a perspective to improve the nature of thinking and behavior of the individual**.** According to Paul and Elder, critical reading strategies contain: (1) identifying a problem or issue, (2) establishing meaning (3) making interpretation in line with the evidence, (4) providing strong assumptions, (5) making applications and (6) taking a different perspective. If the student does not make an interpretation for the text in the simplest level, it is unlikely for him to understand the text (Burnett and Berg, 1988). So, critical reading should be seen as an educational ideal in schools. According to Lewis (1983), the purpose of teaching critical reading is to grow individuals who can reach some judgments about what they read depending on strong evidences and very strong reasoning process more than subjective data**.** Together with this purpose, it helps individuals to develop thinking strategies. It was found that while individuals reading critically had more important notes and markings, those who read less critical took less summary notes. At the same time, critical readers tend to take critical notes and make marking through reading course process (Kobayashi, 2007). Effective critical readers have a number of strategies that are used to improve their critical thinking skills (Walker, Kiefer and Reid, 1994-2012). These techniques represent ways to talk about ideas, graphic organizers and story maps for narrative texts (deVoogd, 2007). On the other hand, Yang (2006), has studied to find out the relationship between reading strategies and comprehension monitoring strategies and how these strategies affect their understanding process. In the study, it was found that reading strategy is a mental process that is used while solving problems caused by the lack of language skills in the process of an individual's understanding of the text. Aregu (2013), unlike, has examined the effects of self-learning strategies on critical reading performances, and the study results showed that learning strategies have significant effects on critical reading performance.

Science Literacy Skill

Because of the complexity of the concept and different learning expectations about desired learning outcomes from students, what science literacy, one of 21st century skills, means exactly in terms of in-class applications and students' learnings is a controversial issue (Smith, Loughran, Berry and Dimitrakopoulos, 2012). Pella and others (1966) who made first studies on scientific literacy, suggested the qualifications that a science literate individual should have. According to them, science literate individuals should (a) have science and the internal relations of society (b) and the ethics controlling the studies of scientists, (c) have the understanding of the science of nature, (d) know the difference between science and technology, (e) know the basic concepts of science and (f) have the understanding of internal relations of individuals and science. Moreover, according to Aron (1983), these individuals have the ability to (a) define that scientific concepts are discovered or created by the actions of human intelligence and imagination; (b) understand the difference between observation and inference; (c) form hypothesis in a planned way and realize testing methods and (d) define their own learning methods, be aware of the reasons of what they believe and prove the evidences they reach on the subject they examine. On the other hand, it is seen that whole of scientific literacy has been defined in four componenets at PISA 2006. Scientific context (living conditions including science and technology, etc.); scientific skills (identifying scientific issues, explaining a phenomenon with scientific explanations and using scientific evidence); learning areas for scientific knowledge (understanding of nature related to the science as well as the understanding of students about the scientific concept) and attitudes towards science (interest in science, scientific inquiry support and to have responsibility for resources and the environment) (OECD, 2006).

In the context of educational activities, it has been revealed that creative and collaborative learning environment (Oluwatelure, 2010), inquiry-based learning community (Nwagbo, 2006; Lewis, 2010), metacognition-based training (Michalsky, Mevarech and Haibi, 2009), authentic scientific inquiry (Hume, 2009) socio-critical and problem solving based chemistry teaching (Marx and Eilks, 2009), teaching based on scientific process (Genç, 2015), the use of scientific texts (Parkinson's and Adendorff, 2004), integrated teaching strategies approach (Webb, 2009) and argumentation method (Washburn and Cavagnetto, 2013)improve science literacy skills. Furthermore, Lin Hong and Huang (2012), studied the relationship between students' affective characteristics such as interest, enthusiasm and participation in science education and scientific literacy. It has been found that interest, enthusiasm and participation that students have in learning science increase their science literacy skill.

Relationship between Critical Reading Skills, Science Literacy Skills and Attitudes towards Science Literacy

Science literacy and critical reading are intertwined with each other. Because, an individual must be able to think along with the evidence to be a science-literate. For this, a combination of the high-level language and thinking skills are needed (Hackling and Sherriff, 2015). Wellington & Osborne (2001), pointed out that reading is a scientific activity. They stated that the ability to read in a careful, critical and a healthy doubtful manner is the basic feature of being a scientist and that the skill to read critically, carefully and with a healthy sceptisizm is the main component of science literacy**.** Learning of science texts thoroughly depends on such skills as making inferences, drawing conclusions, showing a time sequence and chronology, forming hypothesis, realizing the result. Therefore, the above-mentioned critical reading skills make it possible for students to analyze, to interprete and to evaluate in a critical way the concepts, nature and processes of science. For this reason, as stated in the H1 hypothesis, it is assumed that the critical reading skills predict science literacy skills. In H2 hypothesis, it is assumed that science literacy skills predict attitudes towards science literacy. The individual's tendency to questioning causal relationship of scientific propositions in scientific texts, trying to understand how it happens in this process, having curiosity and understanding the knowledge about scientific process, reflect the attitude towards science.

An individual who is successful in terms of science literacy skills is thought to have increasing attitudes toward science. On the other hand, critical reading individuals, as previously stated, use such higher order skills as to analyze, to interpret and to evaluate information about the text they read. An individual who has these skills is likely to have a tendency to wonder the knowledge, concept, process or progresses related to science; to try to understand and question them. Accordingly, in H3 hypothesis, critical reading skill is assumed to have an impact on predicting the attitudes towards science literacy. At the same time, within this relationship pattern, science literacy skill is assumed to have the effect of a mediating variable between attitudes toward science literacy and critical reading skills.

H_1: Critical reading skills of middle school students predict science literacy skills in a meaningful way.

H_2: Science literacy skills of middle school students predict attitudes toward science literacy in a meaningful way.

H_3: Critical reading skills of middle school students predict attitudes toward science literacy in a meaningful way.

H_4. Science literacy skills variable has a mediation effect between attitudes towards reading skills and science literacy.

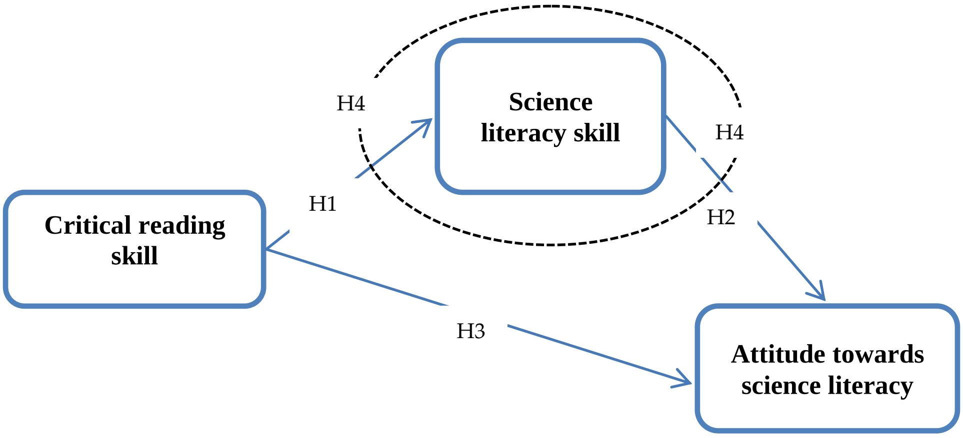


Figure 1. Critical reading skills, the default model of the causal relationship between science literacy and science literacy attitudes towards

Method

Research Model

This research, in which the causal relationship between middle school students' critical reading skills, science literacy skills and attitudes towards science literacy is studied, was designed as correlational model. Studies designed by the correlational model are done to reveal the relationship between two or more variables and to obtain results about cause-effect (Christensen, Johnson and Turner, 2015). In the study, path analysis, which is a structural equation model compatible with the research modelling, has been employed. In line with this model, it was aimed to verify the causal relationship between students' critical reading skills, science literacy skills and attitudes towards scientific literacy, according to data collected through the default model. Out of the variables in the default model, critical reading skills have been dealt with as independent variable; science literacy skills as intermediary (mediator) variable and attitudes towards science literacy as dependent variables.

In the default model, one of the two modellings in SEM, structural model was adopted since verifying the structural relations of variables with each other was intended. Before proceeding to testing the model, two conditions are in question in the fulfillment of the suppositions of SEM (Kline, 2012: 112). These are controlling; (1) the existence of a causal relationship and (2) the structure of data distribution. Kline draws attention to five conditions that must be met in a structural model regarding the existence of a causal relationship between two variables (a) the default cause variable should take place before the effect variables. A temporal priority is in question. (b) there should be a covariance or correlation between the cause and effect variable. (c) in the cause and effect relationships, it must not be affected by other variables that can also affect the effect variable (d) the distribution form of data must be checked. (E) the direction of the causal relationship must be determined accurately (contrary to Y affecting X, X affecting Y actually or the interaction acting of X and Y each other).

The Population-Sample

The population of study constitutes sixth, seventh and eighth grade students studying in secondary schools in Eskişehir. In the sample, 18 secondary schools took part which have been stratified as upper-middle-low socioeconomic status. In the selection of the students in these schools, the convenient sampling method was applied, 1170 middle school students participated in the research. Following the processes such as extracting extreme values and missing data in scales applied to students, 1130 students' data were included in analysis and constituted the study sample. The following table shows the frequency and percentage distributions of demographic information of sample group.

Table 1. Demographics of the sample group

		N	%
Gender	Female	547	48.4
	Male	583	51.6
Grade	Sixth	379	33.5
	Seventh	396	35.0
	Eighth	355	31.4
Total		1130	100

A total of 1130 students participated in the study; 547 of them are (48.4%) female, and 583 (51.6%) are male. 379 of these students are (33.5) at 6th grade; 396(35.0%) are at 7th grade and 355(35.0%) are at 8th grade

Data Collection Tools

Critical reading skills scale:

In this study, in order to identify students' critical reading skills attainment levels, "critical reading skills scale" , which was developed by Unal (2016), was used. The scale consists of 22 items within critical reading context. The scale is in five-point Likert-type; rated as "Always, usually, sometimes, rarely, never". Cronbach alpha reliability value of the scale was .872. Reading comprehension- based text samples test: In order to determine the students' science literacy skills and attitudes, reading comprehension-based text samples test (science literacy test) has been used. The test composed of two parts: a) Science literacy skills test: It is prepared to determine the student's science literacy skills; consisting four texts related to science (working under temperature, genetics, solar and electric) and multiple-choice questions that can be answered by making inferences from these texts. Working under temperature and genetics related texts are taken from sample PISA science questions released by Ministry of Education (2013). Science texts about sun and electric have been prepared based on the opinions of three faculty members expert in the field of science education, two science teachers and a Turkish language expert, within the context of sample PISA questions. In this test, students were given texts as presented below. Students were asked to read the text and answer true-false and multiple-choice questions formed according to this text. A sample science literacy skills test question is presented below.

Question 13.1: Working under Temperature

Murat is working in the repair of an old house. He has left a bottle of water, a bit of metal nail and a piece of lumber in the trunk of his car, After staying under sun for three hours, the temperature inside the car reaches approximately 40 degrees. What happens to objects in the car? For each circle "Yes" or "No"

Is this; happen to the objects?

They all reach at the same temperature. Yes/ No
After a while, the water begins to boil. Yes/ No
After a while, metal nails start flushing. Yes /No
The temperature of the metal nail is higher than the temperature of the water. Yes/ No

Question 13.2: Working under Temperature S420Q03

Murat has drunk a cup of coffee at a temperature 90 ° C, a mineral water at a temperatura 5 ° C during the day. The cups are in the same shape and size and each drink has the same volume. Murat leaves the cups in a room where the temperature is 20 degrees.

What can be the temperature of **coffee** and **mineral water** 10 minutes later?

A) 70 ° C and 10 ° C
B) 90 ° C and 5 ° C

C) to 70 ° C and 25 ° C
D) 20 ° C and 20 ° C

b) Science literacy attitude scale: There are items related to attitude in each of sample PISA science questions in science literacy skill test released by Ministry of National Education (2013). The scale, in which these are included, is called scientific literacy attitude scale. Sample test items of attitude scale used within the context of the text “working under temperature” are as follows: "Understanding how the shape of the container affects the cooling rate of coffee”, “learning about different arrangements of atoms in wood, water, and steel", "knowing why different solids conduct heat differently”. Students were asked to read the propositions set out above and to express their attitude with a question in the form of" to what extent are you interested? ".The grading format of Likert-type scale is "high level of concern", "moderate concern," "concern at the low level" is "does not concern”.

Data Analysis

In order to understand whether data-set is compatible with the default model, path analyzes were performed using maximum diversity calculation in AMOS 21.0. In this analysis, the relationship between the variables in the default model were tried to be determined. Before starting Path analysis, descriptive statistics were made on the implicit (latent) and observed variables in the model. In addition, in planned model, the variables called “working under temperature”, “electric”, ”genetics” and “sun” under science literacy skill (latent variable) are observed variables. Under attitude towards science literacy, another implicit variable, “working under temperature” and “ genetics” represent the observed variable of the model. On the other hand, critical reading skill, too, was included in the model as observed variable. Arithmetic mean and standard deviation values were calculated for each variable. Pearson Product Moment Correlation analysis technique was applied to demonstrate the relationship between critical reading skills, science literacy skills and attitude variables.

Before being subjected to the path analysis, the model went under missing data and extreme values extraction process. In the arrangement of the data contained in continuous variables (critical reading and science literacy attitude scale)the average of the variables in the same series instead of data losses were calculated by 'Series mean’ method. In classified variables (science literacy test), participants data where lost data exists were excluded from the analysis. After the data extraction and editing process, the data was made available for structural equation modeling analysis. Because the researcher is to make the data ready for a software specialized in structural equation modeling studies (Malone & Lubansky, 2012). In assessing the compliance of the model, chi-square fit test ($\chi2$), the comparative fit index (CFI), the incremental fit index (IFI) and the root mean square of about errors (RMSEA) values were calculated (Kline, 2011; Schumacker and Lomax, 2010). The chi-square statistical value ($\chi2 = 3.394$, df = 12, $\chi2$ / df = .283, p = .992) of the model, which is taken into consideration in testing the general harmony between the default model and the data, describes the competency of model (Bentler, 2006). Also, it is known that CFI = 1.00; IFI = 1.010 RFI = .991; fit values get acceptable as fit index of the model gets closer to 1; on the other hand, as [RMSEA = .044] value gets closer to 0, the model is known to show acceptable perfect fit (Arbuckle, 2008).

Results and Discussion

Descriptive statistics and correlation values for the variables in the study are presented in Table 2. While science literacy skills scores are at low-level associated with science literacy attitude scores [r = .148] and critical reading skills [r = .098], scores of attitude towards the scientific literacy are significantly and positively correlated with critical reading skills at moderate-level [r = .346].

The figure 2 showing the path analysis which was made to verify the compliance of theory-based created model with data and coefficients of variables’ relating to each other are presented below.

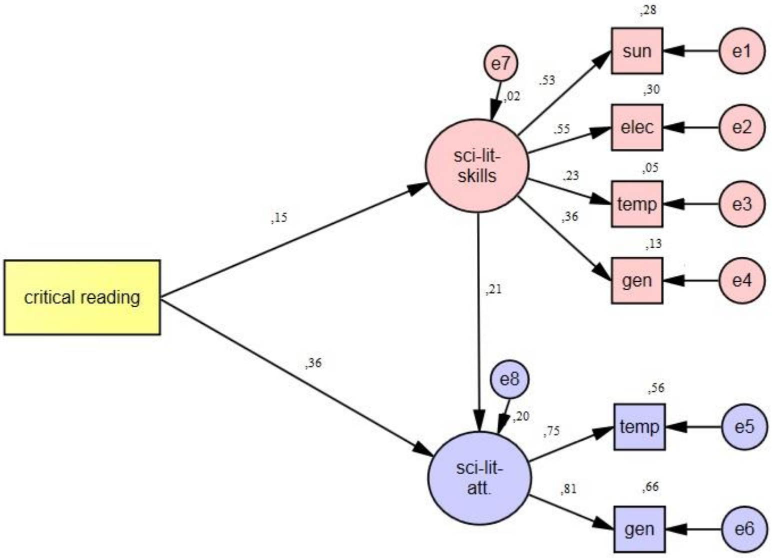


Figure 2. A path model between scientific literacy skills, attitude toward scientific literacy and critical reading skills [Sci-lit-skills: Scientific literacy skills, Sci-lit-att.: Attitude toward scientific literacy, Sun: Sun, Elec.: Electricity,
Temp.: Working under temperature, Gen.: Genetics]

Table 3. Standardized regression coefficients, standard error and significance values of variables in path model

Independent variable	Dependent variable	Std. regression coefficient	Standard Error	(p)	Explained Variance
Critical reading skill	Attitude toward science literacy	.36	.023	.000	% 20.2
Science literacy skill	Attitude toward science literacy	.21	.071	.000	
Critical reading skill	Science literacy skill	.15	.028	.000	% 2.3

As seen in Table 3, the critical reading skill predict attitudes towards science literacy scores (β = .36), and science literacy skills (β = .15) in a significant and positive direction (p <.05). According to this result, Hypotheses H_1 and H_2 are acceptable. On the other hand, science literacy skill (β = .21) predicts science literacy attitudes scores in a meaningful way in a positive direction. This finding indicates the acceptance of H_3 hypothesis. In addition to this, of the variables in the model, critical reading skill and attitudes towards science literacy scores has been found to represent 20.2%. Critical reading skills explains 2.3% of the science literacy skills.

Table 2. Science literacy skills test scores, science literacy attitude scores, average, standard deviation values of critical reading skills and correlation coefficients

Variables	$\bar{X}$	**s**	**(1)**	**(2)**	**(3)**
(1) Science literacy skills test	6.39	2.43	1		
(a) Working under temperature	2.42	.95			
(b) The Sun	1.31	.79			
(c) Electricity	1.07	.78			
(d) Genetic	1.52	1.02			
(2) Science literacy attitude scale	3.20	.64	.148	1	
(a) Working under temperature	3.18	.72			
(b) Genetic	3.21	.72			
(3) Critical reading skills scale	1.07	.78	.098	.346	1

*there is a significant at p<.01, ** p<.05

As stated in hypothesis H_4, the mediation effect of science literacy skill between critical thinking skill and attitudes towards science literacy was tested. Baron and Kenny (1986) stated that three conditions were to be present for a variable to be mediating variable. These are; (1) independent variable or variables'(critical reading skill) affecting the mediating variable (science literacy skill), (2) mediating variable's (science literacy skill) having a meaningful effect on dependent or dependent variables(attitudes towards science literacy), (3) independent variables' (critical reading skill) showing meaningful effects on dependent variables (attitudes toward science literacy)(see Figure 1.). However, in this model, because the predicting level of critical reading skill and scores of attitude towards science literacy is meaningful, science literacy variable has"partial mediating effect". Moving from this point, in the direction of the default theory, this theory can said to be confirmed. Mediation analysis is performed with the aim of predicting the role of causal mechanisms(variables) that transfer an independent variable's impacts on dependent variable (Hicks and Tingley, 2011).

Table 4. Sobel, Aroian and Goodman tests' results related to significance mediation effect between dependent and independent variables

Independent variable	R.C.	S.E.	Mediator variable	R.C.	S.E.	Dependent variable
Critical reading skill	.36	.023	Science literacy skill	.21	.071	Attitude toward science literacy
Sobel test	z 2.58	p .009	*Aroian test*	z 2.56	p .010	*Goodman test* z p 2.62 .008

R.C. Regression Coefficient; S.E. Standard error

According to Sobel, Airon and Goodman mediation test results, z values of science literacy skills (mediating) variable's mediating effect between critical reading skill (independent) variable and attitude towards science literacy (dependent) variables are respectively z[Sobel] = 2.58; z[Airon] = 2.56; z[Goodman] = it is 2.62. According to z values calculated in these tests, it was found that the science literacy skill is meaningful as a mediating variable. Proceeding from these findings, science literacy skill variable has a meaningful mediation effect between critical reading skills and attitudes towards science literacy variables in the model. Critical reading skills of middle school students was confirmed by research data-set of structural model constructed with variables indicated in the study examining assumed causal relationship between attitudes towards science literacy and science literacy skills.

Path analysis results conducted in order to test H1, H2, H3 and H4 hypothesis show the significance of causal relationship between the variables in the default model. As tested in H1 hypothesis, critical reading skills predict the science literacy skills of students in a meaningful way. In other words, these finding indicates that students' critical reading skills play an effective role in the development of science literacy skills. When the literature is examined, science literacy and critical reading is observed to be interrelated skills with each other. Because, to be science-literate, an individual must be able to think along with the evidence. To do this, a combination of the higher order language and thinking skills is required (Hackling and Sherriff, 2015). Some studies describe critical reading as critical thinking skills like analyzing and evaluating what they read, exploring and evaluating the information (Kadir, Subki, Jamal and Ismail, 2014; Akin, Koray and Tavukcu, 2015; Zabiha and Prodel, 2011). In addition, Majima (2012) stated that one of the main components of scientific literacy is the ability to critically analyze and evaluate a scientific claim and evidence. These expressions show the presence of critical thinking skills in the development of scientific literacy. Similarly, Viera & Tenreiro- Viera (2016), reported that science literacy and critical thinking are the basic components of science education. A study, targeting science literacy basically, focusing on critical thinking and including the design, application and evaluation of science learning activities, was carried out. Learning activities which were designed and practiced for 6th grade students, provided a significant impact students' critical thinking skills and science literacy skills. In line with this finding, Blake (2015) explains that science literacy skills require critical thinking. Understanding the science text includes such skills as making inferences, drawing conclusions, showing a time sequence the chronology, forming hypothesis, realizing cause and effect.

Therefore, the above-mentioned critical reading skills of students enable them to analyze, interpret and evaluate the concepts, nature and processes of science in a critical way. Yu (2013), indicates that science literacy skill is intensely based on new words, concepts, information and phenomenons. For him, understanding the contents of scientific texts effectively depends on four factors. These factors are; (1) scientific language, (2) vocabulary, (3) the structure of the text, (4) motivation and (5) the purpose and reading without prior knowledge of. Indeed, Khabir of & Pakzad, (2012), has determined that teaching of critical reading strategy contributes to students' word retention. Depending on Yu's study, it can be said that students' use of vocabulary more effectively thanks to critical reading strategies improve their science literacy skills. Yu's research result supports the findings of the

H1 hypothesis tested in this study. It is clear that as well as critical reading skills and critical thinking skills, critical reading strategies have effect on science literacy skills of students. In addition to this inference, Jurečka & Wander (2012) presented an academic approach in critical evaluation of the scientific literature and other texts. They collected the four-element criteria in two groups. While the first group is originality and competence assessing the nature of the source; the second group is objectivity and validity which assess the nature of the information contained in the source. They develop a concrete evaluation system based on scientific terminology with this approach. The ultimate aim is to improve the students' science literacy skills. Similarly, Geithner & Pollastro (2016) used peer review in their work in order to improve students' critical reading and writing skills. In this study, the effect of students' peer review related to 'Human Physiology' on science literacy skills has been examined. According to them, peer review has been stated as quite common among 21st Century learning activities which increase scientific literacy stated These studies clearly indicates that findings related to critical reading in science education are largely related with science literacy skills. As supporting this relationship, Cavagnetto (2010) stated that science literacy skill contributes to increasing of activities based on argumentations in the context of science education. Additionally, he suggests that students' involvement in these argumentations develops their communication skills, metacognitive awareness, critical thinking skills and thus their science literacy skills. As tested in H_2 hypothesis, science literacy skills predict their attitudes toward science literacy in a meaningful way. This finding shows that students are able to critically analyze, interpret and evaluate scientific texts, and have a causal questioning tendency or attitude towards the information and propositions in these texts. According to this explanation, it is possible to say that the H2 hypothesis is acceptable. On the other hand, as assumed in hypothesis H_3, an individual applying critical reading skill can be said to have a questioning attitude or a tendency towards the information, concept or propositions in science texts

Recommendations

In light of these findings, in science education learning experiences and design of units have been recommended based on students' critical thinking skills and critical reading skills. In learning environment, while teaching the concepts, information and process about science to students, teachers should make use of critical reading strategies that allow students to question, analyze, interprete and evaluate better. This will allow a more effective and permanent science learning and resulty, the student will reach the science literacy, which is one of the goals of science teaching. Student learning has a permanent program that gains access to science literacy is one of the goals. With science literacy, as put forth theoretically, too, the communalization of studies regarding the planning, application and evaluation of thinking-oriented science units based on higher order thinking skills has been considered.

References

Akın, F., Koray, Ö. & Tavukçu, K. (2014). How effective is critical reading in the understanding of scientific texts?. *Procedia - Social and Behavioral Sciences, 174*(12), 2444 – 2451.

Arbuckle, J. L. (2008). *Amos 17.0 user's guide*. Crawfordville, FL: Amos Development Corporation.

Aregu, B. B. (2013). A study of self-regulated learning strategies as predictors of critical reading. *Educational Research and Reviews, 8*(21), 1961-1965.

Arons, A. B. (1983). Achieving wider scientific literacy. *Daedalus, 112*(2), 91–122.

Baron, R. M. ve Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Bentler, P. M. (2006). *EQS 6 structural equations program manua*l. Encino, CA: Multivariate Sofware, Inc.

Blake, C. (2015). Teaching science literacy to all grade levels. http://online.cune.edu/teaching-science-literacy/ adresinden erişilmiştir.

Burnett, E. H. & Berg, P. C. (1988). Reading ınstruction in the schools: ımproving students' critical thinking skills. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 61*(5), 208-210.

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument ınterventions in K–12 science contexts. *Review of Educational Research, 80*(3), 336-371.

Center for Science, Mathematics, and Engineering Education (1998). *Every child a scientist: Achieving scientific literacy for all*. Washington: National Academy Press,

Christensen, L. B., Johnson, R. B. ve Turner, L. A. (2015). *Araştırma yöntemleri: Desen ve analiz* (A. Aypay, Çev.). Ankara: Anı Yayıncılık.

Criscuolo, N. P. (1965). A plea for critical reading in the primary grades. *Peabody Journal of Education, 43*(2), 107-112.

Cromley, J. G. (2009). Reading achievement and science proficiency: International comparisons from the programme on international student assessment. *Reading Psychology, 30*(2), 89-118.

DeVoogd, G. L. (2007). Critical comprehension of social studies texts. *Random House Inc: An annual magazine for Educators, 2,* 21-24.

Fives, H.; Huebner, W.; Birnbaum, A. S.ve Nicolich, M. (2014). Developing a Measure of Scientific Literacy for Middle School Students. *Science Education, 98*(4),549-580.

Geithner, C. A. & Pollastro, A. N. (2016). Doing peer review and receiving feedback: impact on scientific literacy and writing skills. *Advances in Physiology Education, 40*(1), 38-46.

Genç, M. (2015). The effect of scientific studies on students' scientific literacy and attitude. *OMU J. Fac. Educ., 34*(1), 141-152.

Güçlüer, E. ve Kesercioğlu, T. (2012). The effect of using activities improvıng scientific literacy on students' achievement in science and technology lesson. *International Online Journal of Primary Education, 1*(1), 8-13.

Hackling, M. ve Sherriff, B. (2015). Language-based reasoning in primary science. *Teaching science, 61*(2), 14-25.

Hicks, R. ve Tingley, D. (2011). Causal mediation analysis. *The Stata Journal, 11*(4), 1-15.

Holbrook, J. ve Rannikmae, M. (2009). The meaning of scientific literacy. *International Journal of Environmental & Science Education, 4*(3), 275-288.

Hume, A. (2009). Authentic scientific inquiry and school science. *Teaching Science, 55*(2), 35-41.

Jenkins, E. (2010). How might research ınform scientific literacy in schools?. *Education in Science, 239*, 26-27.

Jurecki, K. & Wander, M. C. F. (2012). Science literacy, critical thinking, and scientific literature: Guidelines for evaluating scientific literature in the classroom. *Journal of Geoscience Education, 60*(2), 100-105.

Kadir, N. A., Subki, R. N. S., Jamal, F. H. A., & İsmail, J. (2014). The importance of teaching critical reading skills in a Malaysian reading classroom. *The 2014 WEI International Academic Conference Proceedings*, Indonesia.

Khabiri, M. & Pakzad, M. (2012). The Effect of Teaching Critical Reading Strategies on EFL Learners' Vocabulary Retention. *The Journal of Teaching Language Skills, 4(*31), 73-106.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3. bs.). New York/London: The Guilford Press.

Kobayashi, K. (2007). The influence of critical reading orientation on external strategy use during expository text reading. *Educational Psychology, 27*(3), 363-375

Lawson, A. E. (2010). Basic inferences of scientific reasoning, argumentation, and discovery. *Science Education, 94,* 336-364.

Lewis, J. (1983). Influence of cognitive and ethical development on critical reading. *Journal of College Reading and Learning*, *16*(1), 15-20.

Lewis, K. A. (2010). *Constructing scientific literacy in ınquiry-based communities of science practice*. (Unpublished Doctorate Thesis), University of California, Santa Barbara.

Lin, H-S., Hong, Z.-R.ve Huang, T-C., (2012). The role of emotional factors in building public scientific literacy and engagement with science. *International Journal of Science Education, 34*(1), 25-42.

LoGiudici, R. ve Ende, F. (2010). Teaching toward a more scientifically literate society. *Science Scope, 33*(9), 858-63.

Majima, Y. (2012). The role of critical thinking in acquisition of science literacy and "Subjective demarcation" between science and pseudoscience. *Cognitive Studies, 19*(1), 22-38.

Malone, P. S. & Lubansky, J. B. (2012). Preparing data for structural equation modeling: Doing your homework. Rick H. Hoyle (ed.), *Handbook of Structural Equation Modeling*. New Yok: The Guilford Press.

Marks, R. ve Eilks, I. (2009). Promoting scientific literacy using a sociocritical and problem-oriented approach to chemistry teaching: Concept, examples, experiences. *International Journal of Environmental & Science Education, 4*(3), 231-245.

McGregor, D. ve Kearton, G. (2010). What do researchers say about Scientific Literacy in Schools?. *Education in Science, 240,* 22-23.

M.E.B. (2013). Uluslararası Öğrenci Başarılarını Değerlendirme Programı, PISA Örnek Fen Soruları, retrieved from http://pisa.meb.gov.tr/wp-content/uploads/2015/02/pisa-ornek-sorular- fen.pdf [25.05.2016].

Michalsky, T., Mevarech Z. R. & Haibi, L. (2009) Elementary school children reading scientific texts: effects of metacognitive instruction. *The Journal of Educational Research, 102*(5), 363-376.

Norris S. P. & Phillips, L.M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*, 224-240.

Nwagbo, C. (2006). Effects of two teaching methods on the achievement in and attitude to biology of students of different levels of scientific literacy. *International Journal of Educational Research, 45,* 216–229.

OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006.* Paris: OECD.

Paul, R. & Elder, L. (2008). *The miniature guide to critical thinking concepts and tools*. Dillon Beach, CA: Foundation for Critical Thinking Press.

Schumacker, R. E. ve Lomax, R. G. (2010*). A beginner's guide to structural equation modeling*. New York: Routledge.

Smith, K., Loughran, J., Berry, A. & Dimitrakopoulos, C. (2012). Developing scientific literacy in a primary school. *International Journal of Science Education, 34*(1), 127 - 152.

Ünal, E. (2006). *İlköğretim öğrencilerinin eleştirel okuma becerileri ile okuduğunu anlama ve okumaya ilişkin tutumları arasındaki ilişki*. Yayımlanmamış Yüksek lisans tezi, Eskişehir Osmangazi Üniversitesi.

Vieira, R. M. & Tenreiro-Vieira, C. (2016). Fostering scientific literacy and critical thinking in elementary science education. *International Journal of Science and Mathematics Education,14*(4), 659–680.

Walker, D., Kiefer, K., & Reid, S. (1994 - 2012). Critical reading. Writing@CSU. Colorado State University. Available at http://writing.colostate.edu/guides/guide.cfm?guideid=31.

Wellington, J. & Osborne, J. (2001). *Language and literacy in science education*. Buckingham Philadelphia: Open University Press.

Yu, J. (2013). Science literacy: Reading strategies among secondary reader. Master's Level Graduate Research Conference. Paper 30. Retrieved from http://digitalcommons.brockport.edu/gradconf/2013/program/3 [24.05.2016]

Zabihi, R. & Pordel, M. (2011). An investigation of critical reading in reading textbooks: A qualitative analysis. *International Education Studies, 4*(3), 80-87.

Author Information

Ersin Karademir
Eskişehir Osmangazi University
Faculty of Education, Maths and Science Education
Department Meşelik Campus 26480 Eskişehir
Contact e-mail: *eekarademir@gmail.com*

Ufuk Uluçınar
Uşak University
Faculty of Education, One Septemper Campus
64200,Uşak

JESEH
Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

Evaluating Middle School Students' Spatial-Scientific Performance within Earth/Space Astronomy in Terms of Gender and Race/Ethnicity

Jennifer Wilhelm, Michael Toland, Merryn Cole

Article Info

Article History

Received:
30 August 2016

Accepted:
21 November 2016

Keywords

Middle school education
Geometrical and spatial thinking
Gender and race/ethnicity
Astronomy education

Abstract

Differences were examined between groups of sixth grade students' spatial-scientific development pre/post implementation of an Earth/Space unit. Treatment teachers employed a spatially-integrated Earth/Space curriculum, while control teachers implemented their Business as Usual (BAU) Earth/Space units. A multi-level modeling approach was used in a hierarchical manner to evaluate student performance on the Purdue Spatial Visualization: Rotation test (PSVT-Rot) and on the Lunar Phases Concept Inventory (which included four spatial domains), while controlling for two variables (gender and race/ethnicity) at the student level and one variable (teaching experience) at the teacher level. Results showed Treatment girls achieved higher LPCI Periodic Patterns (PP) spatial domain post-scores than girls in the BAU group. A gender gap was also observed (in favor of boys) within the BAU group for PP domain post-scores, while no gap was shown within the Treatment group. In addition, results for PP suggest Students of Color tended to have lower PP scores than White students (Effect Size = .29), and that higher pretest PP scores tended to lead to higher posttest PP scores, after adjusting for other student and teacher characteristics. The only statistically significant predictor of the PSVT-Rot posttest scores were scores on the respective pretest.

Introduction

Research has shown gender differences on students' spatial understandings in favor of males, particularly for spatial visualization and mental rotation (Kaufman, 2007). Linn and Petersen (1985) determined that males outperformed females at all age levels on mental rotation tasks. Numerous studies have also shown a substantial gap in mathematical achievement between Black and White students (Lee & Wong, 2004; Reyes & Stanic, 1988) which is further intensified among Hispanic and White students (Lubienski, 2002). However, research focusing on spatial reasoning and visualization among students of color is underdeveloped.

Studies have shown relationships between students' spatial abilities and their understanding of scientific phenomena (Black, 2005), especially in the area of Earth/Space science. Rudmann (2002) found students' inclination to learn scientific explanations for the cause of the seasons was restricted by their spatial aptitude. Similarly, Wellner (1995) reported students were more likely to describe a correct cause of lunar phases when they had a strong spatial sense. Other studies claimed understanding celestial motion demands the skill of moving between frames of references (Plummer, Wasko, & Slagle, 2011; Plummer, 2014).

This study builds on previous research (Wilhelm, 2009; Wilhelm, Jackson, Sullivan, & Wilhelm, 2013) and examined differences between groups of students' spatial-scientific development from pre to post implementation of an Earth/Space unit. Wilhelm's (2009; 2013) prior research found that students who participated in spatial experiences within an Earth/Space unit made significant gains on lunar-related concepts. Females tended to lead in significant content development concerning geometric spatial test items. One group of students experienced a purposeful, spatially-integrated Earth/Space unit while the other experienced their Business as Usual (BAU) Earth/Space unit. Differences in spatial-scientific understanding by gender groups and racial/ethnic groups were also investigated within and between BAU and Treatment groups.

The Argument for Developing Spatial Skills in STEM

Research articles in the 1990s have reported a link between students' abilities to report the correct cause of lunar phases with their projective spatial skills (Reynolds, 1990; Wellner, 1995; Bishop, 1996). Other research correlated students' success on science assessments with their spatial ability (Hake, 2002; Sorby, 2006). In addition to this, studies have shown students' improvement in the areas of Chemistry, Geoscience, Physics, and Calculus after they received spatial training (Sanchez, 2012; Miller & Halpern, 2014; Sorby, Casey, Veurink, & Dulaney, 2013).

Recent research has claimed that well-developed spatial thinking is necessary for understanding many astronomical concepts such as celestial motions and lunar phases (Plummer, 2014; Wilhelm, 2009; Wilhelm et al., 2013). Table 1 outlines claims made over the last 25 years linking spatial ability to scientific understanding especially in the area of astronomy.

Table 1. Research that links spatial ability to scientific understanding

Author(s)/Year	**Findings**
Reynolds (1990); Wellner (1995); Bishop (1996)	Students were more likely to report a correct cause of lunar phases when they had strong projective spatial skills.
Pribyl & Bodner (1987); Hake (2002); Sanchez (2012); Sorby (2006); Sorby, Casey, Veurink, & Dulaney (2013); Miller & Halpern (2014)	Students' scores and success on science assessments in the areas of Chemistry, Physics, Geoscience, and Calculus were correlated to their spatial ability.
Black (2005); Plummer (2009, 2014); Plummer, Wasko, & Slagle (2011); Wilhelm (2009); Wilhelm, Jackson, Sullivan, & Wilhelm (2013)	Well-developed spatial thinking is necessary for understanding astronomical concepts such as celestial motions and lunar phases. Spatial thinking includes: Mental rotations, Perspective, Geometric Spatial Visualization, Spatial Projection, Periodic Patterns, and Cardinal Directions.

Black (2005) "hypothesized that mental rotation is the most important in understanding Earth science concepts…humans are handicapped by their single vantage point from Earth of the moving bodies in outer space" (p. 403). Plummer, Wasko, and Slagle (2011) argued that children have difficulties learning to explain daily celestial motion since it requires an understanding across moving frames of references. A mismatch between students' description of apparent motion and their explanation may be due to limited ability to use the necessary spatial abilities to make the logical connection. Instruction may have differentially supported high spatial ability students over low spatial ability (Plummer et al., 2011, p. 1986).

We contend that one cannot understand many astronomical concepts without a developed understanding of four specific spatial domains defined as follows (Wilhelm et al., 2013): a) *Geometric Spatial Visualization (GSV)*-Visualizing the geometric spatial features of a system as it appears above, below, and within the system's plane; b) *Spatial Projection (SP)*-Mentally projecting to a different location on an object and visualizing from that global perspective; c) *Cardinal Directions (CD)*-Distinguishing directions (N, S, E, W) in order to document an object's vector position in space; and d) *Periodic Patterns (PP)*-Recognizing occurrences at regular intervals of time and/or space.

All four of these domains are driven by the facility to mentally rotate objects over time when posed within an astronomical context. For example, the GSV domain concerns visualizing and manipulating the Earth/Moon/Sun system; the SP domain involves mentally maneuvering the sky throughout a day's viewing from various Earthly perspectives; CD domain includes mapping, recording, and predicting lunar positions over time; and PP domain involves noticing the repeated nature of celestial orbital motions.

Gender and Racial Gaps in Spatial Ability

The literature has shown gender differences on students' spatial understandings in favor of males (Kaufman, 2007; Kerns & Berenbaum, 1991; Silverman, Choi, & Peters, 2007; Ansell & Doerr, 2000). Results from the 1996 National Assessment of Educational Progress (NAEP) for United States (US) grade 4 and grade 12 students showed males having significantly higher scale scores than females in the areas of Measurement, and

Geometry and Spatial Sense. "An item-level analysis of percent-correct values revealed some historically common, research-based patterns of difference such as males performing better than females on items that required spatial visualization, the use of measurement tools such as rulers, and working with rational numbers" (Ansell & Doerr, 2000, p. 75). McGraw, Lubienski, and Strutchens (2006) analyzed NAEP results from 1990 – 2003 and found that not only was there a gender gap favoring males in the areas of measurement and geometry, but also that this gap was concentrated at the higher end of score distributions and was most consistent with White students.

Males scored significantly higher than females on tests of spatial visualization as well as 3D mental rotation (Kaufman, 2007). Wilhelm (2009) found that pre-teen female students scored significantly lower than pre-teen male students on spatial pre-tests. However, following a spatially-focused intervention that utilized STEM integrated lessons with many situational opportunities to experience 2D and 3D stimuli, females achieved significantly higher gain scores than their male counterparts. The study speculated that the initial sex differences (on pretests) could be explained by the faster maturation (during preteen years) of the male brain's anatomical regions that handle spatial visual reasoning (Giedd et al. 1999). The implication of the Wilhelm study was that the 2D and 3D instructional intervention allowed females to develop their spatial skills resulting in significant achievement.

In addition to gender differences, research studies have also shown differences in mathematical performance between Black and White students (Lee, 2004; Lee & Wong, 2004; Lubienski, 2002; Reyes & Stanic, 1988) and between Hispanic and White students (Lubienski, 2002). McGraw et al. (2006) analyzed the 2003 NAEP assessment for gender gaps in achievement by race/ethnicity and found "that the differences in scale scores were much greater between racial/ethnic groups than between males and females within the same racial/ethnic group" (p. 140). McGraw et al. (2006) argued that one must examine gender and race/ethnicity as well as social economic status together; otherwise differences within groups will not be documented and interactions will not be found. Despite calls for further research in this area, studies exploring gender and racial/ethnic differences in mathematical performance with potential research-based solutions towards closing the achievement gap have been severely limited.

In order to add to the research base on these issues we examined the following questions: *In what ways will students' curricular and instructional Earth/Space experiences affect their spatial-scientific learning? What, if any, differences in spatial-scientific performance will be observed between gender groups and racial/ethnic groups?*

Methodology

Participants and Instructional Curriculum

Research subjects were sixth-grade students from three US middle schools (Juniper, Butternut, and Willow). Juniper had two Treatment groups (*N* =187) taught by teachers with 4 and 9 years' experience. The Juniper BAU group (*N* = 58) was taught by a teacher with 3 years' experience. Butternut had three Treatment groups (*N* = 228) taught by two first year teachers and one teacher with 11 years of experience. A group of 26 students comprised the Butternut BAU group taught by a teacher with 12 years of experience. Willow had one Treatment group (*N* = 53) taught by a teacher with 13 years' experience. Table 2 displays the teacher and student characteristics. *Pseudonyms were used for all schools; each school self-selected its BAU and Treatment teachers. This, unfortunately, resulted in small BAU numbers (including no BAU classroom at Willow), which was beyond the researchers' control.*

All groups studied Earth/Space concepts related to the Solar System. Treatment teachers employed a spatially-oriented, STEM-integrated Earth/Space curriculum while BAU teachers implemented their regular Earth/Space lessons (see Table 3). The spatially-oriented curriculum (Treatment instruction) was designed to: **(**A) Foster students' understanding of Earth-Space science concepts and 'big ideas' (such as planetary geologic activity and celestial motions and patterns) through the development of innovative projects, lessons, and learning communities; (B) Create experiences for students to do mathematics by challenging them to: i) represent situations graphically and geometrically, ii) observe patterns and functional relationships to make predictions, and iii) develop and employ spatial visualization skills to model phenomena; and (C) Construct opportunities for students to engage in authentic project work, modeling, and data collection and interpretation. The BAU curriculum and instruction tended to utilize videos, simulations, texts, and modeling. Table 3 outlines the time spent on Earth/Space content by each group, the content implemented, and the instructional methods. Juniper teachers executed their Earth/Space units over a nine-week period while Butternut and Willow teachers implemented theirs in approximately four weeks.

Table 2. Teacher and student characteristics

	Teachers	
	Control (BAU) (*n* = 2)	Treatment (*n* = 6)
Gender		
Male	0	1
Female	2	5
Ethnicity		
Caucasian	2	6
Yrs teaching		
Mean	7.50	6.50
Highest degree earned		
BA, BS	0	3
MA, MS	2	3
	Students	
	Control (*n* = 84)	Treatment (*n* = 384)
Gender		
Boys	38	198
Girls	46	186
Grade		
6	84	384
Race/Ethnicity		
Caucasian (Non-Hispanic)	55	244
African American	6	21
Asian American	3	21
Native American	3	10
Hispanic American	5	25
Asian (Not American)	1	21
Other	11	42

Research Questions and Measures

Spatial-scientific reasoning was assessed via pre/post content surveys. The research questions that drove this study were: *In what ways will students' curricular and instructional Earth/Space experiences (Treatment versus BAU) affect their spatial-scientific learning? What, if any, differences in spatial-scientific performance will be observed between gender groups, racial/ethnic groups, and Treatment and BAU groups?*

Due to the small numbers of students comprising groups other than Caucasian, we classified two groups of students: Students of Color (SoC) and White. We acknowledge that analysis at this level has limitations due to the small number of student in these categories. This quasi-experimental study utilized quantitative measures to document students' understanding before and after implementation. The quantitative data sources were the Lunar Phases Concept Inventory (LPCI, Lindell & Olsen, 2002), a multiple-choice survey which assessed eight science domains and four spatial-mathematics domains (*Periodic Patterns (PP), Geometric Spatial Visualization (GSV), Cardinal Direction (CD), Spatial Projection (SP))* as shown in Table 4, and the Purdue Spatial Visualization Test: Rot (PSVT-Rot, Bodner & Guay, 1997), a 20-item multiple choice instrument which assessed mental rotation ability.

Table 3. Time spent on Earth/Space content by each group, content implemented, and instructional method used

Week	Juniper				Butternut and Willow			
	Business as Usual		Treatment		Business as Usual		Treatment	
	Lesson	Method	Lesson	Method	Lesson	Method	Lesson	Method
Week 1	Big Bang Theory; Solar System	PPT Modeling Expanding Universe Balloons	_Overview of Universe_* Why does the Moon appear to change its shape?	"Many Moons" by Thurber, Moon Journaling (5 weeks), Stellarium (planetarium software)	Intro to Solar System	Lecture and note taking	Why does the Moon appear to change its shape? Measuring distance between objects in the sky. Altitude and Azimuth Angles	Moon Journaling (4 weeks) Stellarium (planetarium software) Activity with measurement and graphing
Week 2	Gravity	YouTube video Textbook Centripetal Motion PhET Simulations	How do I measure the distance between objects in the sky? Altitude and Azimuth Angles	Activity with measurement and graphing	Angular measures and measuring the diameter of the Moon; How Far to the Star? Parallax Effect)	Lab work, note taking, and whole class discussion	How can I say where I am on the Earth? Longitude/ Latitude _Rotation/Revolution and Seasons_*	Earth Globe Activity PPT Modeling Activity
Week 3	Stars	Parallax Activity Stellarium	How to say where I am on the Earth. Longitude /Latitude _Rotation/ Rev._	Earth Globe Activity PPT Modeling Activity	Why is Earth the only possible place for life? Seasons Reasons	Lab work using probeware	What can we learn by examining the Moon's surface? Scaling Earth/Moon/ Mars	Exploration of Lunar Images PPT Scaling Activity using Balloons
Week 4	Planets; Earth (day/night)	Foam ball models. Graphing	What can we learn from the Moon's surface?	Exploration of Lunar Images	Moon Phases Eclipses; Tides	Oreo Moon Phases; 3D Earth/ Moon/Sun Activity; Gizmos	Modeling Earth/Moon/ Sun System _Tides_*	PPT 2D and 3D Modeling Activity
Week 5	Seasons	PPT Demos	Scaling Earth/Moon/Mars	PPT Scaling Activity using Balloons				
Week 6	Green House Effect; Water Cycle	Mythbusters Video Book Review	Earth/Moon/Sun System _Tides_*	PPT 2D and 3D Modeling Activity				
Week 7	Moon Phases	Phase Simulations Moonth Activity	What Makes a Planet Geo. Active?	Lab Investigations				
Week 8	Eclipses	PPT	Crater Number Density	Lab Investigations				
Week 9	Projects	Student projects	Experts Lesson on Mars	Video of NASA Expert Scientist;				

* *Not part of the STEM-integrated Treatment curriculum*

Statistical Analysis

This study involved a three-level cross-sectional sample consisting of 468 students (level-1) nested within 8 teachers (level-2) nested within 3 schools (level-3). Note that teachers had either 1 to 3 class periods, but due to the missing data on this variable or students not reporting the correct class period this nested level was not considered and all class periods were collapsed within a teacher. In addition to this, since the race/ethnicity group numbers were small for all non-Caucasian racial groups (as shown in Table 2), it was decided to group the non-Caucasian students into a group category of Students of Color (SoC). Thus, a three-level cross-sectional multilevel model (MLM; Hox, 2010; Raudenbush & Bryk, 2002) was used to examine the effects of pretest score (mean centered), gender (0 = girl, 1 = boy), and race/ethnicity (0 = White, 1 = SoC) (level-1) and Treatment condition (0 = control, 1 = experimental) (level-2) on raw scores. A model-building approach was used to determine the nature and statistical significance of pretest score, gender, race/ethnicity, and Treatment condition on LPCI, each spatial domain that made-up the LPCI (PP, GSV, CD, SP), and PSVT:Rot raw scores. Specifically, a series of multilevel models (MLMs) were specified, estimated, tested, and compared in a hierarchical manner to arrive at the final MLM.

Table 4. LPCI Question Topics and Spatial and Scientific Domains

Question Topics	Spatial Domain	Scientific Domain
A: Time to complete one orbit	Periodic Patterns	Periodicity of Moon's Earthly orbit
B: Time between phases (i.e., time between full and first quarter Moon)	Periodic Patterns	Periodicity of Moon's phases
C: Direction of orbit above the North Pole	Geometric Spatial Visualization; Spatial Projection	Moon's orbit direction around Earth as viewed from space
D: Direction of Moon rise and Moon set	Cardinal Directions	Moon Motion
E: Alignment to produce various phases such as waxing crescent	Geometric Spatial Visualization	Phase and Earth/Moon/Sun positions
F: Time at which various Moon phases rise and set	Cardinal Directions	Phase – sky location - time
G: Explanation of why the Moon's appearance changes over time	Geometric Spatial Visualization	Cause of phases
H: How does the Moon's appearance change when viewed around the world on the same day	Spatial Projection	Effect of lunar phase with change in Earthly location

First, an unconditional (null) model consisting of no predictors was fit to the data. Second, a covariate or main effects only model was fit to the data that consisted of pretest score, gender, race/ethnicity, and treatment condition. Third, a model including the two interactions of primary interest (gender by treatment and race/ethnicity by treatment) were added to the model. To test the difference between nested MLMs, a likelihood ratio test (LRT) or sometimes referred to as deviance difference test was used to test whether each subsequent, larger (i.e. more complicated) model was statistically significantly better than a previous, smaller (i.e. simpler) model. If a model including additional parameters was deemed better fitting than a previous model, it was retained and interpreted. If no difference was found between two subsequent models, then the smaller (reduced) model was retained. If a model including both interactions was deemed better than the main effects only model, then it is known that at least one interaction term was important. To determine which interaction term or both was statistically significantly contributing to the model a backward elimination strategy was used. That is, if the difference in fit between a model without an interaction term versus a model with both interaction terms is nonsignificant, then that interaction term can be eliminated. If the difference in fit is statistically significant, then the interaction term should be retained. The LRTs were based on a full information maximum likelihood estimation method (FIML), while random effects (variances) and fixed effects were estimated using restricted maximum likelihood estimation (REML). Fixed effects were then tested using the convenient Wald test. All statistical significance tests were performed at an alpha level of .05. Hedge's g (corrected for small sample size) was used as an effect size (ES) measure for specific mean comparisons, with MLM coefficient estimates used as the numerator and respective groups posttest variances. All statistical analyses were conducted via SAS version 9.3. In addition to the MLM analysis, we also conducted descriptive statistics to determine gain scores by group for the overall LPCI, each LPCI spatial domain, and the PSVT:Rot. Descriptive results included students by treatment, race/ethnicity, and gender. Including descriptive results allowed us to shed further light regarding how well each student group performed by domain.

Results

Measures

All quantitative assessments were given to both the Treatment and BAU groups immediately prior to and at the conclusion of their Earth/Space unit implementation. Reliability was calculated using the *Cronbach's alpha*; this measures the instrument's internal consistency. The coefficient alpha was calculated for 0.68 and 0.74 for the overall LPCI and the PSVT:Rot assessments, respectively. LPCI and PSVT:Rot values were acceptable. The subset items making up the spatial domains PP (5 test items), GSV (7 test items), SP (4 test items), and CD (5 test items) had coefficient alphas calculated for 0.64, 0.54, 0.41, and 0.17, respectively. The very low alpha for the CD domain illustrates unreliability with these test items; these items have been historically quite difficult for students. For this paper, we will focus on the overall LPCI, the sub-domains PP, GSV, SP, and the PSVT:Rot.

Multilevel Model Results

Table 5 contains the final MLM results for PSVT:Rot, LPCI overall and the LPCI spatial domains (PP, GSV, and SP). We interpreted the MLM results for each outcome as follows. The results for the LPCI overall score showed the best fitting model to the data was the main effects only model. Specifically, LPCI pretest scores, gender, and race/ethnicity were each statistically significant predictors of LPCI posttest scores regardless of treatment condition. That is, higher pretest LPCI scores tended to lead to higher LPCI posttest scores, boys tended to have higher LPCI posttest scores than girls (ES = .18), and Students of Color (SoC) tended to have lower LPCI posttest scores than White students (ES = .23), after adjusting for other student and teacher characteristics.

The results for spatial-mathematics domain PP showed the best fitting model was a model including the interaction term of gender by treatment, which was statistically significant. This interaction term can be understood as meaning that differences in BAU and Treatment groups were dependent on gender of the student, after adjusting for PP pretest scores and student race/ethnicity. Specifically, it could be understood as meaning that gender differences were dependent on treatment condition. That is, boys scored higher than girls in the BAU group (Mean difference = .92, ES = .68), but this gender difference was not maintained in the Treatment group (Mean difference = -0.01, ES = 0). Or, it could be understood to meaning that girls in the Treatment group scored higher than girls in the BAU group (Mean difference = 0.44, ES = .31), while boys in the Treatment group scored lower than boys in the BAU group (Mean difference = -0.49, ES = .36).

In addition, results for PP suggest SoC tended to have lower PP scores than White students (ES = .29), and that higher pretest PP scores tended to lead to higher posttest PP scores, after adjusting for other student and teacher characteristics. Results for domain GSV showed the best fitting model to the data was the main effects only model, which did not include any interaction terms. Results for GSV suggest boys tended to have higher GSV scores than girls (ES = .19) and that higher pretest GSV scores tended to lead to higher posttest GSV scores, after adjusting for other student and teacher characteristics. Results for SP and PSVT-Rot showed the best fitting model to the data was the main effects only model. The only statistically significant predictor of SP and PSVT-Rot posttest scores were scores on the respective pretest.

Descriptive Results

In order to unpack the MLM results, we graphed the gain scores by Treatment and BAU groups for the overall LPCI, the PP, GSV, and SP spatial domain items, and the PSVT:Rot test. Figure 1 displays all Treatment sub-groups (Treatment White and SoC Boys and Treatment White and SoC Girls) to be clustered with similar overall LPCI gain scores (similar clustering can be found for the Treatment sub-groups in Figures 2 - 4 displaying the LPCI spatial domain results). This is not the case for the BAU sub-groups. Within the BAU group, Figure 1 shows the BAU White Boys with the largest gain scores followed by the BAU SoC Boys. BAU White Girls displayed even less gain scores than the BAU Boys, and the BAU SoC Girls showed negative gains. Similar disparaging data is displayed for the BAU group's PP, GSV, and SP spatial domains in Figures 2, 3, and 4, respectively.

Conclusions and Significance

We compared Treatment and BAU groups by LPCI outcomes. Overall LPCI results showed pre-test scores predicted post-test scores, boys performed better than girls, and Whites performed better than Students of Color. We also compared Treatment and BAU groups by LPCI spatial domain outcomes. Domain SP showed no statistically significant differences were observed for gender, race/ethnicity, or treatment type. For domain GSV, it was found that boys, in general, tended to have higher GSV post-test scores. Recall each of the LPCI spatial domains contains mental rotation derivatives. As shown in the review of literature, boys often outperformed girls on mental rotation test items, so it is not surprising that boys, in general, had higher GSV post-scores than girls (Wilhelm, 2009). GSV descriptive results (shown in Figure 3) illustrate White and SoC Treatment students with similar gains, but the same cannot be said for the BAU group where only BAU White boys achieved similar gains to that of the Treatment group.

PP post-scores for the Treatment showed no gender gap. However, boys did outperform girls on PP post-scores within the BAU group. Additionally, Treatment girls scored better than BAU girls on this same domain. Research has shown students (especially females) benefit greatly from situated, project-enhanced learning experiences (Boaler, 2002; Morrow & Morrow, 1995) and this might explain why Treatment girls performed better than BAU girls on the PP domain and why no gender gap was observed within the PP domain for the Treatment group.

Table 5. Final MLM results for predicting LPCI, PP, GSV, SP, and PSVT:rot scores

Parameter	LPCI (n = 462)	PP (n = 462)	GSV (n = 462)	SP (n = 462)	PSVT:Rot (n = 443)
			Fixed effects		
Intercept	7.65***	2.26**	2.99**	1.91**	6.41**
Level-1 (Student)					
Pretest	0.46***	0.41***	0.29***	0.17***	0.65***
Gender	0.61*	0.92***	0.33*	0.14	0.30
Race/Ethnicity	-0.64*	-0.41***	-0.17	0.01	-0.27
Gender by Treatment		-0.93**			
Level-2 (Teacher)					
Treatment	0.69	0.44	0.34	0.19	0.62
			Random effects		
Level-1 (Student)					
Residual variance	8.25***	1.43***	2.19***	1.03***	7.51***
Level-2 (Teacher)					
Intercept variance	0.55	0.06	0.17	< .001	0.18
Level-3 (School)					
Intercept variance	1.10	0.20	0.44	0.17	1.63

Note. LPCI = Lunar Phases Concept Inventory; PP = Periodic Patterns; GSV = Geometric Spatial Visualization; SP = Spatial Projection; PSVT:Rot = Purdue Spatial Visualization Test: Rotations; SoC = Students of Color; Pretest = scores on outcome variable prior to start of study; Gender = girl (0) or boy (1); Ethnicity = white (0) or SoC (1); Treatment = BAU (0) or Treatment condition (1)

* $p < .05$. ** $p < .01$. *** $p < .001$.

Due to limitations of this study (small *N* numbers within BAU groups as well as SoC groups), we can only speculate that the significantly higher scores for the Treatment girls (as compared to the BAU girls) could be due to their project work and the spatially-intensive learning experiences that included daily observations where Treatment students purposefully documented lunar position and appearance while noting patterns and periodicity in journals (Table 3). Although, the only statistically significant predictor of PSVT:Rot posttest scores was the score on the respective pretest, it is interesting to note all groups making similar gain scores except for the BAU girls as shown in Figure 4's descriptive results. In other words, girls in Treatment group performed similarly to boys, but the same cannot be said of BAU girls. Perhaps, there is a way to close the notorious gender gap, after all, when it comes to orchestrating purposeful spatial experiences.

Effect sizes comparing the treatment conditions were estimated for each outcome and ranged from 0.17 (PSVT:Rot) to 0.31 (PP). Although these effect sizes may be small by most standards, they are similar to effect sizes reported elsewhere comparing two groups (McGraw, Lubienski, & Strutchens, 2006). There are obvious limitations to this study in terms of our small BAU numbers. However, our results warrant further studies to examine in more depth how well spatially-oriented, STEM-integrated Earth/Space curricula can advance students' learning, especially for females and students of color.

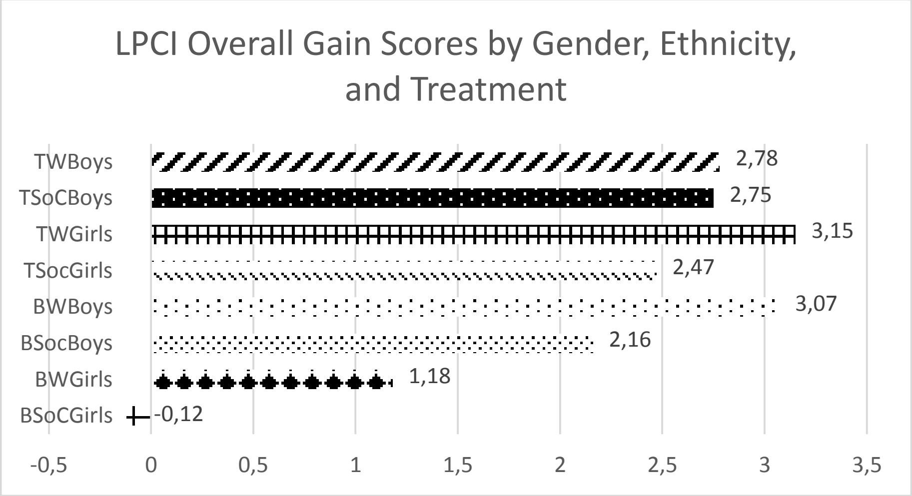


Figure 1. Gain scores by gender [girls/boys], race/ethnicity [white (W)/students of color (SoC)], and treatment [treatment(T)/BAU(B)] for overall LPCI

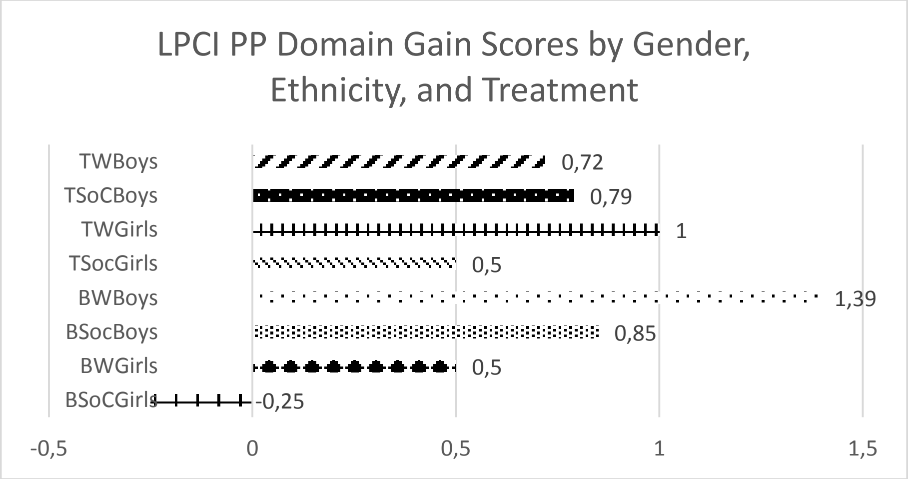


Figure 2. Gain scores by gender [girls/boys], race/ethnicity [white (W)/students of color (SoC)], and treatment [treatment (T)/BAU(B)] for PP domain

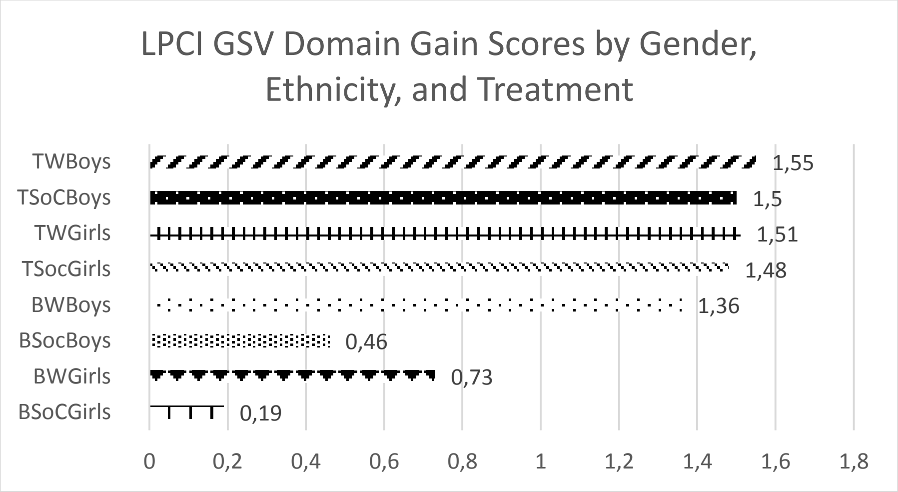


Figure 3. Gain scores by gender [girls/boys], race/ethnicity [white (W)/students of color (SoC)], and treatment [treatment(T)/BAU(B)] for GSV domain

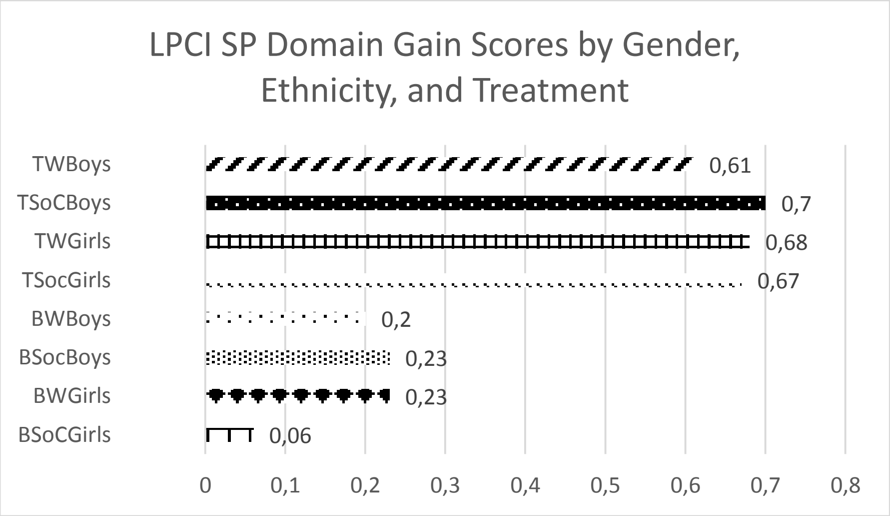


Figure 4. Gain scores by gender [girls/boys], race/ethnicity [white (W)/students of color (SoC)], and treatment [treatment(T)/BAU(B)] for SP domain

The PSVT:Rot gain scores displayed in Figure 5 show all groups making gains between 0.82 and 1.07 except for the BAU girls.

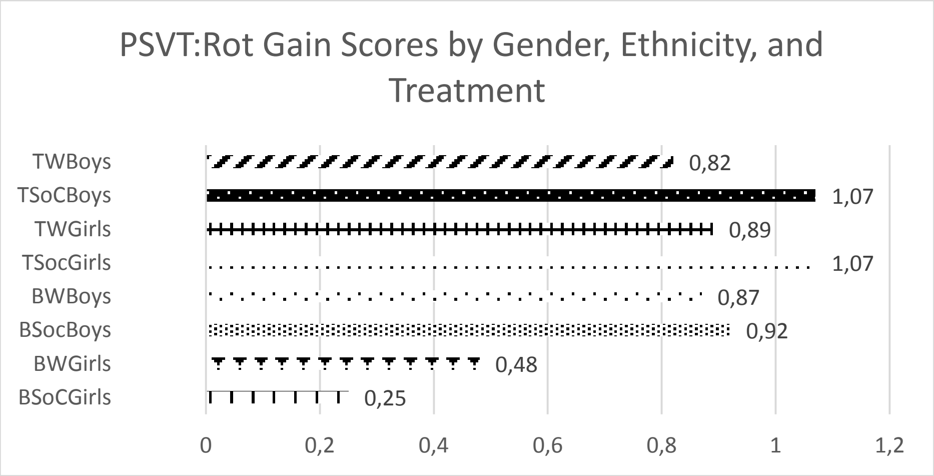


Figure 5. Gain scores by gender [girls/boys], race/ethnicity [white (W)/students of color (SoC)], and treatment [treatment(T)/BAU(B)] for PSVT:rot

The study is unique because it is amongst the first research studies to examine students' spatial-scientific development as they participate in Earth/Space science units. Making the study even more distinctive is discovering how curricular choice and instruction affects student spatial-scientific learning outcomes by gender and race/ethnicity. The authors claimed that one must have well-developed spatial skills in order to understand astronomical phenomena such as lunar phases. Students could come to the classroom already equipped with strong spatial reasoning, ready to understand complicated Earth/Space phenomena; or students will develop the necessary spatial ways of thinking as they make sense of the patterns, geometries, and celestial motions. If we better understand how and which curricular pieces and classroom experiences are instrumental in students' developmental understanding of scientific and spatial content and processes, we can provide more focused interventions to better promote spatial and scientific reasoning with an end effect of better preparedness for *all* students' STEM achievement.

References

Ansel, E., & Doerr, H. (2000). NAEP findings regarding gender: Achievement, affect, and instructional experiences. *Results from the Seventh Mathematics Assessment of the National Assessment of Educational Progress*, 73-106.

Bishop, J. (1996). Astronomy learning and student thinking. *Mercury: Journal of the Astronomical Society of the Pacific, 25*(2). Rocky Mountain Planetarium Association. Retrieved August 1, 2006, from http://www.rmpadomes.org/library/articles/StudentThinking.pdf.

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

Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning* (Volume in the Studies in Mathematical Thinking and Learning Series). Erlbaum Associates, Inc., Mahwah, New Jersey.

Bodner, G. & Guay, R. (1997). The Purdue visualization of rotations test. *The Chemical Educator. 2*(4), 1-17.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., & Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature neuroscience, 2*(10), 861-863.

Hake, R. (2002, August). Relationship of individual student normalized learning gains in mechanics with gender, high-school physics, and pretest scores on mathematics and spatial visualization. *Physics Education Research Conference* (Boise, ID).

Hox, J. J. (2010). Multilevel analysis: Techniques and applications (2nd ed.). New York, NY: Routledge.

Kaufman, S. (2007). Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity? *Intelligence* 35, 211–223.

Kerns, K., & Berenbaum, S. (1991). Sex differences in spatial ability in children. *Behavior Genetics,* 21(4), 383 – 396.

Lee, J. (2004). Multiple facets on inequity in racial and ethnic achievement gaps. *Peabody Journal of Education,* 79(2), 51-73.

Lee, J., & Wong, K. K. (2004). The accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes. *American Educational Research Journal,* 41(4), 797-832.

Lindell, R. & Olsen, J. P. (2002). Developing the Lunar Phases Concept Inventory. *Proceedings of the 2002 Physics Education Research Conference,* Eds. S. Franklin, J. Marx and K. Cummings. PERC Publishing, NY.

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

Lubienski, S. T. (2002). Are we achieving "mathematical power for all?" A decade of national data on instruction and achievement. ERIC, document number ED463166.

McGraw, R. Lubienski, S.T., & Strutchens, M. E. (2006). A closer look at gender in NAEP mathematics achievement and affect data: Intersections with achievement, race/ethnicity, and socioeconomic status. *Journal for Research in Mathematics Education, 37*(2), 129-150.

Miller, D., & Halpern, D. (2014). The new science of cognitive sex differences. *Trends in Cognitive Sciences,* 18*(1),* 37-45.

Morrow, C., and Morrow, J. (1995). Internal influences on gender differences in mathematics. In P. Rogers & G. Kaiser (Eds.), *Equity in mathematics education: Influences of feminism and culture.* London: Falmer Press, 13-26.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Newbury Park, CA: Sage.

Reyes, L. H., & Stanic, G. M. A. (1988). Race, sex, socioeconomic status, and mathematics. *Journal for Research in Mathematics Education*, 19, 26-43.

Reynolds, M. (1990). Two-dimensional versus three-dimensional conceptualization in astronomy education, *Doctoral dissertation*, University of Florida.

Rudmann, D. (2002). Solving astronomy problems can be limited by intuited knowledge, spatial ability, or both. *American Educational Research Association*.

Plummer, J. (2009). A cross-age study of children's knowledge of apparent celestial motion. *International Journal of Science Education, 31*(12), 1571-1605.

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

Plummer, J. (2014). Spatial thinking as the dimension of progress in an astronomy learning progression. *Studies in Science Education*, DOI: 10.1080/03057267.2013.869039.

Pribyl, J., & Bodner, G. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching, 24*(3), 229–240.

Sanchez, C. (2012). Enhancing visuospatial performance through video game training to increase learning in visuospatial science domains. *Psychonometric Bulletin and Review*, *19*(1), 58-65.

Silverman, I., Choi, J., & Peters, M. (2007). The hunter-gatherer theory of sex differences in spatial abilities: Data from 40 countries. *Archives of Sexual Behavior*, *36*(2), 261-268.

Sorby, S. (2006, March). Gender differences in spatial reasoning skills and their effects on success and retention in engineering programs. *Engineers Week.* Retrieved August 1, 2006, from http://www.eweek.org/2002/news/eweek/2006_marathon/buildingskils_2.ppt.

Sorby, S., Casey, B., Veurink, N.,& Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences*, 26, 20-29.

Wellner, K. (1995). A correlational study of seven projective spatial structures with regard to the phases of the moon, *Doctoral dissertation*, University of Iowa.

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.

Author Information

Jennifer Wilhelm
University of Kentucky
Lexington, Kentucky 40506, U.S.A.
Contact e-mail: *jennifer.wilhelm@uky.edu*

Michael Toland
University of Kentucky
Lexington, Kentucky 40506, U.S.A.

Merryn Cole
University of Kentucky
Lexington, Kentucky 40506, U.S.A.

Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

Relationships between Scientific Process Skills and Scientific Creativity: Mediating Role of Nature of Science Knowledge

Gokhan Ozdemir, Ayhan Dikici

Article Info

Article History

Received:
08 July 2016

Accepted:
20 November 2016

Keywords

Scientific process skills
Scientific creativity
Nature of science
Creativity

Abstract

The purpose of this study is to explore the strength of relationships between 7th grade students' Scientific Process Skills (SPS), Nature of Science (NOS) beliefs, and Scientific Creativity (SC) through Structural Equation Modeling (SEM). For this purpose, data were collected from 332 students of two public middle school students in Turkey. SPS, Nature of Scientific Knowledge Scale (NSKS) and SC instruments were used as data collection tools. Zero-order and partial correlation analysis, MANCOVA and ANCOVA analyses were conducted on the data. Two models were hypothesized. In Hypothesized Model 1, it was considered that SPS both directly and indirectly through subscales of NSKS predicted SC; whereas, in Hypothesized Model 2, it was considered that the SPS directly predicted SC. SEM analysis was conducted to test the two hypothesized models. Sobel's *z* tests were conducted to examine the significance of the mediator roles of NSKS's subscales in the relationships between SPS and SC. Results indicated that the relationships between SPS and SC were partially mediated by only the Testable subscale of NSKS; likewise, this relationship was partially mediated by second-order factor NSKS. Finally, we found that the Creative subscale of NSKS has a moderator role on predictive power of the SPS on the SC. Classroom implications obtained from the results are discussed in the paper.

Introduction

Creativity plays an important role in scientific knowledge production process (Hu & Adey, 2002). When scientists undertake research; they use several scientific process skills, creativity and imagination at every phase of their investigations (Hadzigeorgiou, Fokialis, & Kabouropoulou, 2012). Similarly, students use scientific process skills with creative thinking when they are involved in research activities. For example, they develop several possible methods to solve problems. Using scientific process skills such as identifying problem, establishing hypotheses, observing, inferring, choosing method, identifying variables, controlling variables, and making conclusions requires both scientific thinking and creativity in order to develop new methods and solutions to the problems. On the other hand, students' accurate understanding of nature of science (NOS) that accepts the important role of scientific creativity has been appropriately improved through explicit-reflective inquiry activities in which students use their scientific process skills (Carey, Evans, Honda, Jay, & Unger, 1989; Khishfe & Abd-El-Khalick, 2002). Although the relationships are apparent between the two cognitive variables, which are scientific process skills and scientific creativity, the affective variable, and the beliefs about NOS, there has been no research that investigates statistically the strengths of relationships between these three variables. Therefore, the predictive power of scientific process skills on scientific creativity and the potential mediating role of NOS between scientific process skills and scientific creativity are tested in the current study through Structural Equation Modeling (SEM) analyses. Determining the outcomes of this study is important for classroom practices and curriculum developers.

Theoretical Framework and Literature Review

Creativity is usually conceptualized as a skill to produce original and suitable solutions for novel situations and to generate original ideas (Amabile, 1996; Kleibeuker, De Dreu, & Crone, 2013). In the realm of science, original ideas contribute to change and improve our views about the natural world. Therefore, scientific creativity can be interpreted as a skill to foster and raise our understanding of nature (Antink Meyer, & Lederman, 2015). This conceptualization of scientific creativity is the notable property of scientific knowledge because of its developmental nature. At this point, understanding the nature of science (NOS) is to appreciate the role of scientific creativity that is employed in scientific inquiry (Lederman, 1992, 2007; McComas &

Olson, 2002). There is a consensus in the literature that creativity can be improved (Amabile, 1996; Baer & Kaufman, 2006; Cropley, 1992; Kaufman & Beghetto, 2009; Torrance, 1968, 1995). Even a typical academic semester without any special instruction can improve students' scientific creativity (Antink Meyer & Lederman, 2015). Fostering creativity is about teacher behaviors, which includes maintaining an open attitude towards creative ideas or behaviors, showing a humanistic student control, being flexible in thinking and behaving, and valuing independent thinking (Amabile, 1996; Cropley, 1997; Hennessey, 1995; Lubart, 1994; Sternberg & Lubart, 1995). Teachers may encourage students' creativity or cause it to atrophy. Research emphasizes that the relationship between teacher and student is important for students to develop their creativity (Amabile, Hennessey, & Grossman, 1986; Cropley & Cropley, 2009; Erdogdu, 2006; Torrance, 1968, 1995). Beyond the democratic atmosphere in which teachers provide for their students, scientific creativity is also developed through using scientific process skills and accurate understanding of NOS (Hadzigeorgiou et al., 2012). On the other hand, the appropriate teaching for accurate understanding of NOS is achieved through explicit-reflective inquiry activities in which students employ their scientific process skills. The relational nature of these three components, the SPS, NOS and SC, demand investigation in order to uncover the explanatory power of one to another. These relations will be discussed in the sections that follow.

Scientific Creativity and Scientific Process Skills

The typical properties of creativity are creative imagination and creative thinking (Runco, Nemiro, & Walberg, 1998). If we look over the creative thinking models (e.g. Isaksen & Treffinger, 1985; Osborn, 1953), we can see that several dimensions of creative thinking models intersect with those of SPS. For example, finding and/or solving problems, making predictions, designing experiments, seeking solutions are the common properties of SPS and scientific thinking. Meador (2003) presented that all SPS dimensions (Basic science process skills: Observing, comparing, classifying, measuring, and communicating; Intermediate process skills: Inferring and predicting; Advanced process skills: Hypothesizing, defining and controlling variables) correspond to creative thinking components except "measuring" because measuring does not require creativity. Therefore, it can be easily claimed that SPS elements has explanatory power of scientific creativity.

We have to note that the literature on creativity has a common claim that creativity requires domain-specific knowledge and developed skills (Alexander, 1992; Amabile, 1996; Baer, 1991, 2016; Han, 2003; Kaufman & Baer, 2008). Moreover, divergent thinking, which is one of the curricula components of creativity, depends on domain-specific relevance (Barron & Harrington, 1981). Science is a vital domain in which students learn scientific concepts, theories, and laws as well as develop the cognitive skills necessary for scientific creativity. For example, in a simple problem-based learning activity, a student perceives the problem, explores what is known to develop different strategies for the solution in which combines his knowledge with his imagination and skills to produce new knowledge. All these complex cognitive scientific process aid and trigger scientific creativity. At this point, a student's scientific process skills such as identifying problem, establishing hypotheses, observing, inferring, choosing method, identifying variables, controlling variables, and making conclusions should explain the level of scientific creativity of the student. Some research found this relationship between SPS and SC. For example, Aktamıs and Ergin (2007) found that 7th grade Turkish students' SPS scores positively correlate with those of SC scores. In their subsequent study (Aktamıs & Ergin, 2008); they determined that 7th grade students' scientific creativities increased when they were exposed to scientific process skills education for a period of 12 weeks.

Scientific Creativity and Beliefs on Nature of Science

The beliefs on science and scientific knowledge, the understanding about nature of science (NOS), determine how an individual abstracts, structures and manipulates information received from the world around us. For example, a student who has the positivist views of science believes that scientific knowledge consists of absolute truths and everyone reaches the same truths by using the same step-wise methodological procedures (e.g. Edmondson & Novak, 1993; Ryan & Aikenhead, 1992). This understanding ignores imagination and creativity in scientific knowledge construction. This rigid understanding about NOS limits the student's imagination and scientific creativity when involved in a scientific problem activity. However, reformist science curriculums aim to educate students as scientifically literate individuals who have the relativist view of NOS. Accurate and relativist understanding of NOS is a pre-requisite to scientific literacy. A student who has the relativist views of science believes that science is a human endeavor, in which imagination and scientific creativity play a vital role in scientific knowledge production (e.g. Edmondson & Novak, 1993; Ryan & Aikenhead, 1992). For these reasons, scientific knowledge that may change in time is improved with divergent ideas and different methodological procedures. Indeed, the same data can be interpreted differently and all interpretations can be scientifically valid. As a result, one can claim that beliefs about NOS predict scientific creativity.

On the other hand, there are some studies to improve students' beliefs about the nature of science and scientific knowledge through making use of scientific process skills (e.g. Khishfe & Abd-El-Khalick, 2002). The research on this domain indicated that explicit-reflective scientific inquiry activities, where students use scientific process skills, improve their understanding on the nature of science (e.g. Lederman & Abd-El-Khalick, 1998). Some studies support this claim. For example, Ren, Li, Zhang, and Wang (2012), focused on students' creative imagination, found that Chinese students' creative imagination improved when involved in science-related competitions and visits to science-related places. It cannot be disregarded that these science-related activities most probably help students to learn NOS and to develop scientific process skills which support their creative imagination. Another study indicated that students who are successful in science have high scores in creative thinking (Ren et al., 2012). Similarly, in a Turkish sample, both general and scientific creativity scores of 6-8th grade students positively correlated to their academic achievement on a science and technology course (Ayverdi, Asker, Öz Aydın, & Sarıtaş, 2012). Another study focused on creative scientific problem finding (Hu, Shi, Han, Wang, & Adey, 2010), and found students' abilities on this aspect have a developmental trend up until high school. Taking these facts into account, one can claim that beliefs about NOS would have a mediating role between scientific process skills and scientific creativity.

Research Problems

This study is focused on two major questions:
1. To what extent, if any, do middle school students' Scientific Process Skills (SPS) have predictive power to those of Scientific Creativity (SC)?
2. To what extent, if any, do middle school students' Nature of Science (NOS) knowledge have mediating roles in the relationship between those of SPS and SC?

Hypotheses

It is predicted that (a) students' SPS scores will have a direct effect on their SC scores, and (b) students' NOS scores will have a direct effect on SC scores. Finally, (c) students' NOS scores will be significant mediators in the relationship between their SPS scores and SC scores.

Hypothesized Model

Many studies in the literature found that scientific process skills are a predictor on scientific creativity (e.g. Aktamıs & Ergin, 2008). Also, some studies indicated the relationship between NOS and SC (e.g. Antink Meyer & Lederman, 2015). Others showed that how NOS understanding can be improved by science inquiry where students use their SPS (e.g. Lederman, 2004). Consequently, two models are proposed due to the results reported in previous studies in the relationship between scientific process skills (SPS), Nature of Science (NOS), and scientific creativity (SC). To examine these relationships, two path models were proposed. In the Hypothesized Model 1, it was considered that the SPS, both directly and indirectly through NOS's dimensions, predicted SC. In the Hypothesized Model 2, it was considered that the SPS directly predicted the SC. Figure 1 presents the model of these hypothesized structural relationships.

Method

Participants

The participants included 332 students from the 7th grade of two middle schools in Niğde city, Turkey. These schools are typical public middle schools that were chosen at random. 52.1% (f=173) of the students are female and 47.9% (f=159) are male. Students' ages varied between 13 and 14 years old. The arithmetic mean of the ages was 13.14 years and the standard deviation of the ages was .35.

Instruments

Three instruments were used to assess students’ SPS, NOS understanding, and SC. First, the students’ SPS were measured through the Scientific Process Skills Test (SPST) (Okey, Wise, & Burns, 1982). Second, students’ NOS understanding were measured through the Nature of Scientific Knowledge Scale (NSKS) (Rubba & Andersen, 1978). Lastly, students’ SC were measured through the Scientific Creativity (SC) test (Hu & Adey, 2002). The students also provided basic demographic information.

Scientific Process Skills Test

The Scientific Process Skills Test (SPST), originally developed by Okey, Wise, and Burns (1982), was adapted to Turkish by Geban, Aşkar, and Özkan (1992). The Cronbach’s alpha coefficient of the adapted version of the instrument is .81. For the current sample, Cronbach's alpha coefficient was found to be .74, and Kuder–Richardson Formula 20 (KR-20) was found as .78. The Test consists of 36 multiple-choice items. The scale has five sub-dimensions that are defining variables (12 questions), operational defining (6 questions), formulating hypotheses (9 questions), data analyses (6 questions), and research design (3 questions).

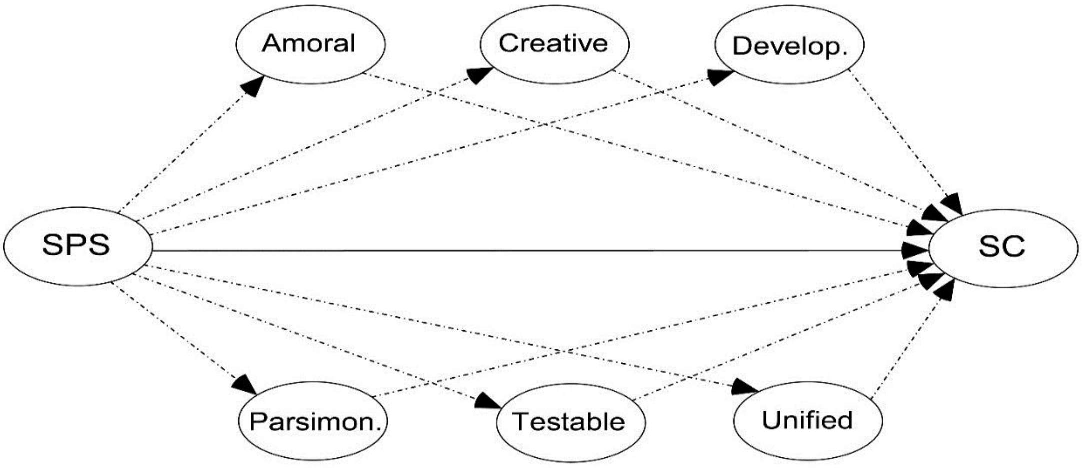


Figure 1. Hypothesized models

Note: Solid and dashed lines represent that Scientific Process Skills (SPS) both directly and indirectly through Nature of Scientific Knowledge subscales (Amoral, Creative, Developmental, Parsimonious, Testable, and Unified) predicted Scientific Creativity (SC) (Model 1), whereas only solid line represents the SPS directly predicted SC (Model 2). SPS = Scientific Process Skills, SC = Scientific Creativity.

Nature of Scientific Knowledge Scale

Originally developed by Rubba and Andersen (1978), the Nature of Scientific Knowledge Scale (NSKS), adapted to Turkish by Kılıç, Sungur, Çakıroğlu, and Tekkaya (2005) with a .74 Cronbach’s alpha coefficient, has a six-factor structure. For the current sample, after “if item deleted” option was applied to the data in SPSS, the Cronbach’s alpha coefficient was found to be .68. Factors are Amoral, Creative, Developmental, Parsimonious, Testable, and Unified respectively. Each factor consists of eight items; with four positive and four negative items. The item statements were each attached to a five-point Likert-type scale, labeled “strongly agree”, “agree”, “neutral”, “disagree”, and “strongly disagree”; randomly arranged as a tryout instrument. The following are sample positive and negative items from the Amoral subscale: “The applications of scientific knowledge can be judged good or bad; but the knowledge itself cannot”, and “Certain pieces of scientific knowledge are good and others are bad”. Creative subscale: “A scientific theory is similar to a work of art in that they both express creativity”, and “Scientific laws, theories, and concepts do not express creativity”. Developmental subscale: “We accept scientific knowledge even though it may contain errors”, and “The truth of scientific knowledge is beyond doubt”. Parsimonious subscale: “Scientific knowledge is stated as simply as possible”, and “Scientific laws, theories, and concepts are not stated as simply as possible”. Testable subscale: “A piece of scientific knowledge will be accepted if the evidence can be obtained by other investigators working under similar conditions", and "Scientific knowledge need not be capable of experimental test". Unified subscales: "The laws, theories, and concepts of biology, chemistry, and physics are related", and "The laws, theories, and concepts of biology, chemistry, and physics are not linked".

Scientific Creativity Test

Developed by Hu and Adey (2002), the Scientific Structure Creativity Model (SSCM) was adapted to Turkish by Deniş Çeliker and Balım (2012), with a .86 Cronbach's alpha coefficient. Cronbach's alpha coefficient was measured to be .81 for the current sample. The Scientific Structure Creativity Model (SSCM) of Hu and Adey (2002) has three dynamic dimensions; product, trait, and process. The Product dimension contains technical product, science knowledge, science phenomena, and science problem; while the Trait dimension contains fluency, flexibility, and originality; and the Process dimension contains thinking and imagination. The following questions were designed by Hu and Adey (2002) in order to measures the combination of attributes in the SSCM.

1. Please write down as many as possible scientific uses as you can for a piece of glass.
2. If you can take a spaceship to travel in the outer space and go to a planet, what scientific questions do you want to research? Please list as many as you can.
3. Please think up as many possible improvements as you can to a regular bicycle, making it more interesting, more useful and more beautiful.
4. Suppose there was no gravity, describe what the world would be like?
5. Please use as many possible methods as you can to divide a square into four equal pieces (same shape).
6. There are two kinds of napkins. How can you test which is better? Please write down as many possible methods as you can and the instruments, principles and simple procedure.
7. Please design an apple picking machine. Draw a picture, point out the name and function of each part.

These SC questions measures more than one dimension of the SSCM. Each dimension is associated with different sub attributes among the 24 attribute combinations. For example, Item 1, "Please write down as many as possible scientific uses as you can for a piece of glass" measures all three dimensions. This question forces students to plan and carry out a scientific investigation. Therefore, this task is associated with: (1) science knowledge in Product dimension; (2) fluency, flexibility, and originality in the Trait dimension; and (3) thinking in the Process dimension.

In scoring, each student's responses, for questions 1, 2, 3, and 4, fluency scores were counted according to the number of responses, regardless of their qualities. For all questions, flexibility scores were counted in terms of the number of response categories (number of designs for Question 6 and number of components for Question 7). For Questions 1, 2, 3, 4, and 6, originality scores were counted in terms of rarity of the responses. If the probability of each response was less than 5%, 2 points were given. If the probability of each response was between 5% and 10%, 1 point was given. If the probability of each response was greater than 10%, it was rated as 0. The originality score for Question 7 was rated out of 5 in terms of the rarity of the each response. Question 5 was omitted in scoring because 12 students recognized that a square can be divided into four equal pieces an infinite numbers of time. So, scoring was deemed impossible.

Two researchers who are authors of this study independently rated students' responses for each question and counted total SC scores for each student. Then, the two scoring results were compared. The score consistency between the raters was calculated as 98%. A few discrepancies between the raters in terms of scoring of the responses were resolved mutually through discussion.

Procedure

All instruments were applied to students by the researchers in accordance with the principles of volunteering. Legal permission was obtained from the Niğde Provincial Directorate of National Education on March 01, 2016. Implementation took place in March 2016. The instruments were administered to students in three different seasons. In the first day, the administration process lasted 45-50 minutes for the SPST. On the second day, students completed the NSKS in 30-35 minutes, and on the third day, students completed the SC test in 50-60 minutes.

Data Analysis

Using the maximum likelihood method of estimation from AMOS 18, three separate Confirmatory Factor Analyses (CFA) were conducted in order to check whether or not the factor structures of the scales would be confirmed in the present sample. In the first CFA, the SPS model, in which five first-order factors (i.e., defining variables, operational defining, formulating hypotheses, data analyses, and research design) were predicted by second-order factor (i.e., SPS). In the second CFA, the NSKS model, in which six first-order factors (i.e., amoral, creative, developmental, parsimonious, testable, and unified) were predicted by a second-order factor (i.e., NSKS). In the third CFA, the single-factor SC model with three indicators (i.e., fluency, flexibility, and originality) was tested.

MANCOVA and ANCOVA analyses were conducted in order to see whether or not there was a significant effect of demographic variables on the scales. Zero-order and partial correlation analysis were conducted in order to see whether or not there were significant relationships among the variables. These analyses were computed using SPSS 18 software. Structural Equation Modeling (SEM) analysis was conducted in order to examine the mediating role of the NSKS subscales with a robust method (Iacobucci, Saldanha, & Deng, 2007; Kim & Bentler, 2006; Preacher & Hayes, 2004). SEM is a statistical technique for estimating and testing hypothesized causal relationships among latent and/or observed variables (Bryne, 2010). Two models were tested in the SEM analysis using AMOS 18. In the first model, SPS both directly and indirectly through NSKS subscales predicted SC; whereas, in the second model, SPS directly predicted SC. The evaluation of model adequacy was based on the minimum value of the discrepancy function (CMIN/x^2), Root Mean Square Residual (RMR), Root Means Square Error of Approximation (RMSEA), Normed Fit Index (NFI), Comparative Fit Index (CFI), Goodness-of-Fit Index (GFI), and Adjusted Goodness-of-Fit Index (AGFI), and its lower and upper confidence interval boundaries (Byrne, 2010; Hair, Black, Babin, & Anderson, 2010; Hu & Bentler, 1995; Schermelleh-Engel, Moosbrugger, & Müller, 2003; Schumacker & Lomax, 2004). All of the items' scores were changed to standard scores. Outliers were not found to be less than -3.0 or higher than +3.0 (Bakeman & Robinson, 2014). Additionally, linearity, multicollinearity, and singularity assumptions for SEM analysis were met. For Hypothesized Model 1, the multivariate kurtosis = 2.18 and critical ratio = .94 (kurtosis for NSKS subscales, amoral = .51, creative = .50, developmental = .68, parsimonious = .55, testable = .04, unified = .26 in absolute value; for SPS subscales, defining variables = .47, operational defining = .85, formulating hypotheses = .83, data analyses = .56, and research design = .77; and kurtosis for SC = .70) indicated that the data distributions were close to normal. If critical ratio values are higher than 5.00, the data are considered non-normally distributed (Bentler, 2005). If the absolute values of the kurtosis index are higher than 10.0, they suggest a problem. If these values are higher than 20.0, then they suggest an extreme problem (DeCarlo, 1997; Kline, 2005). These problems were not observed in our data. Mahalanobisd^2 ranged from 15.94 to 40.77 (p>.001). Because of the large p-values of the Mahalanobis d^2, none of the observations under the assumption of normality should be treated as outliers. With these findings, the data were included in the hypothesized model. Thus, maximum likelihood estimation in the SEM analysis was performed for this study. The model was tested that the SPS both directly and indirectly through the NSKS subscales predicted SC. On the basis of the unstandardized beta coefficients and standard error rates, the Sobel's z tests (Sobel, 1982) were conducted to examine whether or not the mediation models were significant (Baron & Kenny, 1986; Kim & Bentler, 2006; Preacher & Hayes, 2004).

Results

Preliminary Analyses

The Factor Structure of the SPS

Both the first-order factor model and the second-order factor model were tested through CFA in order to validate the factor structures of the SPS. In the first model, the first-order factors (i.e., defining variables, operational defining, formulating hypotheses, data analyses, research design) were predicted by a second-order latent factor (i.e., SPS); whereas, in the second model, the first-order factors were freely estimated regardless of the effect of second-order factor. The results of the CFA demonstrated very good fit for the second order-factor model ($x^2_{(2)}$=1.63; x^2/df=.82; RMR=.02; RMSEA=.01; NFI=.99; CFI=1.00; GFI=.99; AGFI=.98); whereas, the first order-factor model demonstrated a bad fit ($x^2_{(514)}$=1379.71; x^2/df=2.68; RMR=.01; RMSEA=.07; NFI=.58; CFI=.68; GFI=.82; AGFI=.77). Due to the scope of this study, the second order-factor SPS was considered in the present study. Standardized parameter estimations of the second order-factor model ranged from .37 to .69, indicating that the items in the SPS were significantly predicted by their latent variable (all *p*s<.001). Notably, SPS strongly predicted defining variables (β=.37), operational defining (β=.69), formulating hypotheses (β=.62), data analyses (β=.51), and research design (β=.52). For the single factor's internal consistency, Cronbach's alpha coefficient (α) was found to be .74, and Kuder–Richardson Formula 20 (KR-20) was found as .78 for the all multiple-choice items. Item difficulty values ranged from .32 to .71.

The Factor Structure of the NSKS

First, internal consistency coefficients of the NSKS the subscales were analyzed. Cronbach's alpha coefficients were found to be .18 for Amoral, .62 for Creative, .28 for Developmental, .04 for Parsimonious, .57 for Testable, and .64 for Unified. Due to the initial low internal consistency, some items were deleted to re-calculate the consistency considering that the number of items in one subscale must be at least three (Comrey, 1988). In the results of the re-analyses, the internal consistency coefficients of the scales ranged from .63 to .77 (i.e., .63 for amoral, .77 for creative, .65 for developmental, .65 for parsimonious, .70 for testable, and .73 for unified) and the whole scale reliability was computed as .68. Later, both the first-order factor model and the second-order factor model were tested through CFA in order to validate the factor structure of the NSKS. In the first model, the first-order factors (i.e., amoral, creative, developmental, parsimonious, testable, and unified) were predicted by a second-order latent factor (i.e., NSKS); whereas, in the second model, the first-order factors were freely estimated, regardless of the effect of second-order factor. The results of the CFA demonstrated that the first-order factor model ($x^2_{(90)}$=185.47; x^2/df=2.06; RMR=.10; RMSEA=.05; NFI=.86; CFI=.92; GFI=.94; AGFI=.89) fit the data significantly better than the second-order factor model ($x^2_{(7)}$=22.05; x^2/df=3.01; RMR=.46; RMSEA=.08; NFI=.87; CFI=.90; GFI=.98; AGFI=.94). Both because of these reasons and the scope of this study, both first-order factor (i.e., amoral, creative, developmental, parsimonious, testable, and unified) and second-order factor (i.e., NSKS) were considered in the present study. Standardized parameter estimations ranged from .61 to .80, indicating that the items in the NSKS were significantly predicted by their latent variables (all *p*s< .001). Notably, NSKS (i.e., second-order factor) significantly predicted amoral (β=-.53, *p*< .001), creative (β=.23, *p*< .01), developmental (β=.35, *p*< .001), testable (β=.79, *p*< .001), and unified (β=.79, *p*< .001) except for parsimonious (β=.02, *p*> .05) (i.e., first-order factors).

The Factor Structure of the SC

The results of the CFA revealed that the one-factor SC model with three attributes (total scores of fluency, flexibility, and originality) had very good fit to data ($x^2_{(27)}$=0.0; x^2/df=0.0; RMR=.00; RMSEA=.07; NFI=1.00; CFI=1.00; GFI=1.00; AGFI=1.00). The fluency, flexibility, and originality parameter estimations were .75, 79, and.80 respectively. This signifies that SC was considerably predicted by their three attributes (all *p*s< .001). Finally, the Cronbach’s alpha coefficient was computed as .81.

Effects of Demographic Variables on SPS, SC, and NSKS Subscales

The multivariate effect of age on the NSKS subscales was not significant. MANCOVA results demonstrated that the effects of age (η_p^2=.04) on the NSKS subscales were negligible. With η_p^2 coefficients ranging from .00 to .02, the results of the univariate analyses did not replicate the results of MANCOVA. The multivariate effect of gender on the NSKS subscales was significant, but negligible (η_p^2=.05). With η_p^2 coefficients ranging from .00 to .02, the results of the univariate analyses did replicate the results of MANCOVA. Accordingly, male students (M=8.83, SD=.26) were statistically more parsimonious thinking than their female counterparts (M=7.49, SD=.49); whereas, female students (M=11.44, SD=.49) were statistically more unified thinking than their male counterparts (M=10.07, SD=.26). For the SPS, ANCOVA results revealed that the effects of age (η_p^2=02) and gender (η_p^2=01) were trivial. Finally, for the SC, ANCOVA results indicated that the effects of age (η_p^2=01) and gender (η_p^2=.02) were unimportant. Although the effects of demographic variables on SPS, NSKS subscales, and SC were not of interest to the present study, both a partial correlation analysis, through which the demographic variables effects were controlled for, and a zero-order correlation analysis were conducted in order to see whether or not the relationships among the variables at hand significantly changed due to the possible effects of demographic variables. The results of the partial correlation and zero-order correlation analyses are presented in Table 1. As given in Table 1, zero-order correlation coefficients were highly similar to the partial correlation coefficients, indicating that to include demographic variables in the analysis did not significantly alter the general view regarding the relationships among SPS, Creative, Developmental, Testable, and Unified subscales of NSKS, and SC. Amoral and Parsimonious did not significant relationships between the SPS and SC. Besides, relationships among Amoral, Developmental, and Unified, and between Creative and Parsimonious were negatively significant. Importantly, all significant correlation coefficients ranged in magnitude from small (r= .12) to moderate (r= .50) (see Table 1).

Table 1. Zero-order and partial correlation coefficients[a] (N=332)

Variable	M	SD	1	2	3	4	5	6	7	8
1.SPS	14.86	5.22	--	*-.06*	*.13**	*.12**	*-.01*	*.30****	*.26****	*.50****
2.Amoral	8.96	2.83	-.05	--	*.06*	*-.23****	*-.01*	*-.10*	*-.32****	*.04*
3.Creative	10.67	3.23	.14**	.07	--	*.03*	*-.17**	*.18***	*.06*	*.17***
4.Developmental	10.57	2.69	.12*	-.22***	.03	--	*-.07*	*.26****	*.22****	*.25****
5.Parsimonious	8.41	2.89	-.01	-.02	-.17**	-.06	--	*.05*	*.10*	*.02*
6.Testable	10.92	3.03	.31***	-.08	.21***	.26***	.03	--	*.39****	*.38****
7.Unified	10.84	3.02	.26***	-.27***	.08	.23***	.09	.42***	--	*.18***
8.SC	47.59	17.50	.49***	.04	.17**	.25***	.02	.38***	.19***	--

$^{***}p< .001$, $^{**}p<.01$, $^{*}p<.05$

Note: [a] Zero-order correlations are below the diagonal, partial correlations are above the diagonal and italics after the controlled for the demographic variables (i.e., gender and age). SPS= Scientific Process Skills. SC= Scientific Creativity. Amoral, Creative, Developmental, Parsimonious, Testable, and Unified are NSKS's subscales.

These considerable relationships between the SPS, the NSKS's four subscales (i.e., Creative, Developmental, Testable, and Unified), and SC suggest that it is reasonable to examine the mediating roles of NSKS subscales (Creative, Developmental, Testable, and Unified) in relationships between the SPS and SC (Baron & Kenny, 1986). Although the Amoral and Parsimonious subscales did not play a significant role in the relationships between the SPS and SC, we included the Amoral and Parsimonious subscales in SEM analysis as well, which is a robust method in order to see their parameter estimations.

The Mediating Role of the NSKS Subscales

SEM Analysis

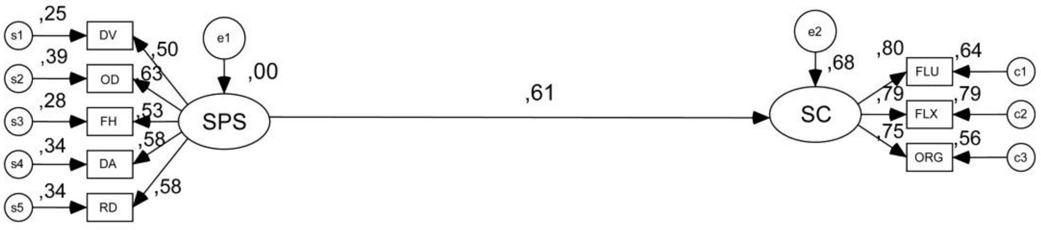
The results of the SEM analysis revealed that Hypothesized Model 2, in which the SPS directly predicted SC ($x^2_{(16)}$=33.14; x^2/df=2.07; RMR=.29; RMSEA=.05; NFI=.97; CFI=.98; GFI=.98; AGFI=.95), fit the data better than Hypothesized Model 1, in which the SPS both directly and indirectly through NSKS subscales (i.e., Amoral, Creative, Developmental, Parsimonious, Testable, and Unified) predicted the SC ($x^2_{(284)}$=1239.96; x^2/df=4.36; RMR=.39; RMSEA=.10; NFI=.58; CFI=.64; GFI=.78; AGFI=.73). Parameter estimations of the Hypothesized Model 2 is given in Figure 2, and the parameter estimates for Hypothesized Model 1 are given in Figure 3. Due to the poor fit of Hypothesized Model 1, NSKS's subscales were individually included in six separate SEM analyses. Parameter estimates derived from these analyses are shown in Figure 4.Goodness-of-fit measures obtained from these analyses were not as good as the fit for Hypothesized Model 2. Goodness-of-fit measures of these analyses are shown in Table 2. Also, Table 2 presents Goodness-of-fit measures of the second order factor NSKS (i.e., one factor NSKS) in the SEM analysis.

Table 2. Goodness-of-fit measures of the hypothesized models

Goodness of Fit Index	First Model NSKS Subscales first order factor model						Second-order NSKS	Second Model
	Amoral	Creative	Develop.	Parsimon.	Testable	Unified		
X^2/sd	3.53	3.44	2.70	4.24	2.99	3.71	4.06	2.07
RMR	.43	.37	.30	.62	.27	.41	.94	.29
RMSEA	.08	.08	.07	.10	.08	.09	.09	.05
NFI	.88	.88	.91	.86	.91	.89	.81	.97
CFI	.91	.91	.94	.88	.94	.91	.85	.98
GFI	.93	.93	.95	.53	.94	.93	.90	.98
AGFI	.88	.89	.91	.44	.90	.88	.84	.95

Note: First Model: The Scientific Process Skills both directly and indirectly through NSKS's subscales (i.e., Amoral, Creative, Developmental, Parsimonious, Testable, and Unified) predicted the Scientific Creativity.

Second Model: The Scientific Process Skills directly predicted on Scientific Creativity. Second-order factor NSKS: The Scientific Process Skills both directly and indirectly through NSKS (i.e., second-order factor) predicted the Scientific Creativity.



p < .001

Figure 2. Direct effect between the SPS and SC

Note: SPS: Scientific Process Skills. SC: Scientific Creativity. Estimate: 4.75, S.E.: .65; C.R.: 7.23, β: .61 SPS: Scientific Process Skills. DV: Defining Variables, OD: Operational Defining, FH: Formulating Hypotheses, DA: Data Analyses, and RD: Research Design. SC: Scientific Creativity. FLU: Fluency, FLX: Flexibility, ORG: Originality.

As seen in Figure 2, the standardized direct effects of SPS on SC were statistically significant (β =.61, *p*<.001). When NSKS subscales (i.e., Amoral, Creative, Developmental, Parsimonious, Testable, and Unified) were all together included in SEM analysis of Hypothesized Model 1, the parameter estimation, which represented the indirect effects of SPS on SC through the NSKS subscales(β=.71,*p*< .001), was still significant and increased. While the effects of four NSKS subscales, Creative, Developmental, Parsimonious, and Testable, on the SC were not statistically significant, Amoral (β=.16, *p*< .05) and Unified (β=-.24, *p*< .05) subscales were significant. However, Sobel (1982) tests did not confirm the mediating roles of Amoral (*z*=1.62, *p*>.05) and Unified (*z*=1.75, *p*>.05) in relationship between SPS and SC. This means that the relationships between SPS and SC were not mediated by NSKS subscales (see Table 3).

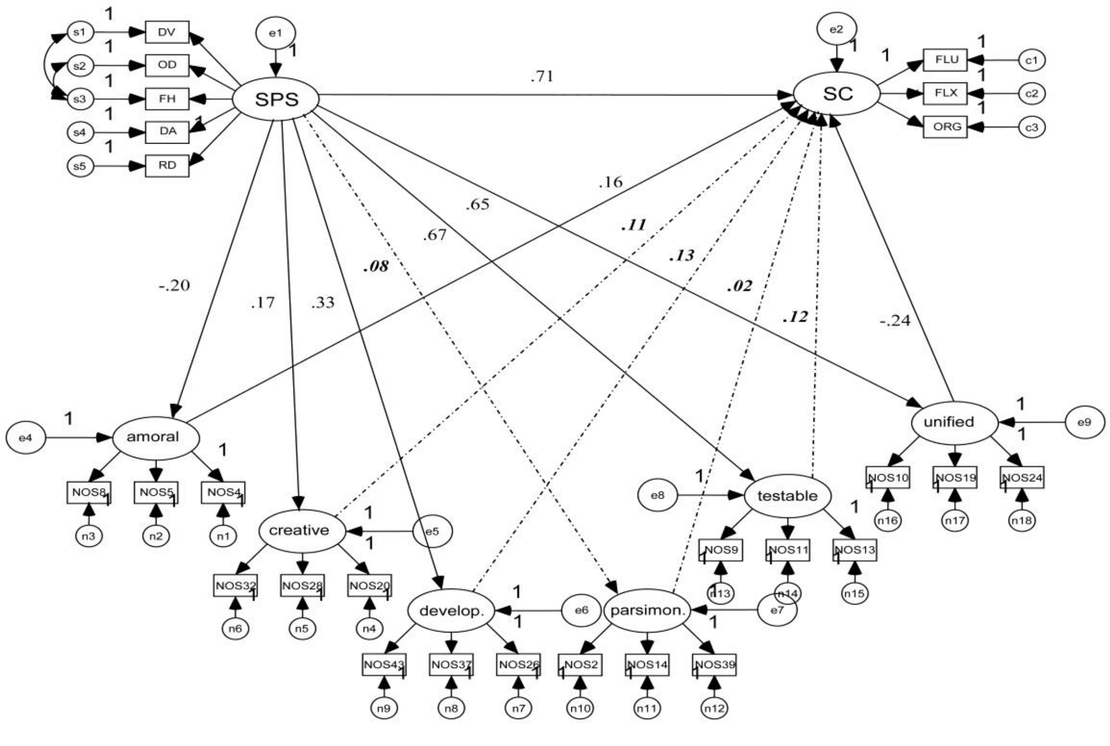


Figure 3. The mediating role of the NSKS subscales

Note: Parameter estimations are standardized values. Bold/italic coefficients and dashed lines are not statistically significant (*p*>.05). SPS: Scientific Process Skills. DV: Defining Variables, OD: Operational Defining, FH: Formulating Hypotheses, DA: Data Analyses, and RD: Research Design. SC: Scientific Creativity. FLU: Fluency, FLX: Flexibility, ORG: Originality. Amoral, Creative, Developmental, Parsimonious, Testable, and Unified are NSKS's subscales.

Table 3. Summary of NSKS subscales in SEM analysis

Dependent Variables		Independent Variables	Estimate	S.E.	C.R.	β	p	Sobel's *z*
amoral	<---	SPS	-.28	.12	-2.31	-.20	.021	
creative	<---	SPS	.32	.17	1.91	.17	.056	
develop.	<---	SPS	.43	.13	3.15	.33	.002	
parsimon.	<---	SPS	.03	.01	2.24	.08	.025	
testable	<---	SPS	1.30	.27	4.82	.67	***	
unified	<---	SPS	1.32	.33	4.03	.65	***	
SC	<---	SPS	6.32	1.69	3.72	.71	***	
SC	<---	amoral	1.04	.45	2.28	.16	.022	1.62
SC	<---	creative	.51	.29	1.78	.11	.075	
SC	<---	develop.	.85	.52	1.63	.12	.103	
SC	<---	parsimon.	.36	.40	.89	.02	.374	
SC	<---	testable	.54	.52	1.04	.12	.295	
SC	<---	unified	-1.05	.53	-1.96	-.24	.049	1.75

$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$

Note: SEM analysis table belongs to the Hypothesized Model 1. SPS: Scientific Process Skills. SC: Scientific Creativity. Amoral, Creative, Developmental, Parsimonious, Testable, and Unified are NSKS's subscales.

For each of the NSKS's subscales, six separate SEM analyses were performed as to whether or not any subscale individually mediated in the predictive power of SPS on SC. The changes of the parameter estimates for each subscale are indicated on the solid line between the SPS and SC in Figure 4.

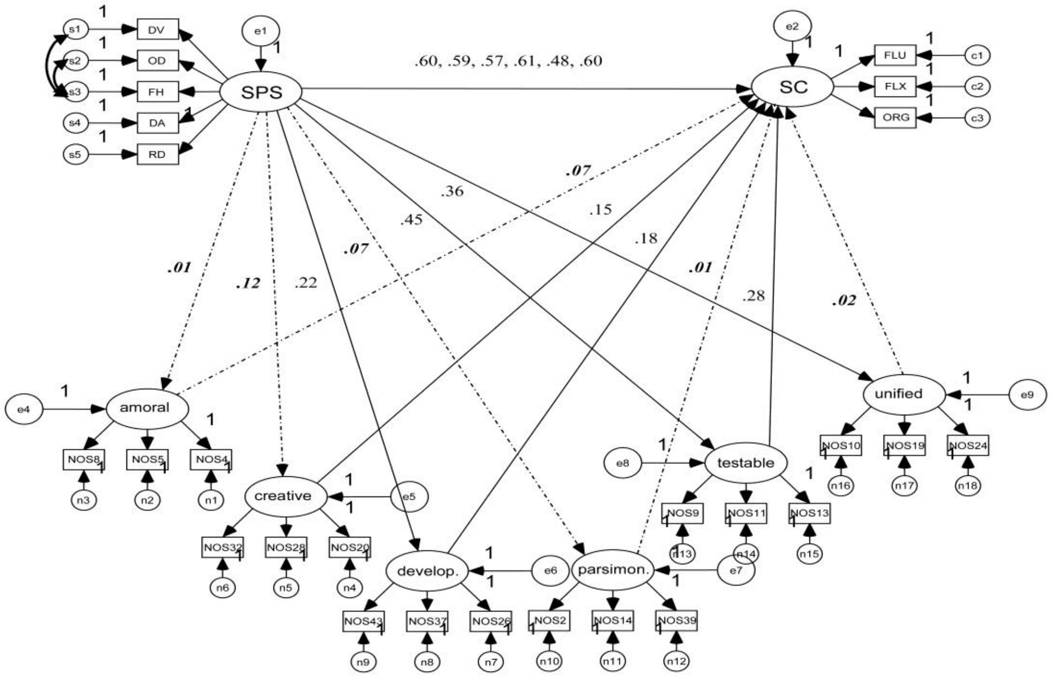


Figure 4. The mediating role of the NSKS subscales

Note: Parameter estimations are standardized values. Bold/italic coefficients and dashed lines are not statistically significant (p>.05). SPS: Scientific Process Skills. DV: Defining Variables, OD: Operational Defining, FH: Formulating Hypotheses, DA: Data Analyses, and RD: Research Design. SC: Scientific Creativity. FLU: Fluency, FLX: Flexibility, ORG: Originality. Amoral, Creative, Developmental, Parsimonious, Testable, and Unified are NSKS's subscales.

As seen in Figure 4, the standardized direct effects of SPS on SC were still significant (β ranged from .48 to .60, *ps*<.001). Namely, when NSKS subscales (Amoral .60, Creative .59, Developmental .57, Parsimonious .61, Testable .48, and Unified .60) were separately included in SEM analysis for Hypothesized Model 1, the significant parameter estimations, which represented the indirect effects of SPS on SC through the NSKS subscales, did not change substantially. The lowest parameter estimations belong to the Testable (β=.48, p<.001) and Developmental (β=.57, p<.001) for indirect effects of SPS on SC through the NSKS subscales. This means that the relationships between SPS and SC may be only mediated by the Testable and Developmental subscales of the NSKS (see Table 4). In addition to these results, Sobel (1982) tests did not confirm the mediating roles of four NSKS subscales, Amoral, Creative, Developmental, Parsimonious, and Unified, in relationship between SPS and SC (p>.05); whereas, the test confirmed the mediating roles of Testable subscale (z=2.61, p<.01) in this relationship(see Table 4). This finding can be interpreted that the Testable subscale has partially mediating role in predictive power of the SPS on the SC.

Table 4. Summary of separate SEM analyses each NSKS subscales

Dependent Variables		Independent Variables	Estimate	S.E.	C.R.	β	*p*	Sobel's *z*
amoral	<---	SPS	.01	.11	.11	.01	.909	
SC	<---	SPS	4.12	.60	6.85	.60	***	
SC	<---	amoral	.39	.39	1.00	.07	.314	.11
creative	<---	SPS	.19	.13	1.42	.12	.155	
SC	<---	SPS	4.57	.64	7.09	.59	***	
SC	<---	creative	.73	.28	2.55	.15	.011	1.24
develop.	<---	SPS	.22	.09	2.35	.22	.018	
SC	<---	SPS	3.95	.58	6.72	.57	***	
SC	<---	develop.	1.18	.52	2.25	.18	.024	1.62
parsimon.	<---	SPS	.02	.01	1.61	.07	.107	
SC	<---	SPS	4.77	.66	7.16	.61	***	
SC	<---	parsimon.	.29	.38	.76	.01	.442	.69
testable	<---	SPS	.65	.16	4.00	.45	***	
SC	<---	SPS	3.45	.60	5.73	.48	***	
SC	<---	testable	1.38	.40	3.44	.28	***	2.61^{**}
unified	<---	SPS	.46	.16	2.77	.36	.005	
SC	<---	SPS	4.25	.66	6.43	.60	***	
SC	<---	unified	.08	.37	.22	.02	.819	.22

$^{***}p<.001$; $^{**}p<.01$; $^{*}p< .05$

Note: SEM analysis table belongs to the Model 1. SPS: Scientific Process Skills. SC: Scientific Creativity. Amoral, Creative, Developmental, Parsimonious, Testable, and Unified are NSKS's subscales.

Again, second-order factor NSKS (i.e., one factor NSKS) SEM analysis was performed for Hypothesized Model 1. Parameter estimates derived from the analysis are indicated in Figure 5. When second-order factor NSKS were included in the SEM analysis of Hypothesized Model 1, the parameter estimation, which represented the indirect effects of SPS on SC through the second-order factor NSKS, was statistically significant (see Figure 5 and Table 5). Sobel (1982) test confirmed the mediating roles of second-order factor NSKS in relationship between SPS and SC (z=2.38, p<.05) (see Table 5). This finding can be interpreted that the NSKS (i.e., second-order factor NSKS) has a partially mediating role in the relationship between the SPS and SC.

The Moderating Role of the NSKS Subscales

In conclusion, NSKS subscales totally influenced the strength of the relationship between the SPS and SC. SPS also affected the influence of NSKS subscales on SC (see Figure 3). Thus, the interaction moderation effects of NSKS subscales on relationship between SPS and SC came to our mind. The findings obtained from the analyses investigating moderation effects indicated that SPS and NSKS subscales were crucial components in SC and showed that the interaction between SPS and NSKS subscales indirectly influenced SC. For this reason, we also investigated NSKS's subscales moderation effect on relationship between SPS and SC.

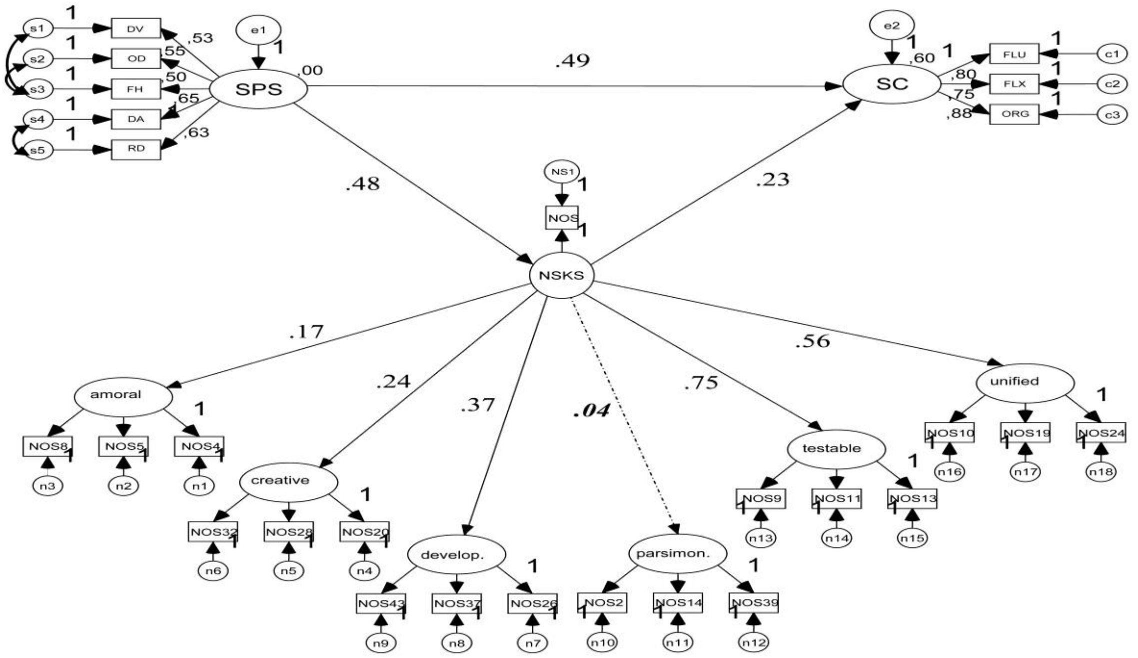


Figure 5. The mediating role of the second-order factor NSKS

Note: Parameter estimations are standardized values. Bold/italic coefficient and dashed line are not statistically significant (p>.05); whereas solid lines are statistically significant (p<.05). SPS: Scientific Process Skills. DV: Defining Variables, OD: Operational Defining, FH: Formulating Hypotheses, DA: Data Analyses, and RD: Research Design. SC: Scientific Creativity. FLU: Fluency, FLX: Flexibility, ORG: Originality. NSKS: Nature of Scientific Knowledge Scale

Table 5. Summary of the second-order factor NSKS in SEM analysis

Dependent Variables		Independent Variables	Estimate	S.E.	C.R.	β	p	Sobel's z
NSKS	<---	SPS	1.54	.33	4.61	.48	***	
SC	<---	NSKS	.54	.19	2.79	.23	.005	2.38*
SC	<---	SPS	3.70	.62	5.93	.49	***	

$^{***}p$<.001; $^{**}p$<.01; $^{*}p$<.05

Note: SEM Analysis Table belongs to the Hypothesized Model 2. SPS: Scientific Process Skills. SC: Scientific Creativity. Creative, Testable, and Unified are NSKS's subscales.

All of the items' scores were converted to standard scores in order to examine the moderation effect of the NSKS subscales. The multiplications were performed between standard scores of the NSKS subscales and SPS. After that, the moderators obtained from these multiplications were included in the SEM analysis. All subscales of NSKS were tested in SEM analysis. In the results, only the Creative of NSKS subscales was significant (β=.18, p<.001); whereas, the others were not significant and standardized parameter estimates ranged from -.08 to .08. Insignificant parameter estimations were .08 for Amoral, .08 for Developmental, .07 for Parsimonious, .05 for Testable, and .03 for Unified (ps>.05). After removing the insignificant subscales of NSKS from the SEM analysis, Goodness-of-fit measures were good fit ($x^2_{(1)}$=3.69, p>.05; x^2/df=3.69; RMR=.09; RMSEA=.09; NFI=.97; CFI=.98; GFI=.99; AGFI=.95). This finding can be interpreted that the Creative subscale has a moderator role on predictive power of the SPS on the SC.

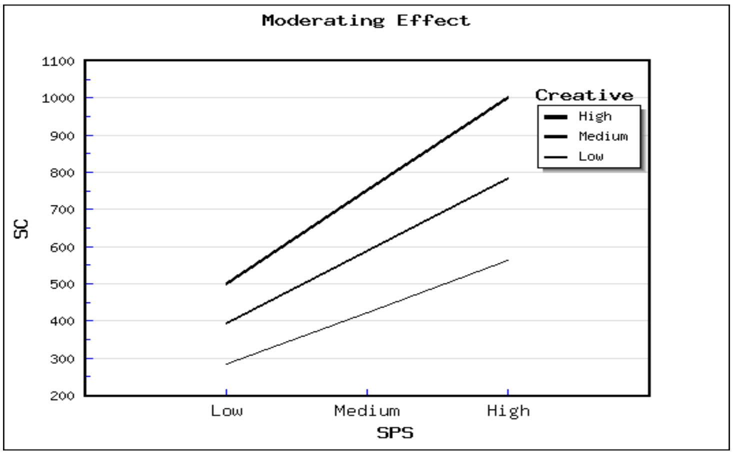


Figure 6. The moderation effects of the Creative from the NSKS subscales

Note: SPS: Scientific Process Skills. SC: Scientific Creativity. Creative is one of the NSKS subscales. The figure was drawn by utilizing http://pavlov.psyc.vuw.ac.nz/paul-jose/modgraph/cont_fig1.php (Jose, 2013).

Table 6. Summary of moderators in SEM analysis

Dependent Variables		Independent Variables	Estimate	S.E.	C.R.	*β*	*p*
ZSC	<---	ZSPS	.54	.05	11.18	.54	***
ZSC	<---	Moderator1	.09	.05	1.62	.08	.103
ZSC	<---	Moderator2	.20	.06	3.45	.18	***
ZSC	<---	Moderator3	-.09	.06	1.59	-.08	.110
ZSC	<---	Moderator4	.06	.04	1.45	.07	.147
ZSC	<---	Moderator5	.05	.05	.94	.05	.344
ZSC	<---	Moderator6	.03	.05	.67	.03	.501

$^{***}p< .001$

Note: Moderators are NSKS subscales. Moderator1: Amoral; Moderator2: Creative; Moderator3: Developmental; Moderator4: Parsimonious; Moderator5: Testable; and Moderator6: Unified. ZSPS: Standard Scores of Scientific Process Skills. ZSC: Standard Scores of Scientific Creativity.

As shown in Figure 6, the three lines are not parallel to each other, which confirm the presence of the moderator effect. Because the lines are not intersecting, the interaction between SPS and Creative is sequential. An individual who has high SC also has high SPS and knows the important role of Creativity in science. Similarly, an individual who has poor SC is poor at SPS and unaware of the important role of Creativity in science. The results of the study indicate that beliefs about the Creativity have a moderating effect on the relationship between the SPS and SC.

Discussion

Today's science education places more attention on innovative thinking, producing and acquiring new ideas, reasoning skills, cognitive development and positive attitudes to science because the current era demands scientifically literate and skilled individuals. In turn, recent literature highlights the role of creativity in science education (Barrow, 2010; Schmidt, 2010). In this context, the results of this study contribute to the literature that students' SC can be developed by developing their SPS where NOS understanding has a partly mediating role between the two.

Two models were tested in the present study. In Hypothesized Model 1, it was considered that the SPS both directly and indirectly predicted SC through NOS's dimensions. In Hypothesized Model 2, it was considered that the SPS directly predicted the SC. The SPS was found to strongly and directly predicted SC in Hypothesized Model 2. These results are consistent with those of Aktamıs and Ergin (2007, 2008) in relationships between SPS and SC. The predictive power of SPS on SC both directly and indirectly increased through NOS's dimensions in Hypothesized Model 1. Due to this reason, NOS understanding dimensions (i.e., NSKS's subscales) were individually included in Hypothesized Model 1. In this relation, only testable subscale of NSKS has a partially mediating role. Testable subscale of NSKS measures students' belief that scientific knowledge is constructed through repeated tests in light of valid observations and open to public examinations. It means that students who know that current scientifically-accepted knowledge can be questioned through ongoing tests and observations demonstrate better SC. By this means, students have the opportunity to reevaluate and interpret new observations and test results through imaginations and creative thinking.

On the other hand, Amoral, Developmental, Parsimonious, and Unified subscales are not mediated in the relationship between SPS and SC. Amoral subscale can be summarized with the role of moral judgments that scientific knowledge cannot be criticized as good or bad, but scientific applications (i.e., technology). Thus, the students' perceptions on this issue naturally did not affect their SC. Developmental subscale emphases the changeable nature of scientific knowledge. Despite it is reasonable to expect that students who believe that scientific knowledge can be changed with new data, new methodological designs, and new interpretations of data…etc., SC of the students will be high. According to our results, Developmental subscale had interpretive power on SC, but it did not take on a mediating role. Parsimonious subscale focuses on the simplicity nature of scientific knowledge, which means similar observations in different situations are conceptualized in a small number of concepts. Also, theories explain the nature in a simple manner. This scale was not correlated to SC. Unified subscale sees close relationships between the laws, theories, and concepts of different disciplines of science. According to our findings, even though SPS had predictive power on the subscale, this subscale did not predict SC. This might be due to the students' grade level, where students know a limited number of concepts from the different fields of science.

Another important result was that although only testable subscale of NSKS had a mediating role in relationship between SPS and SC, this relationship was mediated by second-order factor NSKS (one factor NSKS). This result implies that adequate NOS understanding is required to make a connection between SPS and SC. At this point, the researchers believe that all elements of NOS should be taught in science-related instructions, while students use their SPS in the inquiry activities. This type of learning environment helps students to develop their SC at the same time. Because the predictive power of SPS on SC both directly and indirectly through NOS's dimensions (i.e., NSKS's subscales) was increased in Hypothesized Model 1, we tested the interaction moderation effects of NSKS subscales on the relationship between SPS and SC. This test resulted in an important outcome; that Creative subscale of NSKS has a moderating role on the relationship between SPS and SC. This means that Creative subscale of NSKS regulates the relationship between SPS and SC. Students, who knows the important role of imagination and divergent-thinking in scientific knowledge constructions, possibly try out new methodologies, interpret data with new point-of-views, and go on to develop new solutions and ideas. Consequently, their SC gets higher. In terms of classroom practices, teachers not only put in efforts to develop students' beliefs about NOS understanding focusing on creativity, but they also provide an open-inquiry environment and ill-structured scientific problems for the students in order that they can freely test and share their ideas. As Khishfe and Abd-El-Khalick (2002) stated, "attempts to teach about NOS should be contextualized and woven into inquiry activities and teaching about science content and process skills". Our result indicated that this way of teaching fosters SC.

Limitations of the Study

The present study has four limitations that will guide further studies. Firstly, the sample consisted of two public middle schools' students taking general main courses like mathematics, natural science, Turkish language, and social science; and so the results therefore cannot be generalized to other educational fields and levels. Other fields and levels of education such as vocational, theological, private, elementary, and high school education should also be included within samples for further research. A second limitation of the present study is that the data were obtained from two middle schools' students. Although the school was highly representative of the Turkish science education system, further studies should be conducted based on a larger number of students recruited from different schools across different geographical regions in order to provide more comprehensive results. A third limitation is the cross-sectional nature of the data, with a fourth limitation being data based on self-report measures. The participants' responses on the items of the NSKS's subscales may reflect their hopes about the NOS instead of their actual intentions or knowledge. Consequently, possible effects of social desirability may need to be controlled as part of any future studies.

Conclusions

The results of the present study indicated four major conclusions. Firstly, if students' SPS are developed, their SC also develops. Second, if students' NOS understandings are fully developed, students' NOS beliefs become a mediator in relationships between SPS and SC. Third, because only the Testable dimension of NOS among the other dimensions has a mediating role in relationships between the SPS and SC, students who know that current scientifically-accepted knowledge and open-to-public examinations, can be questioned through ongoing tests and observations demonstrate high levels of SC. The fourth conclusion was that the knowledge and perceptions of the students about the role of creativity in science regulate the relationships between their SPS and SC. However, we know that developing SC is more complex and other ways can be tried to foster SC. "Creative problem solving", "problem solving in the STS context", "creative writing", "creative science inquiry", "creating analogies to understand phenomena and ideas", "challenging students to find connections among apparently unrelated facts and ideas", "mystery solving", and "approaching the teaching and learning of science through the arts" are meaningful ways to improve students' SC in classrooms (see Hadzigeorgiou et al., 2012).

References

Aktamıs, H., & Ergin, Ö. (2007). Investigating the relationship between science process skills and scientific creativity. *Hacettepe University Journal of Education, 33*, 11-23.

Aktamış, H., & Ergin, Ö. (2008). The effect of scientific process skills education on students' scientific creativity, science attitudes and academic achievements. *Asia-Pacific Forum on Science Learning and Teaching, 9*(1), 1-21.

Alexander, P. A. (1992). Domain knowledge: Evolving themes and emerging concerns. *Educational Psychology, 27*, 33–51.

Amabile, T. M. (1996). *Creativity in context*: *Update to the social psychology of creativity.* Boulder, CO: Westview Press.

Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology, 50*(1), 14-23.

Antink Meyer, A., & Lederman, N. G. (2015). Creative cognition in secondary science: An exploration of divergent thinking in science among adolescents. *International Journal of Science Education, 31*(10), 1547-1563.

Ayverdi, L., Asker, E., Öz Aydın, S., & Sarıtaş, T. (2012). Determination of the relationship between elementary students' scientific creativity and academic achievement in science and technology courses. *Elementary Education Online, 11*(3), 646-659.

Baer, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal, 4*, 23-39.

Baer, J. (2016). Creativity doesn't develop in a vacuum. In B. Barbot (Ed.), *Perspectives on creativity development*, New Directions for Child and Adolescent Development (pp. 9-20). http://dx.doi.org/10.1002/cad.20151

Baer, J., & Kaufman, J. C. (2006). Creativity research in English-speaking countries. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The international handbook of creativity*. New York, NY: Cambridge University Press.

Bakeman, R., & Robinson, B. F. (2014). *Understanding statistics in the behavioral science.* Mahwah, New Jersey: Psychology Press.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182. http://dx.doi.org/10.1037/0022-3514.51.6.1173

Barron, F., & Harrington, D.M. (1981). Creativity, Intelligence, and Personality. *Annual Review of Psychology, 32*, 439-476.

Barrow, L. H. (2010). Encouraging creativity with scientific inquiry. *Creative Education, 1*, 1-6.

Bentler, P. M. (2005). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York: Routledge, Taylor & Francis Group.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). An experiment is when you try it and see if it works: A study of grade 7 students' understanding of the construction of scientific knowledge [Special issue]. *International Journal of Science Education, 11*, 514-529.

Comrey, A. L. (1988). Methodological contributions to clinical research. Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology, 56*(5), 754-761.

Cropley, A. J. (1992). *More ways than one: Fostering creativity*. Norwood, NJ: Ablex Publishing Corporation.

Cropley, A. J. (1997). Fostering creativity in the classroom: General principles. In M.A. Runco (Ed.), *Creativity Research Handbook, Vol. 1* (pp. 83-114). Cresskill, N. J.: Hampton Press.

Cropley, A. J., & Cropley, D. (2009). *Fostering creativity: A diagnostic approach for higher education and organizations.* Cresskill, NJ.: Hampton Press.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods,2*(3), 292-307.

http://dx.doi.org/10.1037/1082-989X.2.3.292

Deniş Çeliker, H., & Balım, A. G. (2012). Adaptation of Scientific Creativity Test to Turkish and its Assessment Criteria. *Uşak University Journal of Social Sciences, 5*(2), 1-21.

Edmondson, K. M., & Novak, D. N. (1993). The interplay of scientific epistemological views, learning strategies, and attitudes of college students. *Journal of Research in Science Teaching, 30,* 547-559.

Erdogdu, M. Y. (2006). Relationships between creativity, teacher behaviours and academic success [in Turkish]. *Electronic Journal of Social Sciences, 5*(17), 95-106.

Geban, Ö., Aşkar, P., & Özkan, İ. (1992). Effects of computer simulation and problem solving approaches on high school. *Journal of Educational Research, 86* (1), 5-10.

Hadzigeorgiou, Y., Fokialis, P., & Kabouropoulou, M. (2012). Thinking about creativity in science education. *Creative Education, 3*(5), 603-611. http://dx.doi.org/10.4236/ce.2012.35089.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Han, K. S. (2003). Domain-specificity of creativity in young children: How quantitative and qualitative data support it. *Journal of Creative Behavior, 37,* 117-129.

Hennessey, B. A. (1995). Social, environmental, and developmental issues and creativity. *Educational Psychology Review, 7*(2), 163-183. http://dx.doi.org/10.1007/BF02212492

Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp.76-99). Thousand Oaks, CA: Sage Publications.

Hu, W., & Adey, P. (2002). A scientific creativity test for secondary school students. *International Journal of Science Education, 24*(4), 389-403. http://dx.doi.org/10.1080/09500690110098912

Hu, W., Shi, Q. Z., Han, Q., Wang, X., & Adey, P. (2010). Creative scientific problem finding and its developmental trend. *Creativity Research Journal, 22*(1), 46-52.

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology, 17*(2), 139-153. http://dx.doi.org/10.1016

Isaksen, S. G., & Treffinger, D. J. (1985). *Creative problem solving: The basic course.* Buffalo, NY: Bearly Limited.

Jose, P. E. (2013).*ModGraph-I: A programme to compute cell means for the graphical display of moderational analyses: The internet version, Version 3.0.* Victoria University of Wellington, Wellington, New Zealand. Retrieved from http://pavlov.psyc.vuw.ac.nz/paul-jose/modgraph/

Kaufman, J. C., & Baer, J. (2008). Role of the domain in creative learning in USA. In A. R. Craft, T. Cremin, & P. Burnard (Eds.), *Creative learning 3–11 and how we document it* (pp. 27-34). Staffordshire, UK: Trentham Books.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology,* 13, 1-12. doi:10.1037/a0013688

Khishfe, R., & Abd-El-Khalick, F. (2002). The influence of explicit reflective versus implicit inquiry-oriented instruction on sixth grades' views of nature of science. *Journal of Research in Science Teaching, 39*(7), 551-578.

Kılıç, K., Sungur, S., Çakıroğlu, J., & Tekkaya, C. (2005). Ninth grade students' understanding of the nature of scientific knowledge. *Hacettepe University Journal of Education, 28,* 127-133.

Kim, K. H., & Bentler, P. M. (2006). Data modeling: Structural equation modeling. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 161-175). Mahwah, NJ: Lawrence Erlbaum.

Kleibeuker, S. W., De Dreu, C. K., & Crone, E. A. (2013). The development of creative cognition across adolescence: Distinct trajectories for insight and divergent thinking. *Developmental Science, 16*(1), 2-12.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*(4), 331-359.

Lederman, N. G. (2004). Syntax of nature of science within inquiry and science instruction. In L. B. Flick, & N. G. Lederman (Eds.), *Scientific Inquiry and Nature of Science*. Netherlands: Kluwer Academic Publishers.

Lederman, N. G. (2007). Nature of science: Past, present and future. In S. K. Abell & N.G. Lederman (Eds.), *Handbook of research on science education* (pp. 831-879). Mahwah, NJ: Lawrence.

Lederman, N. G. & Abd-El-Khalick, F. (1998). Avoiding de-natured science: Activities that promote understandings of the nature of science. In W. McComas (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 83-126). Dordrecht, The Netherlands: Kluwer Academic.

Lubart, T. I. (1994). Creativity. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 290-332). San Diego, CA: Academic.

McComas, W. F., & Olson, J. K. (2002). The nature of science in international science education standards documents. In W. F. McComas (Ed.), *The nature of science in science education* (pp.41-52). Rotterdam: Springer.

Meador, K. (2003). Thinking creatively about science: Suggestions for primary teachers. *Gifted Child Today, 26*(1), 25-29.

Okey, J. R., Wise, K. C., & Burns, J. C. (1982). *Integrated process skill test-2.* Department of Science Education, University of Georgia, Athens, GA, 30602.

Osborn, A. F. (1953). *Applied imagination: Principles and procedures of creative thinking.* NY: Charles Scribner's Sons.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717-731. http://dx.doi.org/10.3758/BF03206553

Ren, F., Li, X., Zhang, H., & Wang, L. (2012). Progression of Chinese students' creative imagination from elementary through high school. *International Journal of Science Education, 34*(13), 2043-2059.

Rubba, P. A., & Andersen, H. O. (1978). Development of an instrument to assess secondary school students' understanding of the nature of scientific knowledge. *Science Education 62*(4), 449-458. http://dx.doi.org/10.1002/sce.3730620404

Runco, M., Nemiro, J., & Walberg, H. (1998). Personal explicit theories of creativity. *Journal of Creative Behavior, 32*(1), 1-17.

Ryan, A. G., & Aikenhead G. S. (1992). Students' preconceptions about the epistemology of science. *Science Education, 76*, 559-580.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23-74.

Schmidt, A. L. (2010). The battle for creativity: Frontiers in science and science education. *BioEssays, 32*(12), 1016-1019.

Schumacker, R. E., & Lomax, R. G. (2004). *Beginner's guide to structural equation modeling* (2nd ed.). New York: Psychology Press, Taylor & Francis Group.

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. *Sociological Methodology*, *13,* 290-312.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity.* New York: Free Press.

Torrance, E. P. (1968). *Education and the creative potential*. Minneapolis: University of Minnesota Press.

Torrance, E. P. (1995). *Why fly? A philosophy of creativity.* Norwood, N.J.: Ablex Publishing.

Author Information

Gokhan Ozdemir
Ömer Halisdemir University
51240-Nigde, Turkey
Contact e-mail: *gokhanozdemir@ohu.edu.tr*

Ayhan Dikici
Ömer Halisdemir University
51240-Nigde, Turkey

Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

Does Hands-on Science Practices Make an Impact on Achievement in Science? A Meta-Analysis

Serdar Caglak

Article Info

Article History

Received:
08 August 2016

Accepted:
21 November 2016

Keywords

Hands-on science
Meta-analysis
Effect size
Science achievement
Random-effects model
Mixed-effects model

Abstract

This study aimed to investigate to what extent the use of hands-on science activities influences on students' academic achievement in science. Review of literature revealed several research studies focusing upon such aim and thus, a meta-analysis of these researches was carried out to obtain an overall effect size estimate of hands-on science activities on science achievement. Of the available studies, 15 with multiple outcomes satisfied the pre-determined inclusion criteria. In addition to the estimation of overall effect size using fixed- and random-effects models, subgroup analyses were also run through a mixed-effect model to determine whether heterogeneity in effect size estimates is due to the influence of moderator variables. Results showed that the estimated effect size was statistically significant (Z=8.57, p < .01). The magnitude of the overall effect size estimate indicated that the hands-on activities had a very large impact on students' science achievement (Hedge`s g = 1.55, 95% CI= [1.20-1.91]). The effect size estimates for each moderator variable was also statistically significant at alpha level of .01, but no difference was found within sub-categories of moderator variables.

Introduction

High quality science education is very important for nations to shape their future in accordance with the current global trend and development in science and technology (Kalender & Berberoglu, 2009). Many nations therefore desire individuals with high proficiency levels in science disciplines to cultivate scientifically literate citizens in order to accomplish their national goals, such as having advanced industrial technology and high heels in the global economy (Dillon, 2009). Reform documents in many countries emphasize the importance of facilitating students' learning in science and developing scientifically literate citizens (American Association for the Advancement of Science [AAAS], 1994; Ministry of National Education of Turkey [MoNET], 2006, 2013; National Research Council [NRC], 2012).

Those reform documents and attempts also put emphasize on the students in various level to be active learners and problem solvers though learning by doing in order to achieve the goals of scientific literacy desired for all citizens. In order to investigate and understand real-life phenomena and develop scientific skills and attitudes, science classes should be enhanced with experimental procedures and scientific reasoning (AAAS, 1994). However, traditional classes could not be seen as to promote to this aim. Alternative or complimentary ways are needed. In this regard, hands-on activities in science classes could be complimentary to the direct instruction to encourage students participate in active learning settings in order to achieve the aim of scientific literacy. Hands-on activities particularly present an avenue for facilitating and richening students' success and attainments in science. It allows students to be more active and engaged in science classrooms. As a result, hands-on is appeared as the key component of science education to promote students' interest in science, conceptual understanding, and scientific literacy.

Hands-on can be defined as "*educational experience that actively involves people in manipulating objects to gain knowledge or understanding*" (Haury & Rillere, 1994, p.11). It encompasses learning by experiences and doing (Flick, 1993). Students can manipulate materials through hands-on activities while learning science during the coursework (Flick, 1993). Those materials can be variety of instruments or other things such as bulbs, test tubes, microscopes, thermometers, plants, rocks, insects, leafs and so on. Learning science with those materials in classes can help students gain unique experiences that are inaccessible otherwise (Scharfenberg & Bogner, 2011). Hands-on practices help students deal with real life cases and understand the meaning behind through observation, interaction, and reasoning, and so on (Erentay & Erdogan, 2009).

Students are expected to gain deep understanding of subject matters and be able to do science through those activities. When students are able to work subject materials and manipulate some variables in hands-on activities, they would be able to carry out their own scientific investigations so that science concepts become less abstract and students can have the opportunity for getting deeper understanding along with more open-ended questions. Students' involvement through hands-on activities contribute to develop their discovery, critical thinking, and problem solving skills because students have to rely on the data and evidence obtained from their own investigations (Flick, 1993).

Researchers have acknowledged that hands-on approach is an important specific instructional strategy which aims at engaging students in teaching practices to facilitate students' active learning process (Flick, 1993; Haury, & Rillero, 1994). Within this context, many studies have been conducted to examine the effectiveness of hands-on strategy on students' achievement in science (Ates & Eryılmaz, 2011; Bigler & Hanegan, 2011; Costu, Unal, & Ayas, 2007; Ekmekci & Gulacar, 2015; Gaitano & Bogner, 2011; Glasson, 1989; Randler & Hulde, 2007). The results of these studies have revealed that in general hands-on activities increased student's science achievement levels. For example, a report on science achievement of elementary students in the United States indicated that "students whose teachers reported that their students do hands-on projects every day or almost every day scored higher on average than students whose teachers reported students did hands-on projects in class less frequently" (NCES, 2011, p.10). On the other hand, results from a recent meta-analysis on control-of-variables strategy (CVS) revealed that the utilizing of the CVS had not significantly different compared to studies that did not use hands-on training (Schwichow, Croker, Zimmerman, Höffler, & Härtig, 2015). In an intervention research conducted by the same research team (Schwichow, Zimmerman, Croker, & Härtig (2016), it was found that hands-on tasks did not differ significantly eight grade students' achievement compared to paper-and-pencil training tasks.

Many research have also been conducted to explore the impact of hands-on activities on other outcome variables including attitude (Bilgin, 2006; Koc, & Boyuk, 2012; Sadi & Cakiroglu, 2011), interest (e. g., Holstermann, Grube, & Bogeholz, 2010; Randler & Hulde, 2007; Paris, Yambor, & Packard, 1998), and science process skills (Bilgin, 2006). Results of these studies briefly revealed that the use of hands-on activities in science instruction is an effective way of promoting students' attitudes towards science including both cognitive and affective skills. Additionally, findings from the literature also showed that the use of hands-on activities is a useful and an enjoyable way to teach and learn science compared with traditional or teacher-centered learning activities while mastering the science coursework. Hands-on science activities were observed to associate with students' interest (Bulunuz, 2012; Holstermann et al., 2010). What is missing in science education literature, however, is the systematic review of studies investigating to what extend hands-on activities improves students' academic achievement. Although a large body of studies on hands-on has been conducted in the science education literature, no meta-analysis study exists which examines the overall effectiveness of hands-on activities in developing achievement in science. Available studies mainly focus on the effects of hands-on activities on a science achievement, but individual research of testing the impacts of hands-on using is less concerned without an overall conclusion drawn from a meta-analysis. In this regard, the present study is designed to fill out this gap in the literature by compiling studies and analyzing them to draw an overall conclusion on effectiveness of hands-on activities on academic achievement. Generally speaking, the main purpose of a systematic review is to answer a defined research question using all kind of empirical evidences collected and summarized from studies which satisfy pre-determined inclusion criteria (Borenstein & Higgins, 2009). A meta-analysis is therefore actually a general name of all statistical procedures or methods used to combine and summarize those findings from all available studies in a way that an overall conclusion can be drawn about the effectiveness of an intervention or an experiment (Higgings & Greens, 2008). In this study, the outcome measure is the science achievement of students, but the effectiveness of hands-on science instruction is considered as the treatment or the intervention. Hence, a meta-analysis of research on hands-on activities in Turkish literature becomes the central focus of this study. Even though many studies exist in the literature which investigates the effectiveness of different instructional strategies, the number of studies focusing on the use of hands-on activities and conducted in Turkey is very limited. Within this limitation, this study was conducted to examine to what extent hands-on science instruction is effective and how it influences students' science achievement in Turkey. Hence, findings from this meta-analysis study may be interesting for teachers, researchers, and educational policy makers.

Method

The aim of this study is to explore the overall effectiveness of hands-on activities on science achievement of Turkish students. A meta-analysis technique is therefore used to combine the findings of the relevant research.

Meta-analysis is known as a statistical technique used to determine a common statistical measure - typically an effect size measure - in order to draw an overall conclusion about the effectiveness of an intervention, a treatment, or an experiment (Borenstein et al., 2009; Glass, 1976). Fixed-effects and random-effects models are commonly used in meta-analysis research to obtain a summary effect. In this study, both statistical models were considered, but the most appropriate model was applied to data. Further analyses were conducted using mixed-effects modeling to examine whether the overall effect size estimates show any statistical difference within- and between-levels of study-level (moderator) variables, such as treatment time, grade levels, and so forth. In other words, a mixed-effects model can be used to determine if there is any influence of moderator variables on the heterogeneity in the effect size estimates by taking the all levels of the moderator variables into consideration at once (Viechtbauer, 2010). The R 3.2.3 (R Core Team, 2015) software and a package called metaphor 1.9-8 (Viechtbauer, 2015) were used for all meta-analytic procedures and further statistical analyses in this study.

Acquisition of Studies

Despite the fact that a great amount of research exists in the literature focusing on the effects of different instructional strategies, the number of the experimental studies which used hands-on activities as intervention to teach science has been limited. Even though two studies are enough to conduct a meta-analysis, this quantity may yield a bias estimation of the population effect size (Valentine et al., 2010). Of the available research, a total of 15 studies were found to satisfy the pre-determined inclusion criteria. Three studies included more than one treatment outcomes (e.g. effect size). Thus, a total of 19 individual effect size estimates were considered for computation and used in this meta-analysis.

Inclusion Criteria

Since the literature reveals large amount of research studies, the inclusion criteria below were determined to eliminate the non-relevant studies to be more focused upon. To be included in the meta-analysis, a study must

- be an experimental study
- include treatment and control groups with pre/post-tests or post-tests only design
- assign students randomly to the treatment and control groups
- assess students` achievement (or success, performance, etc.) regarding sub-disciplines of science (e.g., biology, chemistry, and physics)
- manipulate hands-on activities as an intervention or a treatment
- must provide adequate statistical outcomes for effect size computation
- be undertaken with Turkish students between the years of 2000 and 2015

Considering the criteria above, reliable databases such as Web of Science, ERIC, ULAKBIM and National Thesis Archive of the Council of Higher Education of Turkey were firstly selected and they were carefully scanned in March 2015 using the keywords of “hands-on”, “hands-on activities”, hands-on science” “hands-on science teaching and learning”, “hands-on science instruction”, and “learning science with hands-on activities”. A comprehensive review of the literature in these databases resulted in a total of 25 studies. Since full-text of three studies could not be accessed in the search, these studies were obtained via direct communication to the corresponding authors. Close examination of the studies revealed that 10 were not in line with the pre-determined criteria, and thus excluded for the analysis. The total number of the studies used in the meta-analysis was reduced to 15. A flow chart is presented below to indicate how the inclusion process was carried out.

Coding Procedure

The coding procedure of the selected studies was completed by three authors independently. Inter-coder agreement for each of the selected studies was obtained to verify that the data entry for effect size computation procedure was accurate. Consistent and inconsistent patterns between each pairs of the ratings were quantified and later the phi-coefficient was computed to determine whether the agreement among the coders was at an adequate level (.70 < phi). A coding sheet was first created to identify the following attributes of each of the selected study. Some of those attributes were also used as moderator variables to further investigate whether the overall effect size estimates show any statistical difference based on each level of the moderator variables.

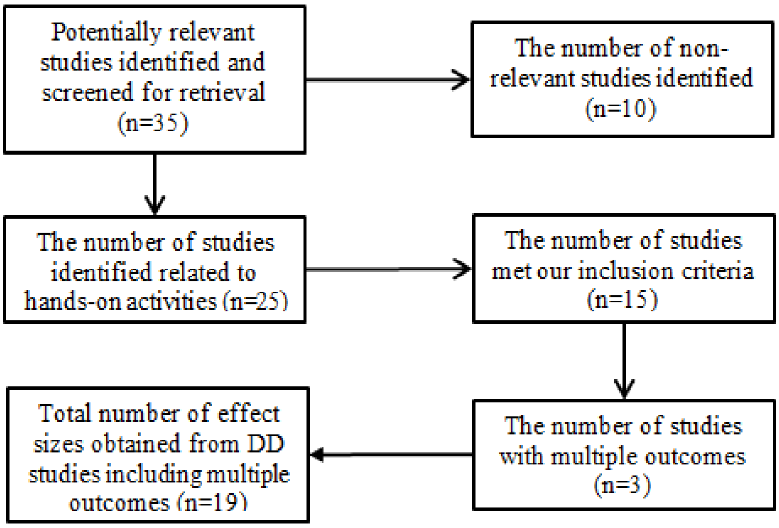


Figure 1. A flow chart for the selection of studies

Demographics and Methods of the Selected Study

In this part of the coding process, some of the important information about the selected studies was recorded such as the number of authors, name(s) of the author(s), title of the study, publication year, the location where the study took place, and also the identification numbers given to the studies.

Subjects

The characteristics and demographics of the samples used in the experiments and treatments may show differences. Since number of the studies on the selected theme was limited in the context of Turkey, the subjects of the selected studies could vary in terms of level and grade (e.g., 6th, 7th, 10th grade etc). Therefore, the effect size estimates must be evaluated based on those differences introduced by the moderator variables. In this part of the coding process, the demographic information about the participants and schools were reported such as grade levels, school type and location, and geographical region.

Sampling

This section provided information about the samples, sampling procedures, and the way used to assign the participants to the experiment and the control groups. In addition, even though external validity referring to generalizability of the findings in the experimental studies are limited and sometimes meaningless, the population of interest in each of the selected studies was also reported to better identify what grade or school level was the target group of interest.

Methodology

The selected studies were categorized based on their type of research (i.e., dissertation, journal article, etc.). In addition, some information, such as reliability and validity, was also obtained based on whether an instrument used to measure the outcome variable of interest, psychometric properties of that instrument if provided.

Treatment Time and Statistical Outcomes

Statistical outcomes from the selected studies were reported for each of the treatment and control groups. It was noticed that the duration of treatment times varied across studies. Therefore, the scale of the original treatment time was converted to a common scale to better compare the effect of treatment time on the overall effect size estimates within each sub-category of the moderator variables.

Subject Area

Science is a very broad name used for several subject areas at different grade levels. Therefore, those subject areas and grade levels of interest in this study would be an important factor to be examined in order to discuss the potential reason why the hands-on activities took place on a specific subject area at most. In addition, the variation in different subject areas may also contribute to the heterogeneity in the effect size estimates across studies. However, science subjects are taught in an integrated science curriculum for the elementary and middle school students in Turkey. Therefore, instead of specifying the subject areas separately, such as biology chemistry, and physics, we classified them as elementary school science, middle school science, and so on to use them in the moderator analysis.

Dependent Variable

In this study, the dependent variable of interest was the students' academic achievement, which could be expressed as success, performance, and learning outcomes interchangeably in different contexts. Researchers generally preferred using achievement tests to assess the academic achievement in their studies. Some of them developed their own instruments to measure students' proficiencies on the selected subject area of interests. Test scores obtained from those measurements were typically reported and also compared for both treatment and control groups to evaluate the effectiveness of the experiment or intervention in those studies.

Independent Categorical Variables

As mentioned earlier, the characteristics of the selected studies were taken into account by several variables identified by the researchers during the coding procedure. Therefore, the authors believed that those participants and experiment related variables might have contributed to the variation in the outcome measures. Eventually, the magnitude of heterogeneity in effect size estimates across studies might have occurred due to the variation in the sample and design characteristics of the experiments taken place in those selected studies. Hence, the potential effects of the independent categorical or moderator variables on the overall effect size estimate were investigated in this study.

Data Analysis

The first stage in a meta-analysis is to estimate the mean effect size and its variance, and also to determine if there is any significant variation in effect size estimates across studies (Pigott, 2012). When the mean effect size estimate is obtained through the fixed-effects model, it is assumed that the studies share a common true effect size, but the differences in effect size estimates are only due to the sampling error. However, the central focus of random-effects model is that it examines the variation in effect size estimates based on the differences across studies. Researchers therefore have a tendency to use the random-effects model since there are multiple sources which may contribute not only to the sampling error, but also to the differences across studies. In other words, studies typically show differences in terms of their methods and/or the characteristics of the samples which potentially affect the heterogeneity among studies (Viechtbauer, 2010). However, statistical evidences are always required to reach this conclusion in order to support such claims.

Restricted Maximum Likelihood (RML) was specified as the between-studies heterogeneity (τ^2) estimator since it is approximately unbiased and quite efficient among the other estimation methods (Viechtbauer, 2005). The statistical significance of the heterogeneity can be determined using the Q-statistic, but it does not provide accurate information about the extent of true heterogeneity since its statistical power depends mainly on the number of the studies. Thus, in addition to the Q-statistic, different statistical indices (I^2 and H^2) were also utilized in this study to assess the total amount of heterogeneity. The I^2 index was proposed by Higgins and Thompson (2002) to better quantify the true heterogeneity from a collection of effect size measures by comparing the Q-value with its expected value (its degrees of freedom) by which the homogeneity is assumed across effect size measures. The I^2 is interpreted as a percentage of heterogeneity, that is, 25% indicates a low, 50% represents a moderate, and 75% poses a high heterogeneity across effect size estimates (Huedo-Medina et al., 2006). In addition, H^2, also known as Birge's ratio (Birge, 1932), was proposed by Higgins and Thompson (2002) as another index of heterogeneity. The H^2 is formulized as the ratio of Q-value to its degrees of freedom, and thus, the H^2 values larger than 1 indicates that between-study variation is more than the within-study variation (Konstantopoulos & Hedges, 2009).

Results

Our data analysis showed that there was a great amount of heterogeneity in the observed effect size estimates obtained from the selected studies [Q (18) = 175.12, $p < .01$]. In addition, the I^2 value also shows that 90.45% of the total variability was due to the total heterogeneity across effect size measures. Moreover, the H^2 value showed that the total variability was approximately 10 times larger than the sampling variability in the effect size measures. Several plots of influence diagnostics were also used (a) to verify that observed effect size measures were heterogeneous, (b) to determine if there was any outlier in the data, (c) to detect if the non-normality concern exists, (d) to check whether the model-data fit was adequate, (e) and to investigate whether publication bias exists (Wang and Bushman, 1998).

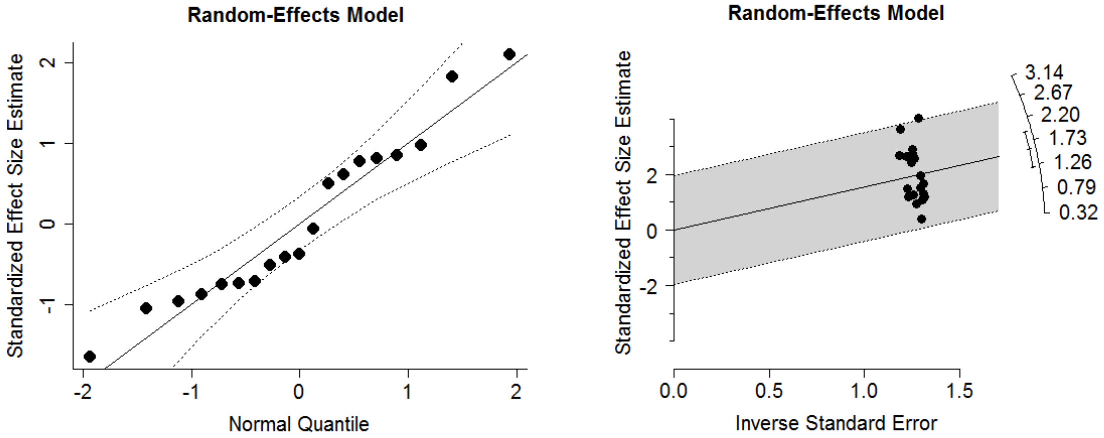


Figure 2. Normal Q-Q and radial (galbraith) plots

Quantile-by-quantile (Q-Q) plots with 95% confidence bands can be useful in meta-analyses to check various aspects and assumptions of the data by plotting the sample quantiles or percentiles of observed effect size measures versus theoretical or predicted quantiles. It is assumed that both measures come from Gaussian distribution, and thus, the quantiles of both the sample and predicted values are supposed to be linearly laid on a 45 degree line on the coordinate axes and which is considered as an indication of normality of the data. Deviations from the 95% confidence bands indicates that the residual heterogeneity in the true effects is non-normally distributed, which means that the observed data do not come from a single normal population. However, the Q-Q plot of the meta-analytic data used in this study did not violate the normality assumption as seen on the left hand side in Figure 2. In addition, the radial *(Galbraith)* plot (on the right hand side) also showed that the total sample size in each individual study was similar to each other, but the heterogeneity in the effect size estimates across studies was present as also indicated with several statistical indices. One of the effect size estimates seemed to be an outlier in the meta-analytic data since the standardized estimate overlapped with the upper level confidence interval band as seen in Figure 2. Several diagnostic procedures (e.g., DFBETAs, Cook's Distance, Covariance Ratios, etc.) were adopted from Viechtbauer and Cheung (2010) to detect the potential outliers in the data. Statistical findings from those diagnostic procedures indicated that no significant outliers were detected as seen in Figure 3, and thus, the statistical analysis was completed using the initial data obtained from the studies which met the inclusion criteria.

Computation of Overall Effect Size

An effect size measure (Cohen's *d*) from the standardized mean difference (*SMD*) family was calculated for each achievement measure described in the selected studies. Names of the study authors, sample sizes used in the treatment and control groups, effect size estimates with 95% confidence intervals, and the overall effect size estimate were presented for the selected studies in Figure 4. *Hedge's g,* an adjusted version of *d*-type effect size measure, was reported for each individual study to obtain more accurate estimates of the effect sizes. The estimated overall effect size is statistically significant ($Z = 8.57$, $p < .01$). According to Cohen (1988)'s

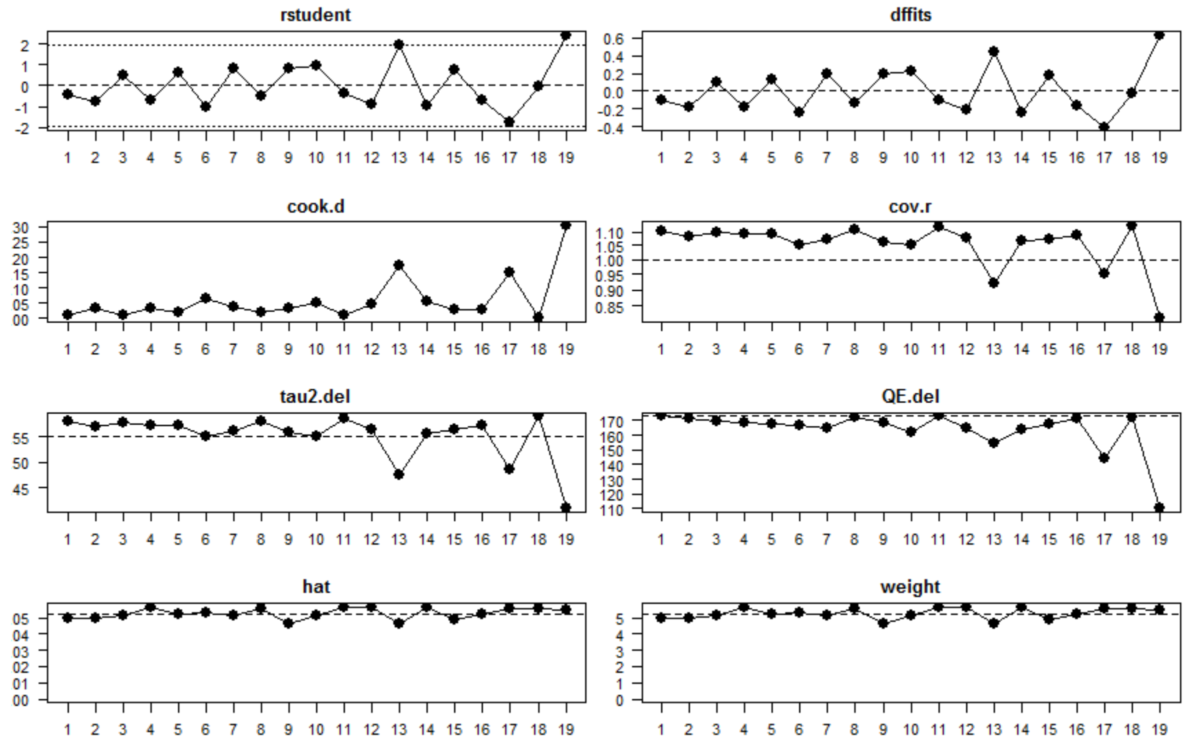
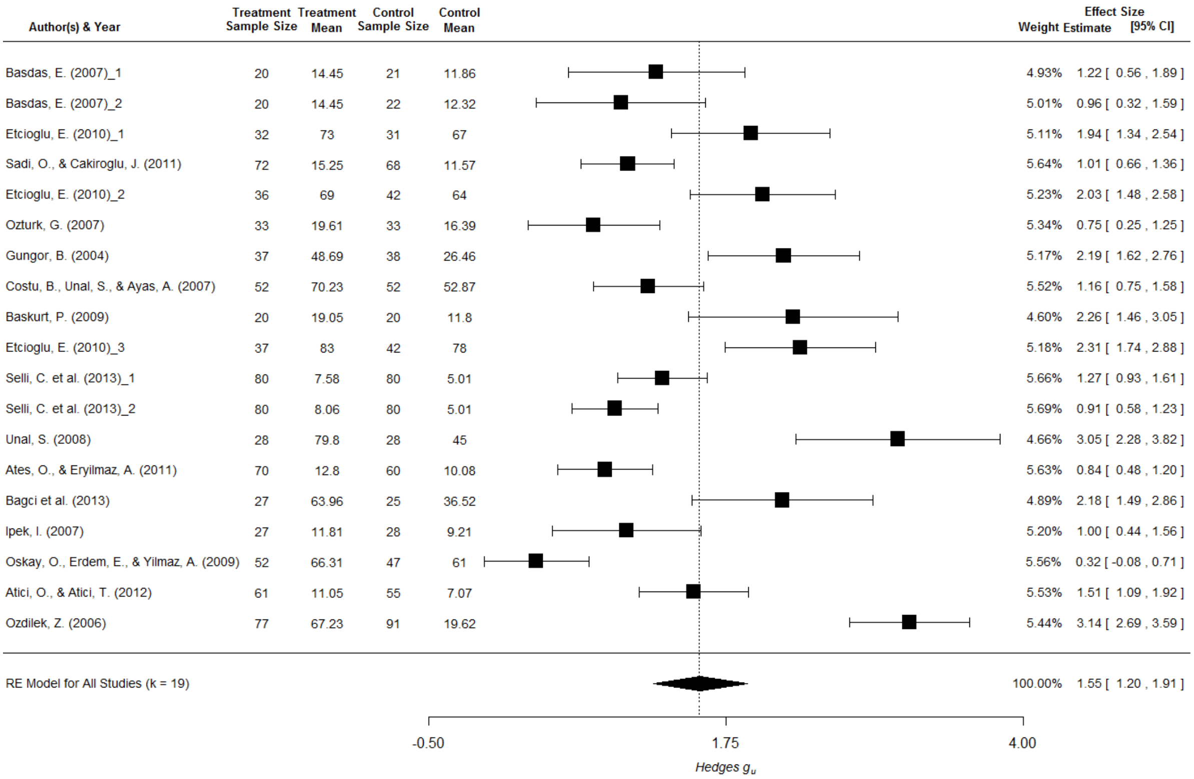


Figure 3. Diagnostics for the outlier detection

classification, its magnitude is very large (Hedge`s *g* = 1.55, 95% CI = [1.20-1.91]). In other words, the magnitude of the true effect is very large, referring that the use of hands-on activities in science teaching has a very large impact on students' academic achievement.



Author(s) & Year	Treatment Sample Size	Treatment Mean	Control Sample Size	Control Mean	Weight	Effect Size Estimate [95% CI]
Basdas, E. (2007)_1	20	14.45	21	11.86	4.93%	1.22 [0.56 , 1.89]
Basdas, E. (2007)_2	20	14.45	22	12.32	5.01%	0.96 [0.32 , 1.59]
Etcioglu, E. (2010)_1	32	73	31	67	5.11%	1.94 [1.34 , 2.54]
Sadi, O., & Cakiroglu, J. (2011)	72	15.25	68	11.57	5.64%	1.01 [0.66 , 1.36]
Etcioglu, E. (2010)_2	36	69	42	64	5.23%	2.03 [1.48 , 2.58]
Ozturk, G. (2007)	33	19.61	33	16.39	5.34%	0.75 [0.25 , 1.25]
Gungor, B. (2004)	37	48.69	38	26.46	5.17%	2.19 [1.62 , 2.76]
Costu, B., Unal, S., & Ayas, A. (2007)	52	70.23	52	52.87	5.52%	1.16 [0.75 , 1.58]
Baskurt, P. (2009)	20	19.05	20	11.8	4.60%	2.26 [1.46 , 3.05]
Etcioglu, E. (2010)_3	37	83	42	78	5.18%	2.31 [1.74 , 2.88]
Selli, C. et al. (2013)_1	80	7.58	80	5.01	5.66%	1.27 [0.93 , 1.61]
Selli, C. et al. (2013)_2	80	8.06	80	5.01	5.69%	0.91 [0.58 , 1.23]
Unal, S. (2008)	28	79.8	28	45	4.66%	3.05 [2.28 , 3.82]
Ates, O., & Eryilmaz, A. (2011)	70	12.8	60	10.08	5.63%	0.84 [0.48 , 1.20]
Bagci et al. (2013)	27	63.96	25	36.52	4.89%	2.18 [1.49 , 2.86]
Ipek, I. (2007)	27	11.81	28	9.21	5.20%	1.00 [0.44 , 1.56]
Oskay, O., Erdem, E., & Yilmaz, A. (2009)	52	66.31	47	61	5.56%	0.32 [-0.08 , 0.71]
Atici, O., & Atici, T. (2012)	61	11.05	55	7.07	5.53%	1.51 [1.09 , 1.92]
Ozdilek, Z. (2006)	77	67.23	91	19.62	5.44%	3.14 [2.69 , 3.59]
RE Model for All Studies (k = 19)					100.00%	1.55 [1.20 , 1.91]

Figure 4. Forest plot for the overall effect size estimation

Subgroup Meta-Analysis and Comparison of Effect Size Estimates

Even though an individual effect size estimate is obtained from a single study, the weighted mean effect size is obtained from an array of studies in a meta-analysis (Borenstein & Higgins, 2013). The individual effect size estimates however may show similarities based on the common characteristics of subject or/and studies which eventually affect the effect size estimates. Therefore, subgroup analysis is an important part of the meta-analytic process to examine whether an intervention is differentiated by each level of the subgroups. One can also statistically compare the mean effect size estimates from the studies that used one variant of the intervention and the other studies that used another variant via mixed-effects modeling approach.

As aforementioned, a great amount of the heterogeneity exists in the individual effect size estimates obtained from the selected studies. Thus, several study-level variables were used to classify the studies in different subgroups in order to explore whether the mean effect size estimates show statistically significant differences across subgroups. Those subgroups were treated as moderators and created based on the study-level variables such as study type, school level, treatment time, geographical region, treatment location and grade level. In total, six different characteristics of the studies were considered for the subgroup meta-analysis. Either the fixed-effects model or the random-effects model was implemented based on the degree of the heterogeneity exists within subgroups to estimate the mean effect size and its variance for each subgroup. Then, the mean effect sizes were statistically compared through the mixed-effects analyses. Effect size estimates (see Table 3) and forest plots (see Figures 6-11) associated with each of the subgroup analyses were given in Appendix A.

Table 1 shows some of the outputs from the mixed-effects analysis. The mean effect size estimates were estimated for each of the subgroups. The within-groups heterogeneities (Q_E) were statistically significant across all of the moderators; the between-groups heterogeneities (Q_M) were not significant at alpha level .05. This omnibus tests indicated that the effect size estimate for each level of the moderator variables were not statistically different from each other. I^2 and H^2 indices also showed that the amount of the heterogeneity across effect size estimates for each moderator variables were quite large. However, the amount of the heterogeneity accounted for by the each level of the moderator variables was not large enough to be statistically significant at alpha level of .05. Except for the treatment time and grade level, the other moderator variables in the model did not explain significant amount of variation in the effect size heterogeneity. These results refer that the differences in the study type, school levels, treatment location, and the geographical region where the experiments or intervention took place are not good predictors of the heterogeneity in the effect size estimates.

Table 1. Moderator analyses through mixed-effects modeling

Moderators	$Q_E{}^a(df)$	$Q_M{}^b(df)$	τ^{2c}	SE^d	I^{2e}	H^{2f}	R^{2g}
Study Type	136.34* (*df*=17)	1.704 (*df*=1)	.518	.203	89.88%	9.88	6.08%
Treatment Time	108.98* (*df*=16)	5.541 (*df*=2)	.441	.182	88.10%	8.40	20.00%
School Level	158.31* (*df*=15)	1.453 (*df*=3)	.614	.253	90.81%	10.88	.00%
Treatment Location	161.17* (df=16)	.750 (*df*=2)	.597	.237	91.28%	11.47	.00%
Geographical Region	131.68* (*df*=14)	3.945 (*df*=4)	.557	.238	90.30%	10.31	.00%
Grade Level	96.73* (*df*=12)	8.464 (*df*=6)	.480	.227	88.19%	8.47	12.97%

*.Estimate is significant at alpha level .05.

[a]. Test for residual heterogeneity.

[b]. Overall omnibus test for moderators.

[c]. Estimated amount of residual heterogeneity.

[d]. Standard error for Q_E.

[e]. Residual heterogeneity/accounted variability.

[f]. Unaccounted variability/sampling variability.

[g]. Amount of heterogeneity accounted for.

Even though the mean effect size estimates for the reference groups were statistically significant, the difference between the reference groups and the other groups of interests in each moderator variable were not statistically significant (see Table 2). 95% of CI estimates for the slope parameters captured the zero point and indicating that those estimates were not also statistically significant. The mean effect size estimates obtained from the dissertation and journal articles, for example, did not show statistical difference from each other since the estimate of the slope parameter (b_1) was between -1.153 and .231. However, the mean effect size estimate obtained from the studies conducted in college or university showed some statistical difference from the other school levels since the associated slope parameter was statistically significant.

Table 2. Parameter estimates in mixed-effects modeling

Moderators	*Parameter**	*Estimate*	*SE*	*z-value*	*p-value*	*CI-LL*	*CI-UL*
Study Type[a]	b_0	1.779	.248	7.188	< .05	1.294	2.265
	b_1	-.461	.353	-1.305	0.192	-1.153	.231
Treatment Time[b]	b_0	1.515	.251	6.037	< .05	1.023	2.007
	b_1	.311	.349	.889	.374	-.374	.995
	b_2	-.987	.557	-1.774	.076	-2.079	.104
School Level[c]	b_0	1.645	.214	6.816	< .05	1.722	2.119
	b_1	-.723	.618	-1.169	.242	-1.934	.489
	b_2	.098	.538	-.182	.856	-1.154	.958
	b_3	.070	.624	.112	.911	-1.153	1.293
Treatment Location[d]	b_0	1.702	.295	5.775	< .05	1.124	2.279
	b_1	-.704	.875	-.804	.421	-2.418	1.011
	b_2	-.201	.390	-.515	.607	-.965	.563
Geographical Region[e]	b_0	1.089	.395	2.758	<.05	.315	1.866
	b_1	.940	.694	1.354	.176	-.421	2.300
	b_2	.349	.475	.734	.463	-.582	1.280
	b_3	.936	.607	1.544	.123	-.252	2.125
	b_4	1.098	.893	1.230	.219	-.652	2.849
Grade Level[f]	b_0	1.276	.376	3.396	<.05	.540	2.013
	b_1	.244	.527	.462	.644	-.790	1.277
	b_2	.614	.504	1.217	.224	-.375	1.603
	b_3	.177	.647	.274	.784	-1.091	1.445
	b_4	-.278	.838	-.332	.740	-1.921	1.365
	b_5	-.368	.634	-.580	.562	-1.611	.876
	b_6	1.866	.821	2.272	<.05	.256	3.476

*. b_0 is defined as the model intercept which indicates the mean effect size estimate for the reference group. b_k (k = 1, 2, …, 6) is the slope parameter for the other levels of moderator variables, which indicates the difference in the mean effect size estimate between the reference group and the other levels of the moderator variables. See, Table 3 in Appendix A for the levels of the each moderator variables.

[a]. In study type, journal is specified as the reference group and the dissertation is the other level of the study type variable.

[b]. Studies with less than 30 days treatment time is the reference group.

[c]. Elementary school level is specified as the reference group.

[d]. Studies conducted in rural areas are classified in the reference group.

[e]. Studies conducted in the Aegean region in Turkey are classified as the reference group.

[f]. Studies implemented for the sixth grade students were classified as the reference group.

Publication Bias

Publication bias is an indication of unrepresentative findings from a readily available study of that population of completed research, which typically puts readers and researchers in danger of drawing wrong conclusion about the actual findings that can be found in that body of research, and thus, becomes a threat to validity of research findings (Rothstein, Sutton, & Borenstein, 2005). Funnel plots are primarily used in meta-analysis for the investigation of publication bias even though it is one of the potential reasons of asymmetry seen in funnel plots (Sterne, Becker, & Egger, 2005). Several measures – standard error, sampling variance, or inverse of those measures – can be plotted against the effect size of interests, and then, asymmetry in the plots is inspected to determine certain forms of publication bias (Viechtbauer, 2010).

Presence of publication bias can be investigated using several diagnostic procedures which assess the degree of asymmetry in the funnel plots, such as regressions test (Egger et al., 1997), trim and fill method (Duval & Tweedie, 2000a; 2000b), file drawer analysis (Fail-Safe *N*) computation using Rosenthal (1979)'s, Orwin (1983)'s, and Rosenberg (2005)'s methods. Sterne and Egger (2010) recommended use of standard error against standardized mean differences in funnel plots if the main objective of a meta-analysis is to examine the treatment effects. Therefore, the existence of publication was examined by taking their recommendation into consideration in this study. The funnel plot was created with trim and fill method. As seen in Figure 5, only one study was missing on the left side of the funnel plot which makes the funnel plot approximately symmetric. Since the moderator analysis showed non-significant results as presented in Table 1, the trim and fill method was only applied to the initial random-effects meta-analytic model from which the overall effect size estimate was found 1.55 [1.20-1.91]. Using Rosenthal (1979)'s file drawer analysis, we found that a total of 4646 studies that would average the null results should be added to this meta-analysis in order to reduce the combined significance level of the effect size to the pre-defined alpha level of .05.

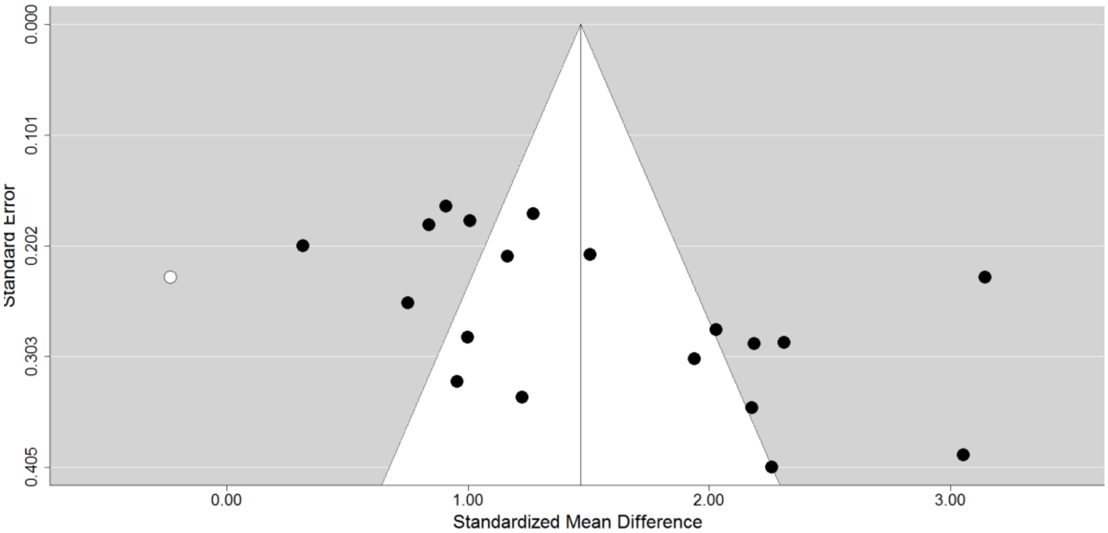


Figure 5. The funnel plot with trim and fill method

Tendency in publishing studies with significant result is the main reason of publication bias. From the selected studies for this meta-analysis, we have found only one study from which the effect size estimate was not significantly different from 0. Significant effect size estimates with large magnitudes were obtained from the rest of the studies. However, this result can be understood at some degree since many of them were either master thesis (n = 8) or doctoral dissertations (n = 7). It is expected that thesis and dissertations tend to show significant results to be approved by the review committee. Therefore, it may be practically hard to control this situation since published studies are reachable after they are published. Journal editors and reviewers may stop publication bias by encouraging authors or researchers for submitting their work with or without significant statistical results. In ideal case, registering every study undertaken regardless of looking for significant results may resolve this problem at some degree (Thorton & Lee, 2000).

Discussion and Conclusion

In this meta-analysis study, the overall effectiveness of hands-on activities on science achievement in various school levels within Turkish context was carefully scrutinized. The estimated effect size was found to be statistically significant at the pre-determined alpha level. The magnitude of the overall effect size estimate (Hedge`s *g* = 1.55, 95% CI = [1.20-1.91]) indicated that hands-on activities had a very large impact on students' science achievement. While the effect size estimates for each moderator variable was also statistically significant, the differences within sub-categories of moderator variables (e.g., study type, treatment duration, etc.) did not show any statistically significant impact on the variation in mean effect size estimates.

Based on our investigation of the selected studies, students in an environment where science was taught using hands-on activities during their coursework shows better academic performance than those who were taught science in a teacher-centered or solely lecture-based atmosphere. This finding is parallel with those of the previous studies showing that hands-on activities promoted students` science achievement level compared to students participated in the lecture-based traditional training (Ates & Eryilmaz, 2011; Bigler & Hanegan, 2011; Costu, Unal, & Ayas, 2007; Ekmekci & Gulacar, 2015; Gaitano & Bogner, 2011; Glasson, 1989; Randler & Hulde, 2007). During hands-on activities, students are expected to complete many tasks including manipulating variables, performing an experiment, making observation, collecting and recording data, reasoning, generating inferences, and sharing their findings with others. Along with these scientific skills, students are able to participate in activities directly. This way of teaching creates such an environment that students can gain skills with high cognitive complexity level, which is one of the characteristics of scientifically literate citizens (Schroeder, Scott, Tolson, Huang, & Lee, 2007). In case of giving students a chance to be actively involved in inquiry through hands-on activities, it should not be surprising to see that students who have the high cognitive skills show better academic performance than the others. It should be noted here that the use of hands-on activities, of course, is not the only way to teach science, but the present study based on overall effect science estimate indicated that hands-on science is one of the effective ways for students to learn science by touching and playing with materials, testing their hypothesis through experiments, sharing their observations and findings to get better understating of subject matters.

Project-based learning (PBL) creates an environment in which students take responsibility of their own learning, get motivated to work in groups, and exchange ideas with others, and draw conclusion based on their own observation in order to reach solution to real world problems (Korkmaz & Kaptan, 2001; 2002). Therefore, it should not be surprising to see hands-on activities mainly take place in this learning process to promote students' understating of subject matters in science and positively affects their academic achievement. A most recent study conducted by Ayaz and Soylemez (2015) confirms that the PBL has a large impact on students' science academic achievement (ES = 0.997). Our findings also suggest that hands-on activities be used for effective science teaching. Despite the fact that hands-on science activities make significant contributions to achievement in science education which is clear in the present meta-analysis study, still some results of previous studies did not show the value of hands-on science over virtual experimental materials (Klahr, Triona, & Williams, 2007; Smetana & Bell, 2012; Tirona & Kalhr, 2003) and paper-and-pencil tasks (Schwichow et al., 2015; Schwichow et al., 2016). Results from Schwichow et al.'s (2015) meta-analysis on control-of-variables (CVS) indicated that manual and virtual manipulation of variables did not directly impact on students understanding of CVS. In a recent study, Schwichow et al. (2016) emphasized the importance of "cognitive manipulation" in hands-on and paper-and-pencil tasks instead of physical and virtual manipulation. They implied that thinking about manipulating variables is crucial for learning CVS in hands-on tasks. For cognitive manipulation, they indicated that "a teacher can utilize hands-on, virtual, or paper-and-pencil tasks as long as the task requires thinking about the manipulation of variables and the consequences of those manipulations" (Schwichow et al., 2016, p. 18). As a result, our findings bring a new question into researchers` agenda to debate on the effectiveness of the hands-on activities in science teaching. We observed a statistically significant and also positive effect of hands-on science practices on students' academic achievement in all of the studies (articles and dissertations) selected for the meta-analysis study in favor or experimental group. Negative impact or no impact was not reported in the selected studies. This might be a reason why the magnitude of the overall effect size estimate is found to be too large in this study. We therefore feel a necessity to mention here that authors, academic journals, journal editors, and reviewers have the tendency of publishing manuscripts with significant results, which basically the main reason of publication bias. The only solution to prevent publication bias is that unpublished studies might have been included to this meta-analysis to obtain a more precise estimate of the effect size in the population of the relevant research on hands-on science. We therefore encourage researchers publish their manuscripts even if no statistically significant results are found.

Limitations of Study

This study was conducted with several limitations. Even though the normality assumption of effect size estimate was not violated, the total number of effect size estimate was not large enough to run more complex statistical model for further investigations and obtain a more precise estimate of the overall effect size estimate. This analysis was limited with studies conducted only in Turkey and does not include all research on hands-on instruction in other countries. Because we analyzed studies that met our analysis criteria, studies in which no control groups were used and/or in which no adequate statistics reported were not included into this meta-analysis. Another limitation was that we focused on hands-on activities and science achievement. However, we could not use other variables such as interest, motivation, science process skills and so on in this analysis since the number of those studies focusing on other variables was not adequate to run meta-analysis with a different outcome variable.

Implications

Since this study has been conducted using the available studies in Turkey, further studies should investigate the same inquiry in other countries to compare the findings with this study. Even, the trend of such research area could also be investigated through sampling of the studies on Hands-on Science from the international literature to draw general picture. The reform attempts in Turkish Education System since 2004 put a considerable attention on student-centered teaching, constructivism, hands-on and minds-on activities, and inquiry based science practices. Results of this study point out to what extend hands-on activities are affective in increasing science achievement and in reaching scientific literacy which present results confirm the value of this step in the general reform. Results that will be obtained from further studies will add new evidences to the debate of the effectiveness of hands-on science in achievement. Thus, educators and teachers would get a better picture of the effectiveness of hands-on activities in science teaching and how it impacts the students` academic achievement.

In this study, we could not include the effects of hands-on activities on some variables including interest, motivation, science process skills on the science achievement due to the lack of adequate research. Further studies are required to report the effects of hands-on on academic achievement through those variables. Since it has been indicated by several studies (Holstermann et al., 2010; Ozel, Caglak, & Erdogan, 2013; Randler & Hulde, 2007) that students` interest in science is an important factor that affects their achievement and science-related future career plans, the use of hands-on activities may help to increase students` interests in subject matters. As for the student group, even though variety of student group was included in the selected studies, no study was observed undertaken with the elementary school students. Unfortunately, it seems that hands-on activities are not frequently used in elementary schools even though students may develop more interest in science at their early ages. However, the results of the present study is the evidence of that the use of hands-on activities should be considered as an effective science teaching strategy to prepare the scientifically literate citizens for the future. In other words, as educators we are responsible to employ, or at least to encourage others for the use of hands-on activities during science teaching in order to accomplish the goal of the science literacy. Therefore, not only the use of hands-on activities in science teaching, but also any other strategy or activity that helps students to develop interests in science and also help them improve their cognitive skills should be included in school curricula in order to prepare our students for the future.

References

American Association for the Advancement of Science (AAAS). (1994). *Benchmarks for science literacy*. New York: Oxford University Press.

Ateş, Ö., & Eryılmaz, A. (2011). Effectiveness of hands-on and minds-on activities on students ‘achievement and attitudes towards physics. *Asia-Pacific Forum on Science Learning and Teaching, 12*(1), 3-22.

Ayaz, M. F., & Soylemez, M. (2015). The effect of the project-based learning approach on the academic achievements of the students in science classes in turkey: A meta-analysis study. *Education and Science, 40* (178), 225-283.

Bigler, A. M., & Hanegan, N. L. (2011). Student content knowledge increases after participation in a hands-on biotechnology intervention. *Journal of Science Education and Technology, 20*(3), 246-257.

Bilgin, I. (2006). The effects of hands-on activities incorporating a cooperative learning approach on eight grade students' science process skills and attitudes toward science. *Journal of Baltic Science Education, 1*(9), 27-37.

Birge, R. T. (1932). The calculation of errors by the method of least squares. *Physical Review, 40*, 207–227.

Borenstein, M., & Higgins, J. P. T. (2013). Meta-analysis and subgroups. *Prevention Science, 14*(2), 134-143.

Borenstein, M., Hedges, L. V., Higgins, J. P. T, & Rothstein, H. R. (2009). *Introduction to Meta- Analysis.* UK: John Willey & Sons.

Bulunuz, M. (2012). Motivational qualities of hands-on science activities for Turkish pre-service kindergarten teachers. *Eurasia Journal of Mathematics, Science & Technology Education, 8*(2), 73-82

Costu, B., Ünal, S., & Ayas, A. (2007). A hands-on activity to promote conceptual change about mixtures and chemical compounds. *Journal of Baltic Science Education, 6*(1), 35-46.

Duval, S. J., & Tweedie, R. L. (2000a). A nonparametric `Trim and Fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89-98.

Duval, S. J., & Tweedie, R. L. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629-634.

Ekmekci, A., & Gulacar, O. (2015). A case study for comparing the effectiveness of a computer simulation and a hands-on activity on learning electric circuits. *Eurasia Journal of Mathematics, Science & Technology Education, 11*(4), 765-775.

Erentay, N., & Erdogan, M. (2009). *22 adımda doğa eğitimi*. Ankara: ODTÜ yayıncılık.

Flick, L. B. (1993). The meanings of hands-on science. *Journal of Science Teacher Education, 4*(1), 1-8.

Haury, D. L. & Rillero, P. (1994). *Perspectives of hands-on science teaching*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Higgins, J.P.T. & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: John Wiley & Sons, Ltd.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539–1558.

Holstermann, N., Grube, D., & Bögeholz, S. (2010). Hands-on activities and their influence on students' interest. *Research in Science Education, 40*(5), 743-757.

Huedo-Medina, T.B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychogical Methods, 11*(2), 193-206.

Glass,G.V. (1976). Primary, secondary and meta-analysis. *Educational Researcher, 5*, 3-8.

Glasson, G. E. (1989). The effects of hands-on and teacher demonstration laboratory methods on science achievement in relation to reasoning ability and prior knowledge. *Journal of Research in Science Teaching, 26*, 121–131.

Gaitano, F., & Bogner, F. X. (2011). Cognitive influences of students' alternative conceptions within a hands-on gene technology module. *The Journal of Educational Research, 104*(3), 158-170.

Kalender, I., & Berberoglu, G. (2009). An assessment of factors related to science achievement of Turkish students. *International Journal of Science Education, 31*(10), 1379-1394.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching, 44*, 183-203.

Koç, A., & Böyük, U. (2012). The effect of hands-on science experiments on attitude towards science. *Journal of Turkish Science Education, 9*(4), 102-118.

Konstantopoulos, S., & Hedges, L.V. (2009). Analyzing effect sizes: Fixed-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Ed.). *The handbook of research synthesis and meta-analysis 2nd Ed. (*pp. 17-37). Russell Sage Foundation: New York, NY.

Korkmaz, H., & Kaptan, F. (2001). Fen eğitiminde proje tabanlı öğrenme yaklaşımı. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 20*, 193-200.

Korkmaz, H., & Kaptan, F. (2002). Fen eğitiminde proje tabanlı öğrenme yaklaşımının ilköğretim öğrencilerinin akademik başarı, akademik benlik kavramı ve çalışma sürelerine etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 22*, 91-97.

Ministry of National Education of Turkey (MoNE)(2006). *Science and technology curriculum of elementary students (6th-8th grades)*. Board of Education, Ankara.

Ministry of National Education of Turkey (MoNE)(2013). *Science curriculum of elementary students (3rd-8th grades)*. Board of Education, Ankara.

National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies.

NCES (National Center for Education Statistics). (2011). The nation's report card: Science 2011 (NCES 2012-465). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, 8, 157-159.

Ozel, M., Caglak, S., & Erdogan, M. (2013). Are affective factors a good predictor of science achievement? Examining the role of affective factors based on PISA 2006. *Learning and Individual Differences, 24*, 73-82.

Paris, S. G., Yambor, K. M., & Packard, B. W. L. (1998). Hands-on biology: A museum-school-university partnership for enhancing students' interest and learning in science. *The Elementary School Journal, 98*(3), 267-288.

Pigott. T. D. (2012). *Advances in meta-analysis*. New York, NY: Springer Science+Business Media, LLC.

Randler, C., & Hulde, M. (2007). Hands-on versus teacher-centered experiments in soil ecology. *Research in Science & Technological Education, 25*(3), 329-338.

R Software Core Team (2015). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

Rosenberg, M. S. (2005). The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution*, 59, 464-468.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638-641.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessments and adjustments*. Hoboken, NJ: John Wiley & Sons, Ltd.

Sadi, Ö., & Cakiroglu, J. (2011). Effects of hands-on activity enriched instruction on students' achievement and attitudes towards science. *Journal of Baltic Science Education, 10*(2), 87-97.

Scharfenberg, F. J., & Bogner, F. X. (2011). A new two-step approach for hands-on teaching of gene technology: effects on students’ activities during experimentation in an outreach gene technology lab. *Research in Science Education*, *41*(4), 505-523.

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T. Y., & Lee, Y. H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, *44*(10), 1436-1460.

Schwichow, M., Croker, S., Zimmerman, C., Höffler, T., & Härtig, H. (2015). Teaching the control-of-variables strategy: A meta-analysis. *Developmental Review*, *39,* 37-63.

Schwichow, M., Zimmerman, C., Croker, S., & Härtig, H. (2016). What students learn from hands-on activities? *Journal of Research in Science Teaching*, *53*(7), 980-1002.

Smetana, L. K., & Bell, R. L. (2012). Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education*, *34*(9), 1337-1370.

Sterne, J. A. C., & Egger, M. (2010). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology, 54*(10), 1046-1055.

Sterne, J.A.C., Becker, B. J., & Egger, M. (2005). The funnel plot. *Publication Bias in Meta-analysis – Prevention, Assessment, and Adjustment.* R. Rothstein, A.J. Sutton, and M. Borenstein (Ed.). Hoboken, NJ: John Wiley & Sons, Ltd.

Thorton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology, 53*, 207–216.

Triona, L. M., & Klahr, D. (2003). Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cognition and Instruction*, *21*(2), 149-173.

Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics, 30*(3), 261–293.

Wang, M. C., & Bushman, B. J. (1998). Using the normal quantile plot to explore the meta-analytic data sets. *Psychological Methods, 3*(1), 46-54.

Wiechtbauer, W. (2010). Conducting Meta-Analysis in R with the metafor package. *Journal of Statistical Software, 3*(36), 1-48.

Wiechtbauer, W. (2015). Metafor: Meta-analysis package for R. R package version 1.9-8. Retrieved from https://cran.rproject.org/package=metafor.

Appendix A

1. Effect Size Estimates for Each Subcategory

Variable and Subgroups	*Parameter*^{a}	*N*^{b}	*ES*^{c}	*SE*^{d}	*LL*^{e}	*UL*^{f}	Variable and Subgroups	*Parameter*^{a}	*N*^{b}	*ES*^{c}	*SE*^{d}	*LL*^{e}	*UL*^{f}
Study Type							*Geographical Region*						
Journal	b_0	9	1.32	.250	.83	1.81	The Mediterranean	b_0	1	2.19	.291	1.62	2.76
Dissertation	b_1	10	1.78	.250	1.29	2.27	The Marmara	b_1	3	2.02	.704	.64	3.40
School Level							The Central Anatolia	b_2	9	1.43	.235	.97	1.89
Middle School	b_0	13	1.58	.189	1.21	1.95	The Black Sea	b_3	2	2.08	.944	.23	3.93
High School	b_1	5	1.13	.306	.54	1.73	The Aegean	b_4	4	1.08	.170	.85	1.52
College or University	b_2	1	3.14	.230	2.69	3.59	*Grade Level*						
Treatment Time							Sophomore	b_0	1	3.14	.230	2.69	3.59
Less than 30 Days	b_0	8	1.52	.265	1.00	2.04	Grade 11	b_1	2	.91	.592	-.26	2.07
Between 30 and 60 Days	b_1	9	1.83	.240	1.35	2.30	Grade 10	b_2	1	1.00	.286	.44	1.56
More than 60 days	b_2	2	.51	.214	.08	.93	Grade 9	b_3	2	1.47	.668	.16	2.78
Treatment Location							Grade 8	b_4	5	1.91	.388	1.14	2.67
Urban	b_0	10	1.51	.296	.93	2.09	Grade 7	b_5	4	1.52	.342	.85	2.19
Suburban	b_1	1	1.00	.286	.44	1.56	Grade 6	b_6	4	1.22	.245	.82	1.70
Rural	b_2	8	1.70	.224	1.26	2.14							

[a] Parameter estimated in Mixed-effects Model

[b] Number of studies in each subcategory

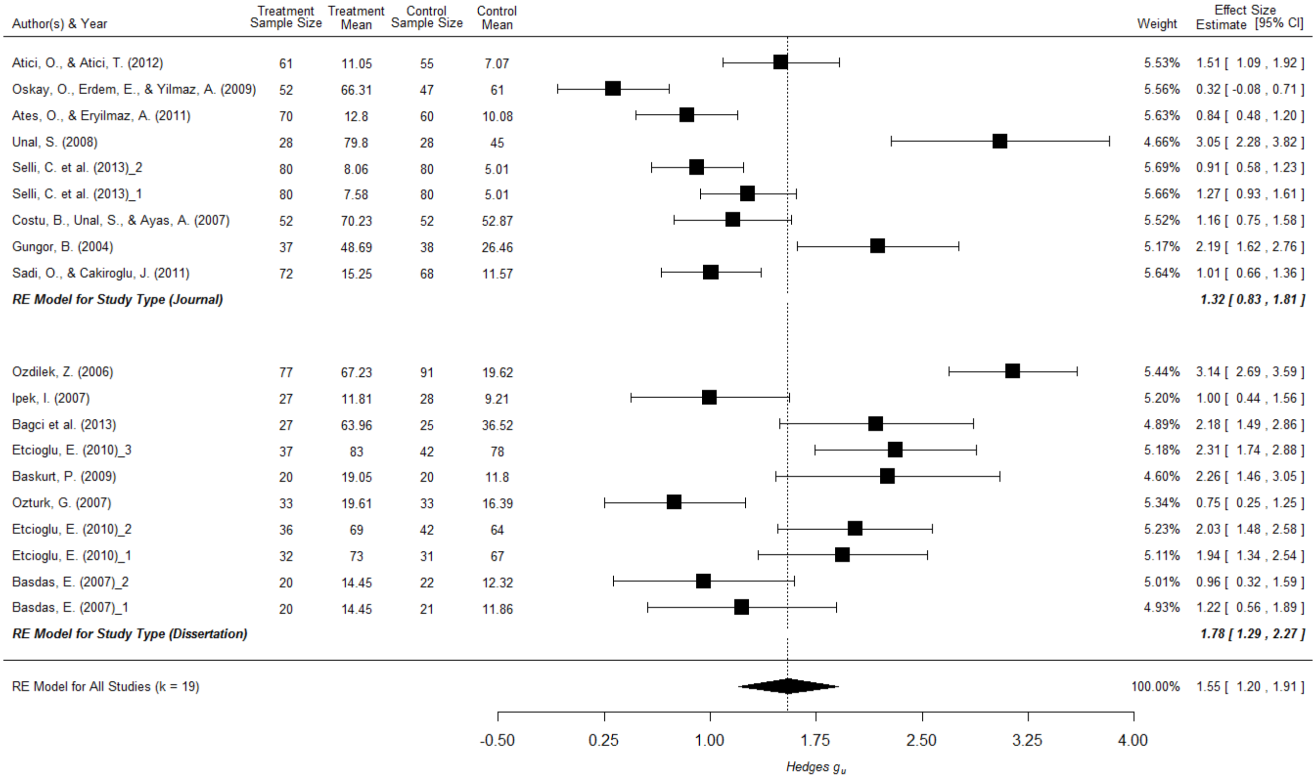
[c] Hedge's *g* effect size estimate

[d] Standard errors of effect size estimation

[e] Lower limit of 95% CI of effect size estimate

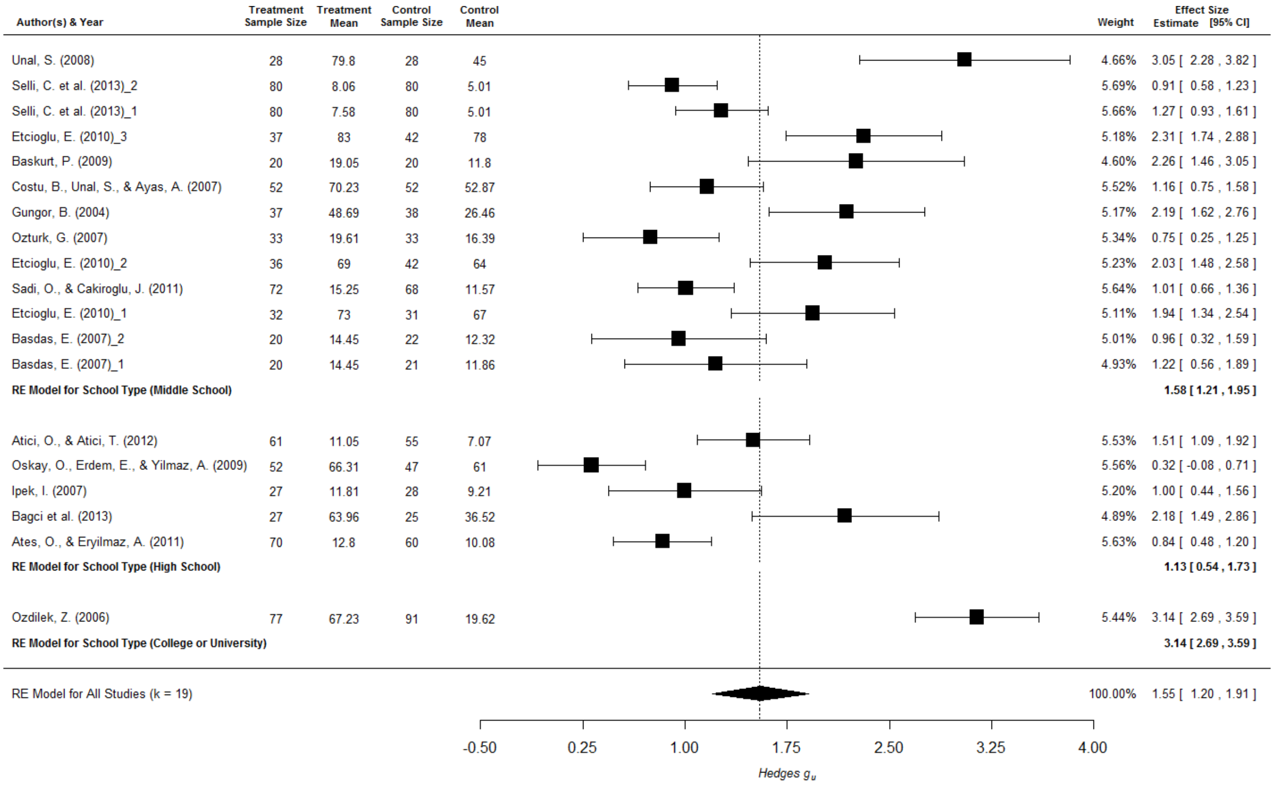
[f] Upper limit of 95% CI of effect size estimate

2. Forest Plot for the Moderator Variable *Study Type*



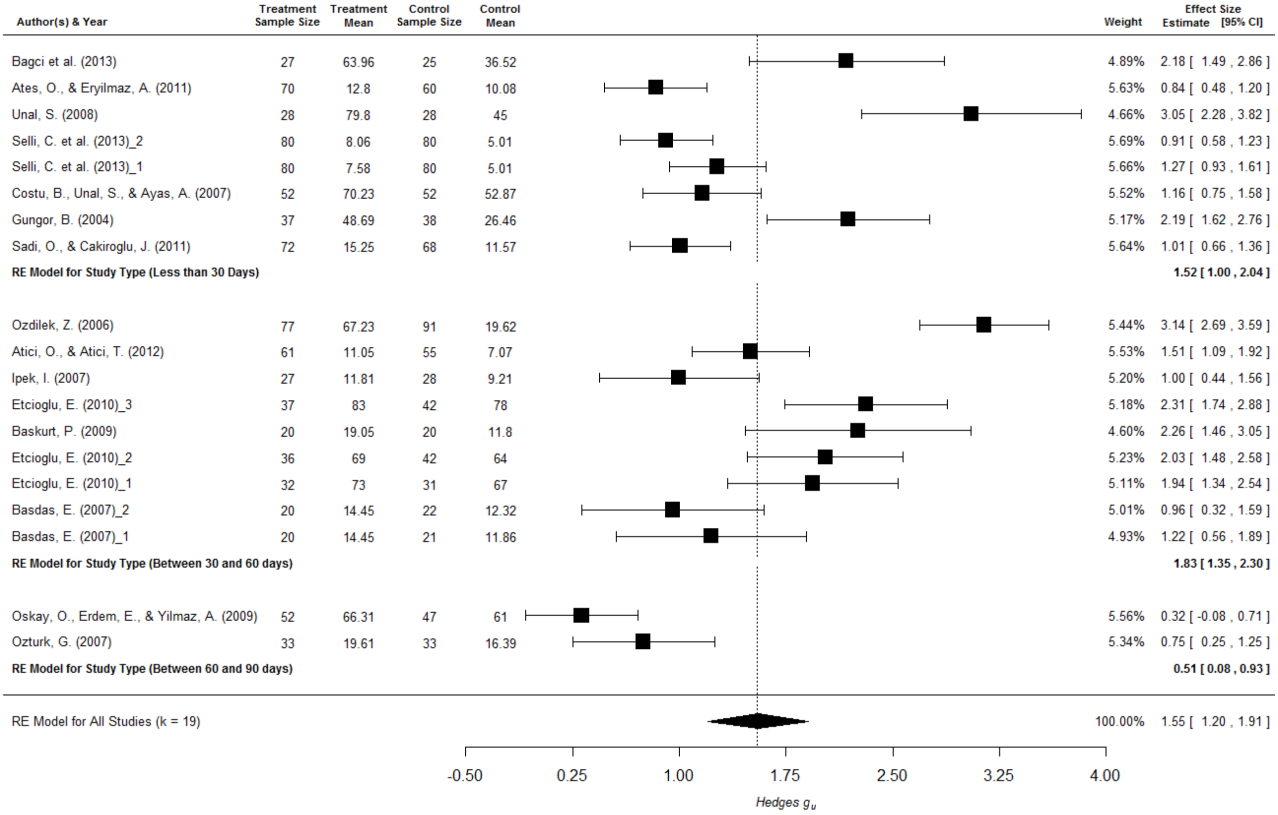
Author(s) & Year	Treatment Sample Size	Treatment Mean	Control Sample Size	Control Mean	Weight	Effect Size Estimate [95% CI]
Atici, O., & Atici, T. (2012)	61	11.05	55	7.07	5.53%	1.51 [1.09 , 1.92]
Oskay, O., Erdem, E., & Yilmaz, A. (2009)	52	66.31	47	61	5.56%	0.32 [-0.08 , 0.71]
Ates, O., & Eryilmaz, A. (2011)	70	12.8	60	10.08	5.63%	0.84 [0.48 , 1.20]
Unal, S. (2008)	28	79.8	28	45	4.66%	3.05 [2.28 , 3.82]
Selli, C. et al. (2013)_2	80	8.06	80	5.01	5.69%	0.91 [0.58 , 1.23]
Selli, C. et al. (2013)_1	80	7.58	80	5.01	5.66%	1.27 [0.93 , 1.61]
Costu, B., Unal, S., & Ayas, A. (2007)	52	70.23	52	52.87	5.52%	1.16 [0.75 , 1.58]
Gungor, B. (2004)	37	48.69	38	26.46	5.17%	2.19 [1.62 , 2.76]
Sadi, O., & Cakiroglu, J. (2011)	72	15.25	68	11.57	5.64%	1.01 [0.66 , 1.36]
RE Model for Study Type (Journal)						***1.32 [0.83 , 1.81]***
Ozdilek, Z. (2006)	77	67.23	91	19.62	5.44%	3.14 [2.69 , 3.59]
Ipek, I. (2007)	27	11.81	28	9.21	5.20%	1.00 [0.44 , 1.56]
Bagci et al. (2013)	27	63.96	25	36.52	4.89%	2.18 [1.49 , 2.86]
Etcioglu, E. (2010)_3	37	83	42	78	5.18%	2.31 [1.74 , 2.88]
Baskurt, P. (2009)	20	19.05	20	11.8	4.60%	2.26 [1.46 , 3.05]
Ozturk, G. (2007)	33	19.61	33	16.39	5.34%	0.75 [0.25 , 1.25]
Etcioglu, E. (2010)_2	36	69	42	64	5.23%	2.03 [1.48 , 2.58]
Etcioglu, E. (2010)_1	32	73	31	67	5.11%	1.94 [1.34 , 2.54]
Basdas, E. (2007)_2	20	14.45	22	12.32	5.01%	0.96 [0.32 , 1.59]
Basdas, E. (2007)_1	20	14.45	21	11.86	4.93%	1.22 [0.56 , 1.89]
RE Model for Study Type (Dissertation)						***1.78 [1.29 , 2.27]***
RE Model for All Studies (k = 19)					100.00%	1.55 [1.20 , 1.91]

3. Forest Plot for the Moderator Variable *School Level*



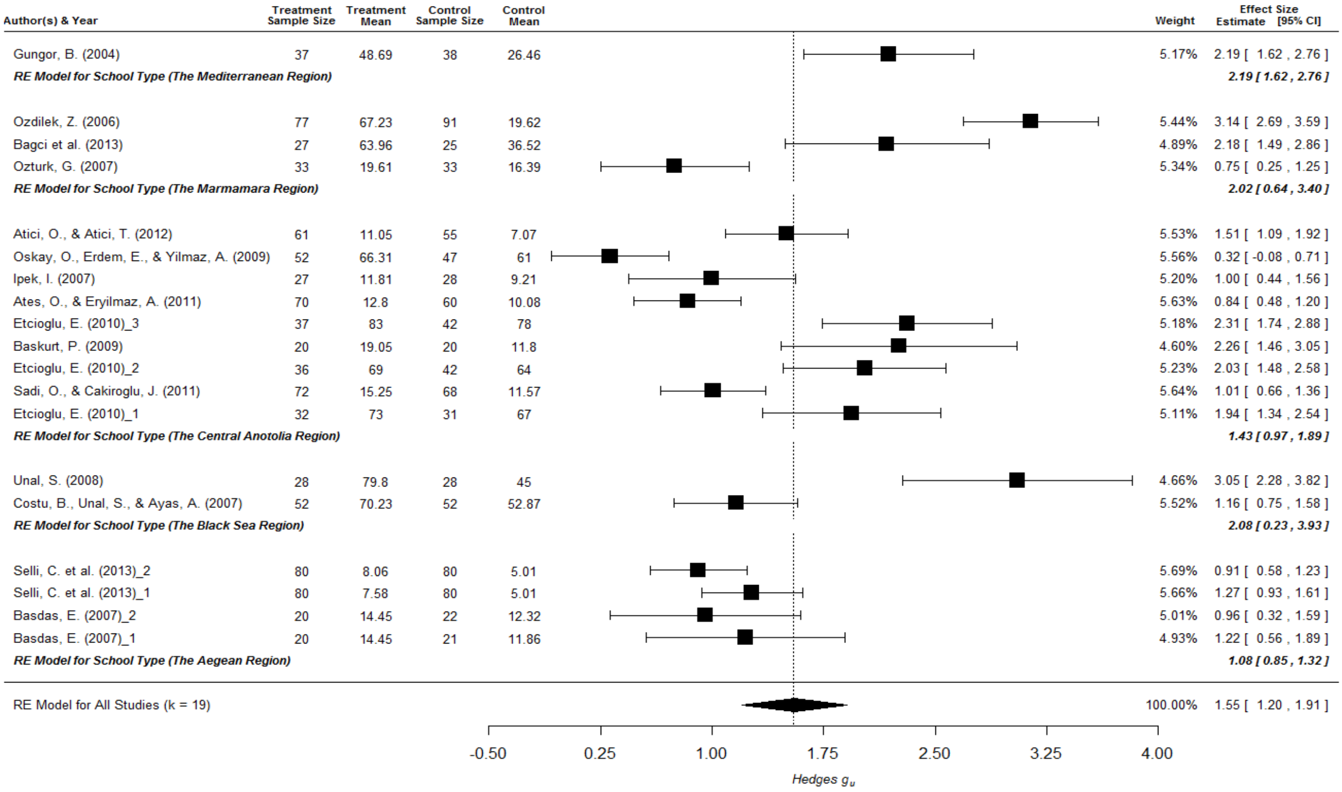
Author(s) & Year	Treatment Sample Size	Treatment Mean	Control Sample Size	Control Mean	Weight	Effect Size Estimate [95% CI]
Unal, S. (2008)	28	79.8	28	45	4.66%	3.05 [2.28 , 3.82]
Selli, C. et al. (2013)_2	80	8.06	80	5.01	5.69%	0.91 [0.58 , 1.23]
Selli, C. et al. (2013)_1	80	7.58	80	5.01	5.66%	1.27 [0.93 , 1.61]
Etcioglu, E. (2010)_3	37	83	42	78	5.18%	2.31 [1.74 , 2.88]
Baskurt, P. (2009)	20	19.05	20	11.8	4.60%	2.26 [1.46 , 3.05]
Costu, B., Unal, S., & Ayas, A. (2007)	52	70.23	52	52.87	5.52%	1.16 [0.75 , 1.58]
Gungor, B. (2004)	37	48.69	38	26.46	5.17%	2.19 [1.62 , 2.76]
Ozturk, G. (2007)	33	19.61	33	16.39	5.34%	0.75 [0.25 , 1.25]
Etcioglu, E. (2010)_2	36	69	42	64	5.23%	2.03 [1.48 , 2.58]
Sadi, O., & Cakiroglu, J. (2011)	72	15.25	68	11.57	5.64%	1.01 [0.66 , 1.36]
Etcioglu, E. (2010)_1	32	73	31	67	5.11%	1.94 [1.34 , 2.54]
Basdas, E. (2007)_2	20	14.45	22	12.32	5.01%	0.96 [0.32 , 1.59]
Basdas, E. (2007)_1	20	14.45	21	11.86	4.93%	1.22 [0.56 , 1.89]
RE Model for School Type (Middle School)						**1.58 [1.21 , 1.95]**
Atici, O., & Atici, T. (2012)	61	11.05	55	7.07	5.53%	1.51 [1.09 , 1.92]
Oskay, O., Erdem, E., & Yilmaz, A. (2009)	52	66.31	47	61	5.56%	0.32 [-0.08 , 0.71]
Ipek, I. (2007)	27	11.81	28	9.21	5.20%	1.00 [0.44 , 1.56]
Bagci et al. (2013)	27	63.96	25	36.52	4.89%	2.18 [1.49 , 2.86]
Ates, O., & Eryilmaz, A. (2011)	70	12.8	60	10.08	5.63%	0.84 [0.48 , 1.20]
RE Model for School Type (High School)						**1.13 [0.54 , 1.73]**
Ozdilek, Z. (2006)	77	67.23	91	19.62	5.44%	3.14 [2.69 , 3.59]
RE Model for School Type (College or University)						**3.14 [2.69 , 3.59]**
RE Model for All Studies (k = 19)					100.00%	1.55 [1.20 , 1.91]

4. Forest Plot for the Moderator Variable Treatment Time



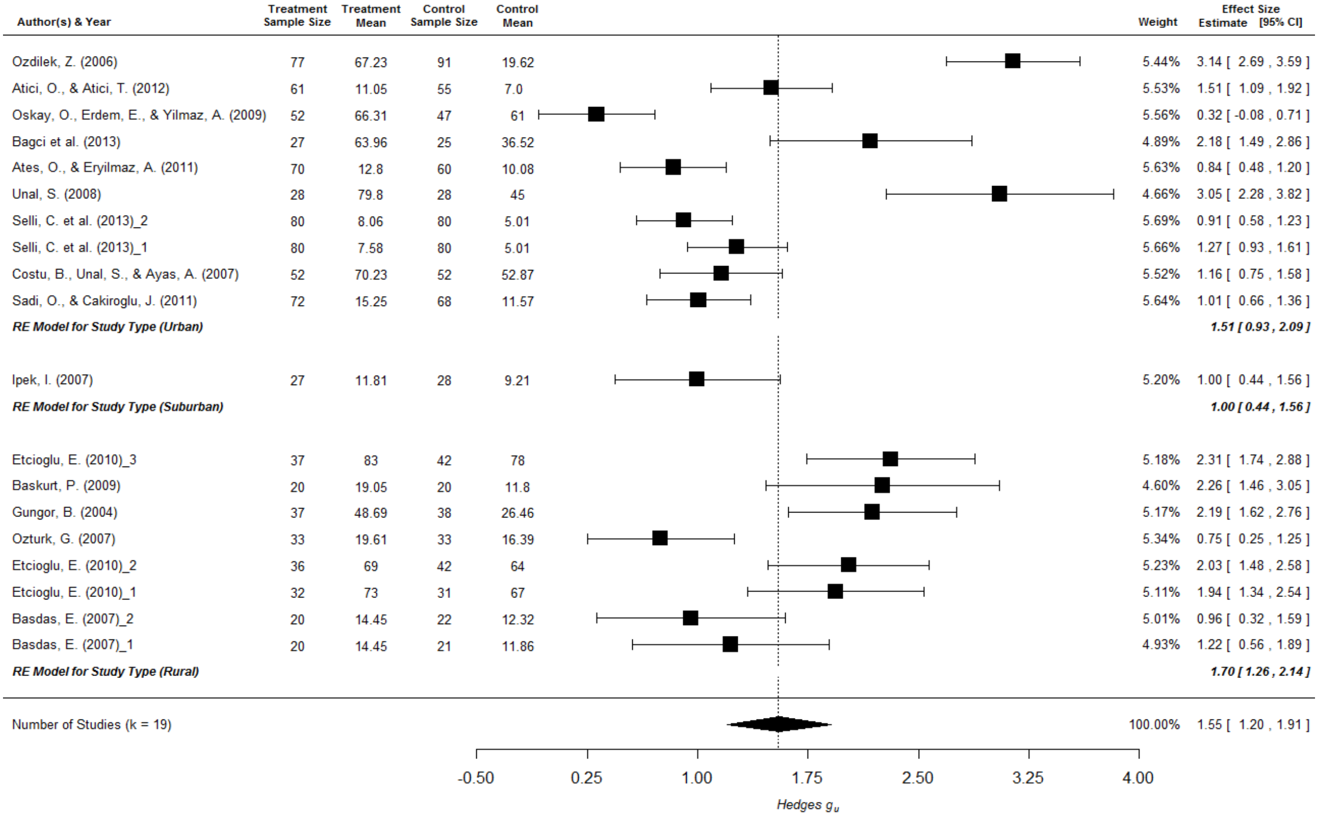
Author(s) & Year	Treatment Sample Size	Treatment Mean	Control Sample Size	Control Mean	Weight	Effect Size Estimate [95% CI]
Bagci et al. (2013)	27	63.96	25	36.52	4.89%	2.18 [1.49 , 2.86]
Ates, O., & Eryilmaz, A. (2011)	70	12.8	60	10.08	5.63%	0.84 [0.48 , 1.20]
Unal, S. (2008)	28	79.8	28	45	4.66%	3.05 [2.28 , 3.82]
Selli, C. et al. (2013)_2	80	8.06	80	5.01	5.69%	0.91 [0.58 , 1.23]
Selli, C. et al. (2013)_1	80	7.58	80	5.01	5.66%	1.27 [0.93 , 1.61]
Costu, B., Unal, S., & Ayas, A. (2007)	52	70.23	52	52.87	5.52%	1.16 [0.75 , 1.58]
Gungor, B. (2004)	37	48.69	38	26.46	5.17%	2.19 [1.62 , 2.76]
Sadi, O., & Cakiroglu, J. (2011)	72	15.25	68	11.57	5.64%	1.01 [0.66 , 1.36]
RE Model for Study Type (Less than 30 Days)						**1.52 [1.00 , 2.04]**
Ozdilek, Z. (2006)	77	67.23	91	19.62	5.44%	3.14 [2.69 , 3.59]
Atici, O., & Atici, T. (2012)	61	11.05	55	7.07	5.53%	1.51 [1.09 , 1.92]
Ipek, I. (2007)	27	11.81	28	9.21	5.20%	1.00 [0.44 , 1.56]
Etcioglu, E. (2010)_3	37	83	42	78	5.18%	2.31 [1.74 , 2.88]
Baskurt, P. (2009)	20	19.05	20	11.8	4.60%	2.26 [1.46 , 3.05]
Etcioglu, E. (2010)_2	36	69	42	64	5.23%	2.03 [1.48 , 2.58]
Etcioglu, E. (2010)_1	32	73	31	67	5.11%	1.94 [1.34 , 2.54]
Basdas, E. (2007)_2	20	14.45	22	12.32	5.01%	0.96 [0.32 , 1.59]
Basdas, E. (2007)_1	20	14.45	21	11.86	4.93%	1.22 [0.56 , 1.89]
RE Model for Study Type (Between 30 and 60 days)						**1.83 [1.35 , 2.30]**
Oskay, O., Erdem, E., & Yilmaz, A. (2009)	52	66.31	47	61	5.56%	0.32 [-0.08 , 0.71]
Ozturk, G. (2007)	33	19.61	33	16.39	5.34%	0.75 [0.25 , 1.25]
RE Model for Study Type (Between 60 and 90 days)						**0.51 [0.08 , 0.93]**
RE Model for All Studies (k = 19)					100.00%	1.55 [1.20 , 1.91]

5. Forest Plot for the Moderator Variable *Geographical Region*



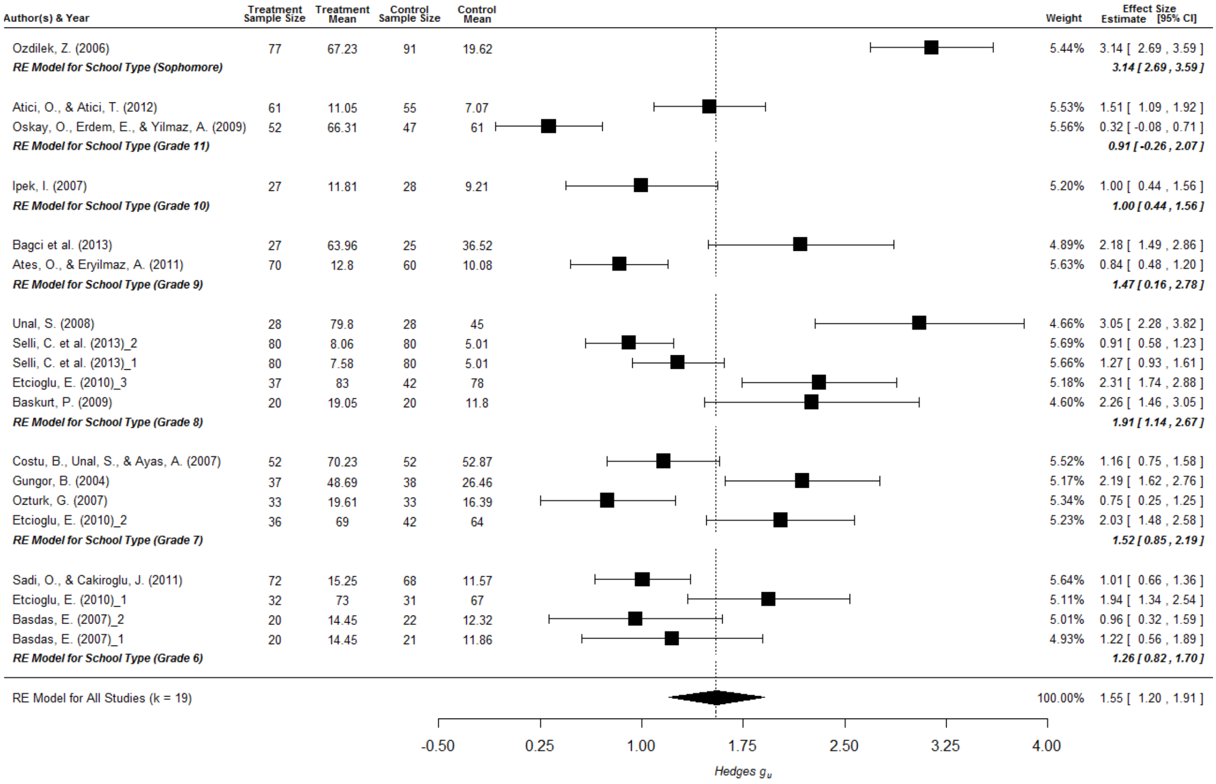
Author(s) & Year	Treatment Sample Size	Treatment Mean	Control Sample Size	Control Mean	Weight	Effect Size Estimate [95% CI]
Gungor, B. (2004)	37	48.69	38	26.46	5.17%	2.19 [1.62 , 2.76]
RE Model for School Type (The Mediterranean Region)						***2.19 [1.62 , 2.76]***
Ozdilek, Z. (2006)	77	67.23	91	19.62	5.44%	3.14 [2.69 , 3.59]
Bagci et al. (2013)	27	63.96	25	36.52	4.89%	2.18 [1.49 , 2.86]
Ozturk, G. (2007)	33	19.61	33	16.39	5.34%	0.75 [0.25 , 1.25]
RE Model for School Type (The Marmamara Region)						***2.02 [0.64 , 3.40]***
Atici, O., & Atici, T. (2012)	61	11.05	55	7.07	5.53%	1.51 [1.09 , 1.92]
Oskay, O., Erdem, E., & Yilmaz, A. (2009)	52	66.31	47	61	5.56%	0.32 [-0.08 , 0.71]
Ipek, I. (2007)	27	11.81	28	9.21	5.20%	1.00 [0.44 , 1.56]
Ates, O., & Eryilmaz, A. (2011)	70	12.8	60	10.08	5.63%	0.84 [0.48 , 1.20]
Etcioglu, E. (2010)_3	37	83	42	78	5.18%	2.31 [1.74 , 2.88]
Baskurt, P. (2009)	20	19.05	20	11.8	4.60%	2.26 [1.46 , 3.05]
Etcioglu, E. (2010)_2	36	69	42	64	5.23%	2.03 [1.48 , 2.58]
Sadi, O., & Cakiroglu, J. (2011)	72	15.25	68	11.57	5.64%	1.01 [0.66 , 1.36]
Etcioglu, E. (2010)_1	32	73	31	67	5.11%	1.94 [1.34 , 2.54]
RE Model for School Type (The Central Anotolia Region)						***1.43 [0.97 , 1.89]***
Unal, S. (2008)	28	79.8	28	45	4.66%	3.05 [2.28 , 3.82]
Costu, B., Unal, S., & Ayas, A. (2007)	52	70.23	52	52.87	5.52%	1.16 [0.75 , 1.58]
RE Model for School Type (The Black Sea Region)						***2.08 [0.23 , 3.93]***
Selli, C. et al. (2013)_2	80	8.06	80	5.01	5.69%	0.91 [0.58 , 1.23]
Selli, C. et al. (2013)_1	80	7.58	80	5.01	5.66%	1.27 [0.93 , 1.61]
Basdas, E. (2007)_2	20	14.45	22	12.32	5.01%	0.96 [0.32 , 1.59]
Basdas, E. (2007)_1	20	14.45	21	11.86	4.93%	1.22 [0.56 , 1.89]
RE Model for School Type (The Aegean Region)						***1.08 [0.85 , 1.32]***
RE Model for All Studies (k = 19)					100.00%	1.55 [1.20 , 1.91]

6. Forest Plot for the Moderator Variable *Treatment Location*



Author(s) & Year	Treatment Sample Size	Treatment Mean	Control Sample Size	Control Mean	Weight	Effect Size Estimate [95% CI]
Ozdilek, Z. (2006)	77	67.23	91	19.62	5.44%	3.14 [2.69 , 3.59]
Atici, O., & Atici, T. (2012)	61	11.05	55	7.0	5.53%	1.51 [1.09 , 1.92]
Oskay, O., Erdem, E., & Yilmaz, A. (2009)	52	66.31	47	61	5.56%	0.32 [-0.08 , 0.71]
Bagci et al. (2013)	27	63.96	25	36.52	4.89%	2.18 [1.49 , 2.86]
Ates, O., & Eryilmaz, A. (2011)	70	12.8	60	10.08	5.63%	0.84 [0.48 , 1.20]
Unal, S. (2008)	28	79.8	28	45	4.66%	3.05 [2.28 , 3.82]
Selli, C. et al. (2013)_2	80	8.06	80	5.01	5.69%	0.91 [0.58 , 1.23]
Selli, C. et al. (2013)_1	80	7.58	80	5.01	5.66%	1.27 [0.93 , 1.61]
Costu, B., Unal, S., & Ayas, A. (2007)	52	70.23	52	52.87	5.52%	1.16 [0.75 , 1.58]
Sadi, O., & Cakiroglu, J. (2011)	72	15.25	68	11.57	5.64%	1.01 [0.66 , 1.36]
RE Model for Study Type (Urban)						***1.51 [0.93 , 2.09]***
Ipek, I. (2007)	27	11.81	28	9.21	5.20%	1.00 [0.44 , 1.56]
RE Model for Study Type (Suburban)						***1.00 [0.44 , 1.56]***
Etcioglu, E. (2010)_3	37	83	42	78	5.18%	2.31 [1.74 , 2.88]
Baskurt, P. (2009)	20	19.05	20	11.8	4.60%	2.26 [1.46 , 3.05]
Gungor, B. (2004)	37	48.69	38	26.46	5.17%	2.19 [1.62 , 2.76]
Ozturk, G. (2007)	33	19.61	33	16.39	5.34%	0.75 [0.25 , 1.25]
Etcioglu, E. (2010)_2	36	69	42	64	5.23%	2.03 [1.48 , 2.58]
Etcioglu, E. (2010)_1	32	73	31	67	5.11%	1.94 [1.34 , 2.54]
Basdas, E. (2007)_2	20	14.45	22	12.32	5.01%	0.96 [0.32 , 1.59]
Basdas, E. (2007)_1	20	14.45	21	11.86	4.93%	1.22 [0.56 , 1.89]
RE Model for Study Type (Rural)						***1.70 [1.26 , 2.14]***
Number of Studies (k = 19)					100.00%	1.55 [1.20 , 1.91]

7. Forest Plot for the Moderator Variable *Grade Level*



Author(s) & Year	Treatment Sample Size	Treatment Mean	Control Sample Size	Control Mean	Weight	Effect Size Estimate [95% CI]
Ozdilek, Z. (2006)	77	67.23	91	19.62	5.44%	3.14 [2.69 , 3.59]
RE Model for School Type (Sophomore)						***3.14 [2.69 , 3.59]***
Atici, O., & Atici, T. (2012)	61	11.05	55	7.07	5.53%	1.51 [1.09 , 1.92]
Oskay, O., Erdem, E., & Yilmaz, A. (2009)	52	66.31	47	61	5.56%	0.32 [-0.08 , 0.71]
RE Model for School Type (Grade 11)						***0.91 [-0.26 , 2.07]***
Ipek, I. (2007)	27	11.81	28	9.21	5.20%	1.00 [0.44 , 1.56]
RE Model for School Type (Grade 10)						***1.00 [0.44 , 1.56]***
Bagci et al. (2013)	27	63.96	25	36.52	4.89%	2.18 [1.49 , 2.86]
Ates, O., & Eryilmaz, A. (2011)	70	12.8	60	10.08	5.63%	0.84 [0.48 , 1.20]
RE Model for School Type (Grade 9)						***1.47 [0.16 , 2.78]***
Unal, S. (2008)	28	79.8	28	45	4.66%	3.05 [2.28 , 3.82]
Selli, C. et al. (2013)_2	80	8.06	80	5.01	5.69%	0.91 [0.58 , 1.23]
Selli, C. et al. (2013)_1	80	7.58	80	5.01	5.66%	1.27 [0.93 , 1.61]
Etcioglu, E. (2010)_3	37	83	42	78	5.18%	2.31 [1.74 , 2.88]
Baskurt, P. (2009)	20	19.05	20	11.8	4.60%	2.26 [1.46 , 3.05]
RE Model for School Type (Grade 8)						***1.91 [1.14 , 2.67]***
Costu, B., Unal, S., & Ayas, A. (2007)	52	70.23	52	52.87	5.52%	1.16 [0.75 , 1.58]
Gungor, B. (2004)	37	48.69	38	26.46	5.17%	2.19 [1.62 , 2.76]
Ozturk, G. (2007)	33	19.61	33	16.39	5.34%	0.75 [0.25 , 1.25]
Etcioglu, E. (2010)_2	36	69	42	64	5.23%	2.03 [1.48 , 2.58]
RE Model for School Type (Grade 7)						***1.52 [0.85 , 2.19]***
Sadi, O., & Cakiroglu, J. (2011)	72	15.25	68	11.57	5.64%	1.01 [0.66 , 1.36]
Etcioglu, E. (2010)_1	32	73	31	67	5.11%	1.94 [1.34 , 2.54]
Basdas, E. (2007)_2	20	14.45	22	12.32	5.01%	0.96 [0.32 , 1.59]
Basdas, E. (2007)_1	20	14.45	21	11.86	4.93%	1.22 [0.56 , 1.89]
RE Model for School Type (Grade 6)						***1.26 [0.82 , 1.70]***
RE Model for All Studies (k = 19)					100.00%	1.55 [1.20 , 1.91]

Appendix B

Included Studies in the Meta-Analysis

Ateş, Ö., & Eryılmaz, A. (2011). Effectiveness of Hands-on and minds-on activities on students' achievement and attitudes towards physics. *Asia-Pacific Forum on Science Learning and Teaching*, *12*(1), 3-22.

Atıcı, Ö. ve Atıcı, T. (2012). Fotosentez konusunun öğretiminde uygulanan laboratuvar yönteminin öğrenci tutum ve başarısına etkisi. *The Journal of Turkish Educational Sciences*, *10*(1), 143-166.

Başdaş, E. (2007*). İlköğretim fen eğitiminde, basit malzemelerle yapılan fen aktivitelerinin bilimsel süreç becerilerine, akademik başarıya ve motivasyona etkisi*. (Unpublished Master Thesis). Celal Bayar University, Manisa, Turkey.

Başkurt, P. (2009). *İlköğretim 8. sınıf fen ve teknoloji dersi kuvvet ve hareket ünitesinin basit malzemelerle yapılan fen aktiviteleri ile öğretilmesinin başarıya, kalıcılığa ve tutuma etkisi*. (Unpublished Master Thesis). Gazi University, Ankara Turkey.

Costu, B., Unal, S., & Ayas, A. (2007). A hands-on activity to promote conceptual change about mixtures and chemical compounds. *Journal of Baltic Science Education*, *6*(1), 35-46.

Etcioğlu, Ö. (2010). *Kuvvet ve hareket ünitesine yönelik basit araç gereçlerle geliştirilen rehber materyal kullanmanın öğrenmeye etkisi ve öğrenilenlerin günlük hayata aktarılma düzeylerinin incelenmesi*. (Unpublished Master Thesis). Erzincan University, Erzincan, Turkey.

Güngör, B. (2004). *"Yönetici moleküller" konusunun öğretilmesinde deneysel yönteme göre geliştirilen öğretim tekniğinin uygulanması ve geleneksel öğretimle karşılaştırılması*. (Unpublished Master Thesis). Balıkesir University, Balıkesir, Turkey.

Hardal, Ö. (2003). *The effects of hands-on activities on ninth grade students' achievement and attitudes towards physics*. (Unpublished Master Thesis). The Middle East Technical University, Ankara, Turkey.

Inciser, I. (2007). *Implementation of conceptual change oriented instruction using hands on activities on tenth grade students' understanding of gases concepts*. (Unpublished Master Thesis). The Middle East Technical University, Ankara.

Oskay, Ö. Ö., Erdem, E. ve Yılmaz, A. (2009). Kimya laboratuvar uygulamalarının öğrencilerin kimyaya yönelik tutum ve başarılarına etkisi üzerine bir çalışma. *Electronic Journal of Social Sciences*, *8*(27), 222-321.

Öztürk, G. (2007). *Öğrencilerin basit malzemelerle yaptıkları deneylerin kuvvet-enerji kavramını öğrenmelerine ve fene karşı tutumlarına etkisi*. (Unpublished Master Thesis). Marmara University, Istanbul, Turkey.

Selli, C., Yıldırım, G., Kaymak, A., Karacicek, B., Ogut, D., Gungor, T., Erem, E., Ege, M., Bumen, N. & Tosun, M. (2014). Introducing basic molecular biology to Turkish rural and urban primary school children via hands-on pcr and gel electrophoresis activities. *Biochemistry and Molecular Biology Education*, *42*(2), 114-120.

Telli, A., Yıldırım, H. İ., Şensoy, Ö. ve Yalçın, N. (2004). İlköğretim 7. sınıflarda basit makinalar konusunun öğretiminde laboratuar yönteminin öğrenci başarısına etkisinin araştırılması. *Gazi Eğitim Fakültesi Dergisi*, *24*(3), 291-305.

Unal, S. (2008). Changing students' misconceptions of floating and sinking using hands-on activities. *Journal of Baltic Science Education*, *7*(3), 134-146.

Author Information

Serdar Caglak
Eskisehir Osmangazi University
Eskisehir, Turkey
Contact e-mail: *serdar.caglak@gmail.com*

JESEH
Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

Effect of Cooperative Learning Strategy on Students' Acquisition and Practice of Scientific Skills in Biology

Hanadi Chatila, Fatima Al Husseiny

Article Info

Article History

Received:
21 September 2016

Accepted:
18 December 2016

Keywords

Scientific skills
Cooperative learning
Biology education

Abstract

Recent research findings have shown that cooperative learning improves students' thinking skills as it allows them to communicate actively with each other (Johnson, Johnson and Smith, 2014). Therefore, cooperative learning has been proposed by many educators to be implemented in classrooms to produce lifelong learners and critical thinkers (Lunenburg, 2011). The current study investigates the effect of cooperative learning in Biology classroom, on students' learning and achievement of scientific skills. A convenient sample of 120 students from two grade levels, seven and ten, participated in the study in a private school in Beirut, where biology was taught to each class of the two different grades using two different teaching methods: cooperative learning (experimental group) and individualistic-direct learning (control group). Pre- and post- tests were administrated to both groups of each grade to compare students' achievement particularly in scientific skills items before and after intervention. Results of the study show that cooperative learning has a significant effect on students' achievement in learning and practicing scientific skills in grade ten, however no significant effect was shown in the acquisition of new scientific skills for grade seven students.

Introduction

Many studies around the world have documented students' difficulties in learning biology that may affect their motivation and achievement (Agboghoroma and Oyovwi, 2015; Bahar et al., 1999; Çimer, 2012). Those difficulties encountered by students may be attributed to two main factors: the difficulty of grasping biology concepts and "working scientifically" skills, and the overloaded curriculum.Studies have reported that students face difficulties in many abstracts concepts or topics in biology, at both high schools and university levels, including the concepts of hormones, cells, genes and chromosomes, mitosis and meiosis, and the nervous system (Agboghoroma and Oyovwi, 2015, Agorram et al, 2010, Chattopadhyay, 2005 and Tekkaya et al., 2001), in addition to other topics like water and gas transport in plants, protein synthesis, photosynthesis and respiration, gaseous exchange, energy, organs, physiological processes and oxygen transport (Çimers, 2012). Adding to the abstract content, the nature of science itself, which requires learning and applying "working scientifically skills" throughout the course of study, is a major problem for students. Another cause of difficulties is the overloaded curriculum (Tekkaya et al., 2001), which indirectly affects both teachers and students. Due to the massive and abstract content in curriculum, teachers usually take care of delivering the content regardless students' interest and motivation, which unfortunately prevents meaningful learning and results in learning materials by memorization, and therefore causes several learning problems (Zeidan, 2010).

Science education research, including biology education, conducted in the past few decades has focused on the integration of knowledge, skills and attitudes to develop a better understanding of scientific concepts (Zeidan and Jayosi, 2015.) In other words, the emphasis has been on how students learn, and how they built their personal understandings of scientific concepts. Lunenburg (2011) considered that the use of student-centered teaching strategies in classrooms within an overall inquiry-based pedagogy is an effective way to enhance students' academic performance, critical thinking, and problem solving skills. So, through inquiry, students may learn both skills and concepts, and develop positive attitudes towards science.

Chiappetta & Koballa (2010) considered that the learning and application of "science process skills" are always associated with scientific inquiry. Many factors have an impact on students' achievement such as classroom environment, attitude and motivation, and above all teaching methods and strategies. Educators around the world have been investigating various teaching strategies in science classes to improve students' outcomes. One of the most favored investigated teaching strategies among science educators is cooperative learning; it is considered as one of the most efficient instructional methods that enable students to work together in solving scientific problems, as it improves students' thinking skills and abilities, and has the potential to promote academic achievement, enhance social skills, and improve self-esteem by engaging students in an active learning environment (Vijayratnam, 2009).

Cooperative Learning and Academic Achievement

Research around the word has highlighted the effectiveness of cooperative learning in promoting deep learning and higher achievement in the classroom, especially science classroom at all levels (Johnson & Johnson, 1989; Johnson et al., 2014; Lord, 2001; Lin, 2006; Vijayratnam, 2009 and Wolfensberger and Canella, 2015). Knowing that cooperative learning encourages student involvement and engagement in their own learning, it provides all students with opportunities to make their thoughts visible to others, allows them to talk about their own ideas, and permits them to consider the ideas of others, which enhances their higher order thinking skills (Johnson et al., 2014). In the light of this, Chang and Mao (1999) noted that effective cooperative learning leads to active learning that enables students to move beyond the text, memorization of basic facts, and consequently promotes learning and practicing higher-level skills. This would lead, apart from academic benefits, to enhance learners' self-esteem, and interpersonal relationship and attitudes toward school and peers (Bilgin and Geban ,2006).

For an effective cooperative learning experience, Richards and Rodgers (2001) and Johnson et al. (2013) suggested five major premises :(1) Positive interdependence when learners work together to attain the group objective. (2) Individual accountability, when each team member is considered responsible for his or her own understanding of the work, which in turn contributes to the objective of the team. (3) Interpersonal skills (communication, trust, leadership, decision-making and conflict resolution), where team members argue, solve problems, and work together. (4) Face to face interaction and (5) Processing where learners assess and reflect on their team work ability and skills.

Therefore, it is assumed that cooperative learning has a positive effect on students' cognitive, emotional and social skills, such as promoting higher achievement, greater use of higher level critical reasoning competencies and strategies, higher self-esteem, and greater collaborative skills and attitudes necessary for working effectively with others. However, literature includes some studies reporting no significant effect on learning.

For example, Chang and Lederman (1994) applied cooperative learning to investigate students' physical science achievement and found that cooperative learning had no significant difference on students' achievement. Similar findings were reported by Seymour (1994) and Tingle & Good (1990) as cited in Chang and Mao (1999). Sadler (2002) investigated the effect of cooperative learning on college students' biology academic achievement and reported no significant difference in academic achievement and cooperative learning verses direct lecture methods.

Biology in Lebanese schooling system

In Lebanon, the Science curriculum is rich in content, details and concepts that students must learn through exploring, investigating and describing their experiences to achieve meaningful learning. There is a great emphasis on knowledge and skills in the Lebanese science curriculum, disfavoring the affective domain in education. Table 1 shows the general objectives of the Lebanese "science curriculum" (CRPD, 1997).

In the Lebanese schooling system, Biology is delivered separately from other sciences (chemistry and physics) at the intermediate level starting from grade seven up to grade nine under the name of Earth and Life Science. The table below show the distribution of periods per week, knowing that the duration of each periods is 50 minutes .

Noting that the number of periods in grade nine was two per week, until recently in 2015 it was increased to three periods per week following the request of teachers to meet the requirement of the official exam at the end of grade nine. As for the secondary level, Biology is delivered differently between grades, under the names of "life Science" and "scientific literacy". The table below show the distribution of periods at the secondary level.

Table 1: General Objectives of the Lebanese “Science Curriculum” (CRPD, 1997)

	General Objectives of the Lebanese “Science Curriculum” (CRPD, 1997)
Introduction	“Science plays an important role in our everyday life. It manifests itself in all aspects of human activity. Consequently, it is important that students become lifelong learners of science, starting with science at school, but extending science learning beyond the school years”
General objectives	• Develop the learners' intellectual and practical scientific skills • Deepen the learner’s awareness in the ability of humans to understand, invent, and create • Understand the nature of science and technology, their development across history, and their impact on human thought • Ensure that learners have acquired the facts, concepts, and principles necessary to understand natural phenomena • Motivate students to apply basic scientific principles in all sciences. • Explain the scientific concepts and principles behind commonly used machines and devices • Acquire knowledge about health, environment, and safety practices and behave accordingly • Realize that some natural resources can be depleted, and make the learner aware of the role of science in sustaining these resources • Encourage learners to use scientific knowledge and skills in novel situations, especially in everyday life • Emphasize the role of scientists in the advancement of human kind. • Encourage learners to be open to the ideas of scientists from different cultures, and to their contributions in the advancement of science • Encourage learners to abide by such scientific values as honesty and objectivity • Develop the learners’ scientific curiosity and orientation toward scientific research • Encourage learners to work independently and cooperatively in solving scientific problems • Make the learners aware of career possibilities in different science related areas

Table 2: Distribution of Biology Periods in the Lebanese Schooling System.

Grade	Number of periods/week
Seven	3
Eight	2
Nine	3

Table 3: Distribution of Biology Periods at the secondary level.

Grade	Number of periods/week/sections		
Secondary first	2		
Secondary second	Sciences Sections 2	Humanities Sections 1	
Secondary third	Life Science 6	sociology and Economics 2	Humanities 1

In the Lebanese system, there are two official exams, the first one at the end of the intermediate level (grade nine) and the second one at the end of the secondary level (grade twelve). Biology teachers (Earth and Life Science, and Life Science teachers) in most schools are concerned in delivering the content required by the curriculum. This was supported by research studies, so teachers focus mainly on the content, regardless students’ needs, interest and motivation, and attributing their decision to do so to time limitation and the need to prepare students for official exams (Boghtchalian Karadaghlian, 2014). Therefore, it is well noticeable that lecturing is the main mode of instruction in Lebanese schools, and thus deeper learning is neglected (Al

Husseiny, 2014; Zeidan 2014). Biology concepts are explained by text readings using textbook documents, with minor usage of animations and video presentations, which results in passive learning.

Research Problem

Being Science /Biology Educators at the Lebanese University, faculty of education, the researchers of this study have noticed that both teachers and students are complaining from difficulties in Biology classes. We interviewed ten “Life and Earth” and “life Science” qualified secondary teachers about those difficulties faced by their students in all grades, from grade seven to grade twelve. All teachers agreed that the main issue in their teaching is the teaching and learning of scientific skills. Students have little problem with the content, and the department of education has removed many topics from the curriculum in order to make it lighter and to allow extra time for the abstract content to be grasped. However, according to “Life and Earth” and “Life Science” teachers, students are always struggling with the learning and application of “scientific skills”. Researchers have different views about the categorization of scientific skills (Phang and Tahir, 2012).

Table 4: The Lebanese framework for competences and skills in Biology (CERD, 1998).

Domain	Skills
A- Acquiring Knowledge	A1- Recall Knowledge - Recall the acquired knowledge related to specific facts, terminology, law, theories, model… A2- Apply knowledge - Select the knowledge and use it in a new situation - Apply knowledge in a new context
B- Practicing Scientific Process	B1- Collect Information - Select information related to a real situation or to its representation in a table, text, graph, media… B2- Interrelate information to define a problem and/or formulate a hypothesis - Organize data in order to prove a relation - Compare new data to previous data - Identify a cause and effect relation - Define a problem - Formulate a hypothesis B3- Test a hypothesis - Identify the consequences implied by a hypothesis that could be verified - Design an experiment - Use data to test a hypothesis B4- Synthesize B5- Demonstrate critical thinking - Criticize experimental results, an argument, design an experiment
C- Mastering of Techniques	C1- Use laboratory or field materials and apply laboratory techniques C2- Perform an experiment following a given design C3- Carry out measurements, construct a model or make drawing based on observation…
D- Communicating	D1- Utilize proper scientific terminology - Use appropriate specific terminology to express information, observation, tabulated data, drawing, graph, or flow chart, in verbal or written form D2- Use various modes of scientific representation - Represent data by a table, a graph, a drawing, a chart, a symbol, or a formula.

The classification includes two levels: basic skills and integrated skills. Chiappetta and Koballa (2010) classified basic scientific skills as follows: observing, classifying, space/time relation, using numbers, measuring, inferring and predicting. As for integrated skills, they include: defining operationally, formulating models, controlling variables, interpreting data, hypothesizing, and Experimenting. In the Lebanese system, the

Centre of Educational Research and Development has set a framework for teachers to follow in the educational guide (1998). The framework includes four domains of competences: Domain A “Acquiring Knowledge”, Domain B “Practicing Scientific Process”, Domain C “Mastering of Techniques”, and Domain D “Communicating”. The table below represents the Lebanese framework for competences and skills in Biology. The current study investigates the effect of cooperative learning in Biology classroom, on students’ learning and achievement of scientific skills, namely measured by the achievement of “Practicing Scientific Process” Domain B of the Lebanese framework. Three skills were the focus of the study: “analyze a document (text, graph or table)”, deduce”, and “draw out a conclusion”. The Centre of Educational Research and Development (2012) has defined the requirement of these action verbs used in biology classes, as follows:

“Analyze: Decompose a whole into its constituent elements to make evident to variation.
Deduce: Draw using logical reasoning new information from given or existing information.
Draw out: Draw from a set of given and without reasoning a relation, a role, a law…” (CRDP, 2012. P 1)
The main research question addressed in the study: Does cooperative learning enhance students’ acquisition and practice of scientific skills in Biology classroom?

Consequently, the following sub-questions are investigated:
Does cooperative learning strategy enhance:

- the practice of scientific skills, namely for grade ten students?
- the acquisition of new scientific skills, namely for grade seven students?

It is assumed that cooperative learning enhances the performance of students, regardless of their class grade level, and improves their achievement in domain B representing the acquisition of “Practicing Scientific Process” skills.

Method

This study employed a quasi- experimental design in which two intact sections of each grade were assigned to control and experimental conditions, and an independent variable, the teaching method, was manipulated.
The researcher used the pre-test/post-test control group design. This design greatly minimizes threats to the internal validity of the experiment. In addition, self-assessment of the student’s and teamwork in the experimental groups were measured using two self- assessment grids with a five-point scale. The teachers observed the teams in the experimental groups, and filled out an observation grid to ensure that cooperative teamwork was conducted correctly.

Participants

School

A private high school in Beirut was selected based on its convenience in terms of location and time, and the willingness of the biology teachers to participate in the study. In addition, the school is a high school, thus contains all the grades which are of the sample of this study. Also, the school is well equipped with tools that facilitate cooperative learning in terms of wide classes.

Students

A total of 120 students (N=120) enrolled in two different grades (grade seven and grade ten participated in this study.) Grade seven represents the first class in the third cycle of the Lebanese educational system, while grade ten is the first class in fourth cycle. The table below represents the distribution of students among grades.

Table 5 shows the distribution of students among the groups.

Grade	Control Group	Experimental Group	Total Number of Students
7	30	30	60
10	30	30	60
Total	60	60	120

All participants were native speakers of Arabic and learning English as a first foreign language. English is the language of instruction in Biology.

Teachers

Two qualified biology teachers participated in the study; an Earth and life Science grade seven teacher, who taught both experimental and control groups, and a Life Science grade ten teacher who also taught both experimental and control groups.

Procedure

The duration of the study was a total of nine weeks. Students in both grades had two fifty minutes Biology sessions per week, which made a total of 18 sessions throughout the study. The study included two parts:

The first part was a training period for both the teachers and the students. It extended for two weeks. During this period, students practiced how to work cooperatively and distribute the roles among the team members. Moreover, the researchers followed up on the work of the teachers by making observations in both sections of each grade, to check if the teachers are planning their teaching periods according to the assigned and if the students are mastering how to work cooperatively. This part of the study aimed to help both students and teachers to master the learning strategy used in this study so that the results of the second part can be reliable.

The second part was the implementation period of the study, which was seven weeks long. This phase used the same process as described in the first phase, except that data was collected during this part. Both groups, experimental and control were given an equal amount of time and worked on the same hands-on activities but in different strategies (individually or cooperatively). They had the same assignments and were given equal opportunities to practice their learning objectives. For the control groups, the lessons were explained using the traditional individual learning (teacher demonstration approach). This method includes asking open-ended questions, oral reading of textbook, classroom discussion, and oral reviews. The teachers used the textbook and other materials including worksheets to help students construct their conceptual knowledge. Those sheets included hands-on activities that students must solve. Students in both experimental groups were teamed up in six groups of fives by the stratified random method in order to form heterogeneous groups. It is believed that when the size of team members increases, the range of abilities, expertise, skills, and number of minds available for acquiring and processing information increase (Johnson and Johnson, 1989). The experimental groups were taught using Johnson and Johnson Model of cooperative learning, where the same hands-on activities were used, but students worked on them cooperatively rather than individually. Students in those sections discussed open-ended questions in groups, read the content knowledge in cooperative groups, in addition to doing the hands-on activities cooperatively. When the groups completed their work and reached a consensus, the teachers asked the readers of the groups to explain their answers and discuss them with other members of the class. Assessment grids were used in the study, they were prepared, piloted and validated by the researcher based on group interaction in the cooperative learning teams (appendix):

A student self-assessment grid was given to each student to measure self-assessment of their learning; it included nine items (e.g. I accomplished my task; I organized my thought before and while speaking) in which each member of the team was to respond to a five-point Likert scale. A mark of 1-5 was applied on a scale that goes from a very positive assessment of their ability to a very negative assessment. The grid was administered twice during the study: once in the first session and once in the last session. The sum of the marks on each item were calculated for each student and named: "student self-assessment 1" referring to the first administration and "student self-assessment 2" for the second one. Moreover, each team was required to fill a team self -assessment grid and submit it before the end of the cooperative learning session. This included nine items (e.g. All of the team members contributed ideas; everyone in the team responded kindly to disagreements) assessing the teamwork and was filled according to a five-point scale. The aim of this grid was to ensure that students were practicing the cooperative skills to help teachers in team processing and providing feedback about the work. The sum of the marks on each item were calculated for each team and named: "team's self-assessment". In parallel to team self-assessment grids, an observation grid was designed to be used by the teacher and named "Teacher's team-assessment grid"; it was filled by the teachers during the cooperative work sessions for the experimental groups, where they chose randomly two teams each session for observation. The grid included nine five-point Likert scale items. Throughout the whole study, teachers had assessed all teams in each class, and then the marks of each item were calculated for each team and compared with the team's self-assessment to ensure the reliability of the teamwork. Means of the teams' and teacher's assessment scores were calculated for each experimental group in each grade, and were then compared to ensure the reliability of the teams' work. The results of teams' self-assessment and teachers' assessment of all the teams of both grades seven and ten show high values of self-assessment among the teams, an average of 18.2 and 18.8 respcetively. While teachers' assessment mean scores are 17 and 18.2 respectively. Those results show close scores between the teams' self-assessment and teachers' assessment of the teams, which indicates that cooperative learning was effectively implemented and reflected the honesty of students in assessing their work. In addition, students were comfortable in assessing themselves especially since these scores don't count in their task assessment. Thus, the results assure the reliability of the cooperative learning strategy used in the study.

Data Collection and Analysis

Students' achievement was measured by tests prepared by the researchers with the collaboration of other science educators. The tests included a variety of questions covering the three domains of evaluation in the Lebanese curriculum: Domain A (knowledge), B (cognitive), D (communication). For both grade seven and ten, three main "action verbs" were presented in the pre and post-tests and repeated twice, representing the higher order skills: "analyze", "deduce" and "draw out". The tests were piloted and based on the piloting results; minor modifications were made. Descriptive statistics for the domain B questions items of pre and post-tests were computed for each control and experimental group in each grade (seven and ten). To check whether the dependent variable (teaching method used) had an effect or not on student achievement in domain B questions, t-tests were conducted to determine whether there were significant differences between means on domain B selected items on pre and post-tests of the control and experimental groups in each grade.

Results and Discussion

The results show that grade seven students in the experimental and control groups have very close means in the pre-test domain B items (10.37/20 and 10.5/20 respectively). In the post-test, both groups show an increase in the scores (15.5/20 and 14.12/20 respectively). However, the increase in the means of experimental group is greater than that of the control group. Table 6 shows the mean scores and standard deviations for the domain B scores in pre- and post-tests of both experimental and control groups in grade seven.

Table 6: Mean and Standard Deviation of Domain B scores of pre-and post-test for grade seven groups

	Group	N	Mean	SD
Pretest	G7 E.	30	10.375	3.2250
	G7 C	30	10.500	3.2088
Posttest	G7 E	30	15.504	2.8728
	G7 C	30	14.118	3.4436

Similarly, concerning grade ten, table 7 shows the mean scores and standard deviations for the domain B scores in pre- and post-tests of both experimental and control groups. Students in both groups have very close means in the pre-test (9.67/20 and 8.69/20 respectively). In the post-test, both groups show a significant increase in the domain B scores (13.93/20 and 11.15/20 respectively). However, the increase in the means of experimental group is greater than that of the control group. Table 7 shows the mean scores and standard deviations for the domain B scores in pre- and post-tests of both experimental and control groups in grade ten.

Table 7: Mean and Standard Deviation of Domain B scores of pre-and post-test for grade ten experimental and control groups.

	Group	N	Mean	SD
Pretest	G10 E.	30	9.671	4.4861
	G10 C.	30	8.686	4.1851
Posttest	G10 E.	30	13.929	3.8909
	G10 C.	30	11.153	5.8610

A two-tailed t-test at the level 0.05 of significance shows no significant difference between cooperative learning and individualistic learning in grade seven on domain B of pre-or post-tests (t-value of 0.89 and 0.1 respectively). Table 8 shows the the results of T-test for pre-test and post-test domain B items for grade seven experimental and control groups.

Table 8: T-test for pre-test and post-test domain B items for grade seven experimental and control groups.

	Group	N	Mean	SD	Sig (2-tailed)	Mean difference
Pretest	G7 E.	30	10.375	3.2250	0.885	- 0.1250
	G7 C.	30	10.500	3.2088		
Posttest	G7 E.	30	15.504	2.8728	0.108	1.3866
	G7 C.	30	14.118	3.4436		

On the other hand, a two-tailed t-test at the level 0.05 of significance shows no significant difference between cooperative learning and individualistic learning in grade ten on domain B of pretest (t-value of 0.399 > 0.05). However, there is a significant difference on domain B of post-test (t-value of 0.042 < 0.05). Table 9 shows the results of T-test for pre-test and post-test domain B items for grade ten experimental and control groups.

Table 9: T-test for pre-test and post-test domain B items of grade ten experimental and control groups.

	Group	N	Mean	SD	Sig (2-tailed)	Mean difference
Pretest	G10 E.	30	9.671	4.4861	0.399	0.9857
	G10 C.	30	8.686	4.1851		
Posttest	G10 E.	30	13.929	3.8909	0.042	2.7760
	G10 C.	30	11.153	5.8610		

As mentioned earlier, cooperative learning has no effect on students' domain B scores in grade seven experimental group. This result shows that cooperative learning has neither a positive nor a negative effect on students' achievement. The findings of this study are in line with the results of other studies carried out about the effect of cooperative learning on students' academic achievement in science. This result is not consistent with the literature concerning the effect of cooperative learning on academic achievement for primary and secondary school students, with few exceptions, such as Sherman (1989) and Chang and Lederman (1994) who applied different types of cooperative learning models on middle school students and came up with the same results: cooperative learning had no effect on students' academic achievement in science, but had an effect on other aspects like confidence and attitudes.This lack of significant differences between the control and the experimental groups in grade seven in this study may be due to the fact that Domain B skills are newly introduced to grade seven students, since it is their first year of learning biology as a separate discipline, and in this grade they are introduced to skills such as "analyze", "deduce" and "draw out". Students might need more time to show a significant acquisition of the concept. Therefore, it is recommended that more research is needed to investigate this finding. However, results of grade ten students are different. They show significant difference in post-test for domain B items. The findings are in parallel with the literature review, namely the reported studies that confirm a significant correlation between cooperative learning and achievement and that cooperative learning engages students in the learning process and improves critical thinking, reasoning, and problem-solving skills of the learner (Chang, & Mao, 1999; Bilgin and Geban ,2006; Nezami , Asgari and Dinarvand, 2013).

Conclusion

The main purpose of the study is to investigate the effect of cooperative learning strategy on teaching and learning scientific skills in grades seven and ten. The results show a significant improvement in students' achievement of scientific skills in grade ten; however, grade seven students show improvement but not significant. Therefore, we may conclude that students' grade level and the complexity of concept introduced have impacts on students' outcomes. It is well clear that cooperative learning has a positive effect on teaching and learning scientific process skills, even though it is not always significant, but it does improve the learning and practice process of the acquired skills and help the learning of new skills. The results of this study are in line with findings reported in the literature.

Recommendations

Changes are needed in biology teaching. Teachers should pay more attention to teaching and learning scientific skills, and incorporate a variety of teaching strategies to overcome the major problems faced by students in the acquisition and practice of scientific process skills. Cooperative learning is proved by research to be an effective strategy to enhance the learning and practice of those skills. It is recommended to investigate in a larger scale the effectiveness of cooperative learning on teaching and learning skills, by including a larger sample for both grades for a longer period of time.

Notes

Data used in this study were taken from a Master Research study conducted by Fatima Al Husseiny under the supervision of Dr. Hanadi Chatila at the Lebanese University, entitled *Enhancing Critical Thinking through Cooperative Learning in Biology.*

References

Agboghoroma, T E., & Oyovwi, E.O (2015). Evaluating Effect of Students’ Academic Achievement on Identified Difficult Concepts in Senior Secondary School Biology in Delta State. *Journal of Education and Practice, 6 (30), 117-125.*

Agorram, B; Clement, P; Selmaoui, S; Khzami,S.E., & Chafik, J (2010). University students’ conceptions about the concept of gene: Interest of historical approach. *US-China Education Review. 7 (2), 9 - 15.*

Al Husseiny, F. I. (2014). *Enhancing Critical Thinking through Cooperative Learning in Biology. Research* Master Dissertation, Lebanese University.

Aydin, S. (2011). Effect of cooperative learning and traditional methods on students' achievements and identifications of laboratory equipments in science-technology laboratory course. *Educational Research and Reviews*,*6*(9), 636-644

Bahar M, Johnstone A.H., & Hansell M.H. (1999). *Revisiting learning difficulties in biology. Journal of Biological Education., 33(2), 84-86.*

Bilgin, I., & Geban, O. (2006). The effect of cooperative learning approach based on conceptual change condition on students’ understanding of chemical equilibrium concepts. *Journal of Science Education and Technology, 15(1),* 31-46.

Boghtchalian Karadaghlian, S (2014). The Effect of Computer Based Games on the Academic Achievement and Motivation towards Biology of Grade 8 Students. Unpublished thesis. Lebanese University, Beirut, Lebanon.

Center for Educational Research and Development CERD (1997). Lebanese national curriculum. Beirut, Lebanon.

Center for Educational Research and Development CERD (1998). Life Science Pedagogical Guide, Teacher’s guide, Secondary Education, First Year.

Center for Educational Research and Development CERD (2012). Verbes d’action : Physique, chimie et Sciences de la Vie. Définitions et exemples.

Chang, C. Y., & Mao, S. L. (1999). The effects on students' cognitive achievement when using the cooperative learning method in earth science classrooms. *School Science and Mathematics, 99*(7), 374-379.

Chang, H. P., & Lederman, N. G. (1994). The effect of levels of cooperation within physical science laboratory groups on physical science achievement. *Journal of Research in Science Teaching, 31*(2), 167-181.

Chattopadhyay, A. (2005). Understanding of genetic information in higher secondary students in northeast India and the implications for genetics education. *Cell Biology Education*, *4*(1), 97-104.

Chiappetta, E. L., & Koballa, T. R. (2010). Science instruction in the middle and secondary schools: Developing fundamental knowledge and skills for teaching (6 th ed.). NJ: Pearson Prentice-Hall.

Çimer, A, (2012).What makes biology learning difficult and effective: Students’ views. *Educational Research and Review, 7(3), 61-71.*

Iyer, R. B. (2013). Relation between cooperative learning and student achievement. *International Journal of Education and Information Studies*,*3*(1), 21-25.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.

Johnson, D., W.; Johnson, R., T., & Smith, K., A. (2014). Cooperative Learning: Improving University Instruction By Basing Practice On Validated Theory. *Journal on Excellence in College Teaching, 25 (4),*1-26.

Lin E (2006). Learning in the science classroom. Sci. Teach., 73(5): 35- 39.

Lord, T. R. (2001). 101 reasons for using cooperative learning in biology teaching. *The American Biology Teacher*, *63*(1), 30-38.

Lunenburg, F. C. (2011). Critical thinking and constructivism techniques for improving student achievement. *National Forum of Teacher Education Journal, 21 (3),* 1-9.

Nezami, N. R., Asgari, M., & Dinarvand, H. (2013). The effect of cooperative learning on the critical thinking of High School students. *Technical Journal of Engineering and Applied Sciences*, *3(19)*, 2508-2514.

Phang, F. A., & Tahir, N. A. (2012). Scientific skills among pre-service science teachers at universiti teknologi malaysia. *Procedia-Social and Behavioral Sciences*, *56*, 307-313.

Richards, J., & Rodgers, T. (2001). Approaches and methods in language teaching. Cambridge: Cambridge University Press.

Sadler, K. C. (2002). The Effectiveness of Cooperative Learning as an Instructional Strategy To Increase Biological Literacy and Academic Achievement in a Large, Non-Majors College Biology Class. Retrieved from the ERIC database. (ED 471 548).

Sherman, L. W. (1989). A comparative study of cooperation and competitive achievement in two secondary biology classrooms: The group investigation model versus an individually competitive goal structure. *Journal of Research in Science Teaching*, *26* (1), 55-64.

Tekkaya C, Özkan Ö., & Sungur S (2001). Biology concepts perceived as difficult by Turkish high school students. Hacettepe Univ. J. Educ., 21, 145-150.

Tran, V. D. (2014). The effects of cooperative learning on the academic achievement and knowledge retention. *International Journal of Higher Education*, *3*(2), 131-140.

Vijayaratnam, P. (2009, November). Cooperative learning as a means to developing students' critical and creative thinking skills. *Proceedings of the 2nd International Conference of Teaching and Learning, Nilai, Negri Sembilan,* 16-18.

Wolfensberger, B., & Canella, C. (2015). Cooperative Learning about Nature of Science with a Case from the History of Science. *International Journal of Environmental and Science Education*, *10*(6), 865-889.

Zeidan, A. (2010). The relationship between grade 11 Palestinian attitudes toward biology and their perceptions of the biology learning environment. *International Journal of Science and Mathematics Education, 8(5),* 783 – 800.

Zeidan, A., & Jayosi,M., R. (2015). Science Process Skills and Attitudes toward Science among Palestinian Secondary School Students *World Journal of Education. 5 (1),* 13-24.

Zeidan, M (2014). *The Effect of Self-Regulation (Checklist as a Tool of Learning) on Students' Motivation and Achievement in Biology.* Unpublished thesis. Lebanese University, Beirut, Lebanon.

Author Information

Hanadi Chatila
Lebanese University, Faculty of Education.
Beirut, Lebanon.

Fatima Al Husseiny
Lebanese University, Faculty of Education.
Beirut, Lebanon.

Appendix

STUDENT SELF-ASSESSMENT GRID

Name:………………………………………………
Date……………………………………………
Grade……… / Section: ……
The Title of the activity: ……………………………………..……………………………………………

- *The objective of this grid is to self -assess the student's learning skills during team work.*

Using "5" as the highest point and "1" as the lowest, decide to what degree you were successful in each of the following areas. Circle one number.

Student's Self-assessment of Learning Skills					
1-I accomplished my tasks.	1	2	3	4	5
2-I shared ideas and opinions.	1	2	3	4	5
3-I organized my thoughts before and while speaking.	1	2	3	4	5
4-I used appropriate terms when stating ideas.	1	2	3	4	5
5-I asked for facts and reasoning.	1	2	3	4	5
6-I offered to explain and clarify statements.	1	2	3	4	5
7-I clarified statements using examples.	1	2	3	4	5
8-I can summarize what have been said without referring to notes.	1	2	3	4	5
9-I can relate the material to previous information or experience.	1	2	3	4	5

TEAM'S SELF-ASSESSMENT GRID

Date……………………………………………
Grade……… / Section: ……
The Title of the activity: ……………………………………..……………………………………………

- *The objective of this grid is to self-assess our team's work.*

A. Using "5" as the highest point and "1" as the lowest, decide to what degree your team was successful in each of the following areas. Circle one number.

Team Self- Assessment					
1- All of the team's members contributed ideas.	1	2	3	4	5
2- All of the team's members listened carefully to the ideas of other team members.	1	2	3	4	5
3- All of the team's members encouraged other members to contribute their thoughts and opinions.	1	2	3	4	5
4- Everyone in the team shared ideas/information.	1	2	3	4	5
5- Everyone in the team helped others.	1	2	3	4	5
6-Everyone in the team accepted help.	1	2	3	4	5
7-Everyone in the team responded kindly to disagreements.	1	2	3	4	5
8- Everyone in the team understood the activity.	1	2	3	4	5
9- We finished the task on time.	1	2	3	4	5

Team Members' Names

……………………………………………………………………………..………
……………………………………………………………………………..………
…………………………..……………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………

TEACHER'S TEAM-ASSESSMENT GRID

Date……………………………………………
Grade……… / Section: ……
The Title of the activity: ………………………………………..……………………………………………

- *The objective of this grid is to assess the team's work by the teacher during her observation.*

Team Members' Names

……………………………………………………………………..…………
……………………………………………………………………..…………
……………………………..…………………………………………………
………………………………………………………………………………
……………………………………………………………………..…………

Team Work Assessment					
1- All of the team's members contributed ideas.	1	2	3	4	5
2- All of the team's members listened carefully to the ideas of other team members.	1	2	3	4	5
3- All of the team's members encouraged other members to contribute their thoughts and opinions.	1	2	3	4	5
4- Everyone in the team shared ideas/information.	1	2	3	4	5
5- Everyone in the team helped others.	1	2	3	4	5
6-Everyone in the team accepted help.	1	2	3	4	5
7-Everyone in the team responded kindly to disagreements.	1	2	3	4	5
8- Everyone in the team understood the activity.	1	2	3	4	5
9- The team finished the task on time.	1	2	3	4	5

JESEH
Journal of Education in Science, Environment and Health
Volume 3, Issue 1, 2017 ISSN: 2149-214X

The Effect of Active-Participant Experiments upon the Skills of Nursery Class Students to Recognize Measuring Instruments

Ozkan Sapsaglam, Aykut Emre Bozdogan

Article Info

Article History

Received:
01 September 2016

Accepted:
30 December 2016

Keywords

Preschool period
Science activities
Experiments
Measuring instruments

Abstract

Preschool children learn through their senses. Children learn language, daily life skills, concepts and many other things through their senses. Thus, preschool educational environments and preschool educational activities should stimulate children’s senses. In this context, preschool science activities and experiments have positive effects upon children’s development and learning as they contain skills aimed at various senses like observation, relationship establishment, interpretation, inference and discussion. The objective of this study is to determine the effect of active-participant experiments upon the skills of preschool children to recognize and accurately select measuring instruments that are encountered in daily life and used in measuring various magnitudes. In the study, action research method was used. The study was conducted in the school year of 2015-2016 with totally 19 children (8 male and 11 female) aged 48-60 months. A total of 6 experiments (mass, weight, volume, length, temperature and time) were performed in 3 weeks to measure two magnitudes each week. In order to collect the study data, the researchers developed a measuring instrument of totally 6 questions, 3 of which were distracting. As a result of the study, it was determined that the experiments were effective upon the skills of children to recognize and accurately select measuring instruments that were used in measuring daily life magnitudes like mass, weight, volume, length, temperature and time.

Introduction

Preschool period is the second period where development is the fastest after prenatal period in human life. In periods where development is fast, learning is fast as well. In the preschool period, children mainly learn through their senses and acquire abundant knowledge and skills through their senses. In the first years of life, children seek and try to learn knowledge and skills to use in their future life. Jean Piaget defines children as little scientists that try to explore the world. The quest of children in the learning process can only be responded by an environment that is rich in stimulus. A quality preschool educational environment and quality preschool educational activities will provide the stimulus needed by children in the learning process. The development of young children is realized through their experiences in centers like school and home within micro systems (Bronfenbrenner and Morris, 2006). Children begin to investigate and explore their environment as from the first years of life (California Department of Education, 2013) and they have the skills of comprehending and learning knowledge in scientific fields like life science, location science and physical science at an early age (Erden and Sönmez, 2011; National Association for the Education of Young Children, 2014). However, the preschoolers sometimes overgeneralize (Kabadayı, 2006) the objects and sometimes over regularize (Kabadayı, 2012) them in their responses in different categories, while object labeling process in their environment. The study results show that preschool children can think about science and use scientific knowledge by making comparisons and estimations (Carey 2009; National Research Council, 2007; Schulz and Bonawitz, 2007).

Free researches and experiments conducted in the first two years of life will enable the development of children’s senses (Charlesworth and Lind, 2003). Children’s natural behaviors in these years will increase their interest in science in the future. Rich stimulus experiences (seeing, hearing, tasting, touching, and smelling) allow children to become better observers and more curious individuals (Akman, Üstün and Güler, 2003). It is believed that children can explore scientific knowledge and skills only in a free environment (Sundberg and Ottander, 2014).

Gelman, Brenneman, Macdonald, and Roman (2009) collect scientific applications aimed at preschool children under five titles as; 1) observation, prediction, control, 2) comparison, opposition, experiment, 3) word, discussion and language, 4) accounting, measuring and math's and 5) record and documentation.

Rather than directly conveying scientific knowledge to children; preschool science education aims to allow them to learn these knowledge by experiencing and practicing (Aktaş-Arnas, 2002). Science education allows children to understand and learn their environment. With the help of science activities, children develop their learning skills, explore, obtain information about scientific methods like measuring, learn scientific language and special signs, conduct and explain experiments and participate in discussions (Curriculum For Excellence, 2004: 253; Dere and Ömeroğlu, 2001). Experiments play an important role in preschool science education activities. Experiments that are conducted in preschool education classes allow children to materialize abstract concepts and develop positive attitudes toward science (Alisinanoğlu, Özbey and Kahveci, 2011). Preschool children can learn more easily through concrete ways. Experiments are necessary for materializing abstract scientific concepts that children desire to investigate (Kandır et al., 2012). Children acquire not only scientific knowledge, but also scientific process skills with the help of experiments that are conducted in preschool science activities. From the middle of the 20th century, inclusion of all scientific process skills in science education programs has become important (Erten, Kiray and Sen-Gumus, 2013). Scientific process skills of preschool children include; observation, measuring, comparison, classification and communication (Alisinanoğlu, Özbey and Kahveci, 2007:14; Charlesworth and Lind, 2003; Kandır, Can Yaşar, İnal, et al., 2012:17; Keil, Haney and Zoffel, 2009). Choi (2016) explains the science learning of preschool children as in Figure 1.

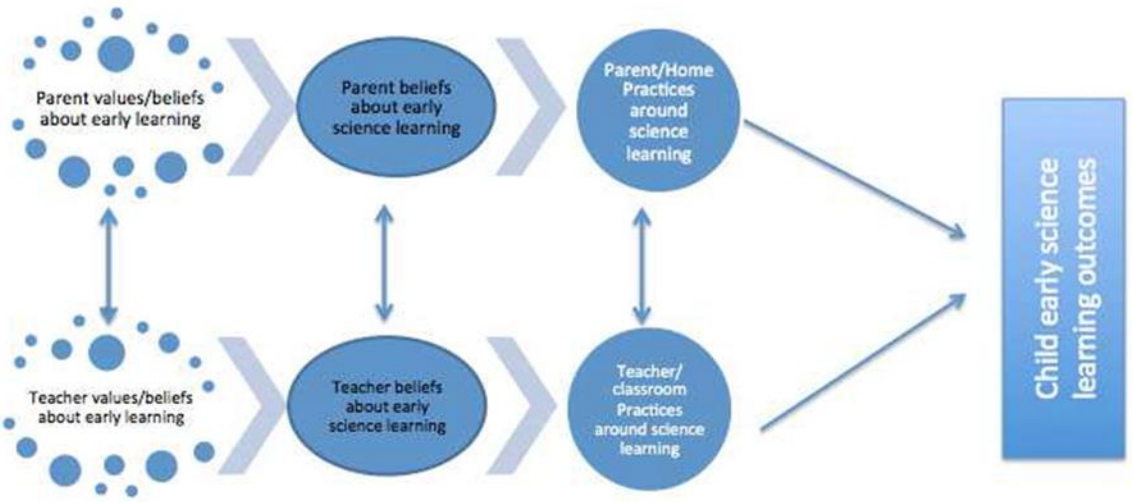


Figure 1. Children's development of early science learning (Source, Choi, 2016)

Being one of scientific process skills; measuring is defined as making qualitative and quantitative distinctions between objects and features via an measuring instrument (TDK, 2016). Measuring, in the most general sense, is a comparison and accounting and it contains the use of standard and non-standard units for defining measurable qualities like linear dimensions, area, volume, time, temperature and mass. Children can measure the length of a table, extent or width of a room, height of a door, their own length and a lot of things in their physical environment. (Martin, 2011: 87). Measuring skill allows preschool children to solve many daily life problems that require this skill. Thus, it is important for children to acquire this skill at a early age. Examining the studies on preschool science education in literature; it is seen that the studies focus on methods and techniques being used by preschool teachers in science education (Ayvacı, Devecioğlu and Yiğit, 2002; Güler and Bıkmaz, 2002; Karaer and Kösterelioğlu, 2005; Taş, 2010), content of preschool science education (Brunton and Thornton, 2010), competence of preschool teachers in science education (Karamustafaoğlu and Kandaz, 2006; Özbey, 2009), attitudes of preschool teachers toward science education (Kallery, 2004; Ünal and Akman, 2006), role of teachers in science education (Martin et al., 2005), awareness of teachers (Chang, 2012), teachers' sense of science (Saçkes et al., 2009) and children's exploration of the properties of water (Siri, Ziegler and Max, 2012). The study results also show that preschool children and their teachers do not spare enough time for science activities (Early et al., 2010). Relevant literature does not include any studies aimed at the skills of preschool children to recognize and select measuring instruments that are used in measuring various magnitudes via active-participant experiments. Thus, both the study and its potential results are considered important.

Study Objectıve

The objective of this study is to determine the effect of active-participant experiments upon the skills of preschool children (aged 48-60 months) to recognize and accurately select measuring instruments (meter, equal arm balance, dynamometer, graduated cylinder, thermometer and chronometer) that are encountered in daily life and used in measuring various magnitudes (length, mass, weight, volume, temperature, time).

Method

Action research, one of the qualitative research methods, was used in the study. Action research is suitable for the individuals who are directly related to the process to be developed and it can also be used to solve the professional problems a person meet in his field of expertise or to increase the quality of the work he does (Büyüköztürk, et. al., 2010). It was thought in the study conducted that the use of this method was suitable to determine the effect of active-participant experiments upon the skills of preschool children (aged 48-60 months) to recognize and accurately select measuring instruments that are encountered in daily life and used in measuring various magnitudes. Table 1 shows the study pattern.

Table 1**.** Study Pattern

Groups	Pretest	Procedure	Posttest	Permanence test
Experiment	Informational questions	Experimental study (3 weeks)	Informational questions	Informational questions (5 weeks later)

Procedure

1. Week: Visual measuring instrument of totally 6 questions, 3 of which were distracting, aimed at determining whether or not children were able to recognize appropriate measuring instruments used in measuring various magnitudes like mass, weight, volume, length, temperature and time in daily life was displayed for children by using a projector and a projection curtain and children were individually asked to answer the questions. Answers given by children to the questions were recorded by the researcher in writing. The process lasted for approximately 8 minutes for each child.

2. Week: In the experimental procedure stage, 2 different experiments were prepared for introducing equal arm balance, which is used measuring the mass, and dynamometer, which is used in measuring weight, and the experiments were applied with active participation of children. Each experiment lasted for approximately 35 minutes.

3. Week: The experimental procedure was sustained and 2 different experiments were prepared for introducing meter, which is used in measuring length, and chronometer, which is used in measuring time, and the experiments were applied with active participation of children. Each experiment lasted for approximately 30 minutes.
4. Week: In this final week of the experimental procedure, 2 different experiments were prepared for introducing thermometer, which is used in measuring temperature, and graduated cylinder, which is used in measuring volume, and the experiments were applied with active participation of children. Each experiment lasted for approximately 35 minutes.

5. Week: The measuring instrument that was applied in the 1st week after completing the experimental procedure was reapplied to children as a posttest in the computer environment. In that process, children individually answered the questions that were mirrored by computer and projection. Their answers were recorded by the researcher. The process lasted for approximately 7 minutes for each child.

10. Week: 5 weeks after the application of the posttest, the same measuring instrument was reapplied to children as a permanence test in the computer environment. Children individually answered the questions that were mirrored by computer and projection. Their answers were recorded by the researcher. The process lasted for approximately 7 minutes for each child.

Working Group

Working group of the study was determined by using the technique of typical case sampling, which is among purposeful sampling methods. Typical case sampling is among the purposeful sampling methods used for determining the sample group in qualitative researches. In typical case sampling, it is aimed to have an opinion about a particular field by studying average conditions (Yıldırım and Şimşek, 2008). The study included totally 19 children (8 male and 11 female) aged 48-60 months, who were thought to show similar features to other children in this age group. All children in the working group showed a normal development.

Data Collection Tool

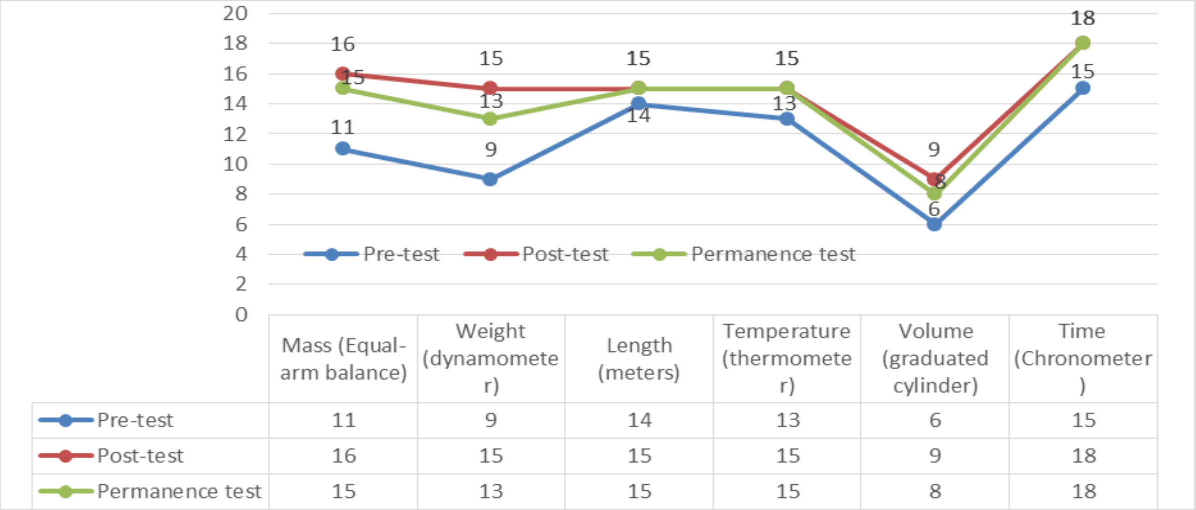
In order to collect the study data, the researchers developed a visual measuring instrument of totally 6 questions, 3 of which were distracting, aimed at determining whether or not children were able to recognize appropriate measuring instruments used in measuring various magnitudes like mass, weight, volume, length, temperature and time in daily life. The measuring instrument was developed based on the opinions of 2 preschool teachers and 3 academicians. In this context, the visual measuring instrument involved 6 questions aimed at using appropriate measuring instruments in measuring these concepts encountered in daily life and popular cartoon characters for children **(Appendix 1).** The visual measuring instrument was displayed for children using a projector and a projection curtain and children were individually asked to answer the questions. Answers given by children to the questions and their reasons were recorded by the researcher in writing.

Data Analysis

Statistical analyses of the data that were collected for sub-problems, which were tried to be answered within the frame of the general study objective, were conducted by using frequency. Numerical data were interpreted in graphics. Besides, opinions of children about measuring instruments were conveyed without making any changes.

Findings

Diagram 1 shows the data concerning the state of preschool children (aged 48-60 months) to recognize and select measuring instruments used in measuring various magnitudes that are encountered in daily life.



	Mass (Equal-arm balance)	Weight (dynamometer)	Length (meters)	Temperature (thermometer)	Volume (graduated cylinder)	Time (Chronometer)
Pre-test	11	9	14	13	6	15
Post-test	16	15	15	15	9	18
Permanence test	15	13	15	15	8	18

Diagram 1. Distribution of Preschool Children (Aged 48-60 Months) to Recognize and Select Measuring Instruments Used in Measuring the Magnitudes of Mass, Weight, Length, Temperature, Volume and Time (f)

Examining Diagram 1; 11 children responded “right” and 8 children responded “wrong” to the question, “mass should be measured with equal-arm balance” in the pretest. However, at the end of the experimental study, it was observed that 16 children responded “right” and 3 children responded “wrong” to this question. In the permanence test, 15 children responded “right” and 4 children responded “wrong”. Similarly, 9 children responded “right” and 10 children responded “wrong” to the question, “weight should be measured with dynamometer” in the pretest. However, at the end of the experimental study, it was observed that 15 children responded “right” and 4 children responded “wrong” to this question. In the permanence test, 13 children responded “right” and 6 children responded “wrong”. 14 children responded “right” and 5 children responded “wrong” to the question, “length should be measured with meter” in the pretest. However, at the end of the experimental study, it was observed that 15 responded “right” and 4 children responded “wrong” to this question. In the permanence test, the posttest scores remained the same. 13 children responded “right” and 6 children responded “wrong” to the question, “temperature should be measured with thermometer” in the pretest. However, at the end of the experimental study, it was observed that 15 children responded “right” and 4 children responded “wrong” to this question. In the permanence test, the posttest scores remained the same. 6 children responded “right” and 13 children responded “wrong” to the question, “volume should be measured with graduated cylinder” in the pretest. However, at the end of the experimental study, it was observed that 9 children responded “right” and 10 children responded “wrong” to this question. In the permanence test, 8 children responded “right” and 11 children responded “wrong”. Finally, 15 children responded “right” and 4 children responded “wrong” to the question, “time should be measured with chronometer” in the pretest. However, at the end of the experimental study, it was observed that 18 children responded “right” and only 1 child responded “wrong” to this question. In the permanence test, the posttest scores remained the same.

Examining Diagram 1; it was observed that children had preliminary information mainly about the necessity for measuring time with chronometer (f=15), which was followed by the necessity for measuring length with meter (f=14), temperature with thermometer (f=13), mass with equal-arm balance (f=11) and weight with dynamometer (f=9). On the other hand, children had minimum preliminary information about the necessity for measuring volume with graduated cylinder (f=6). However, as a result of the experimental study, an increase was observed in the rate of children to recognize and accurately select appropriate measuring instruments in measuring these concepts. In this context, 18 children stated, “time should be measured with chronometer”, 16 children stated, “mass should be measured with equal-arm balance” and 15 children individually stated, “weight should be measured with dynamometer, temperature with thermometer and length with meter” in the posttest. 9 children also stated, “volume should be measured with graduated cylinder”. As a result of the permanence test, there was a decrease at the least and it was determined that children continued to recognize and accurately select appropriate measuring instruments in measuring these concepts. In this context, 18 children stated, “time should be measured with chronometer”, 15 children individually stated, “mass should be measured with equal-arm balance, temperature with thermometer and length with meter”, 13 children stated, “weight should be measured with dynamometer” and 8 children stated, “volume should be measured with graduated cylinder” in the permanence test. Table 2 shows the opinions of preschool children (aged 48-60 months) about their state of recognizing and selecting measuring instruments that are used in measuring various magnitudes encountered in daily life.

Conclusion

The study results reveal that experiments being conducted are effective upon preschool children (aged 48-60 months) to recognize and accurately select measuring instruments used in measuring various magnitudes like mass, weight, length, temperature, volume and time that are encountered in daily life. In this context, it is seen that children’s active participation in experiments under the guidance of teachers provides an opportunity for permanent learning. Experiments are among basic learning methods in science education (Yıldız, Aydoğdu, Akpınar and Ergin, 2006:72). Experiments can be performed on many different topics in preschool education classes and these experiments will allow children to materialize abstract concepts (Alisinanoğlu et al., 2011; Kandır et al., 2012; Uyanık Balat and Önkol, 2010). Teachers have a tendency to show children scientific facts via experiments and science activities and they frequently apply to experiments in science activities (Merino et al., 2014; Özbek, 2009). Fleer (2013) suggests that teachers realize science education via counseling, material procurement, discussion or experiments. Preschool teachers primarily use scientific experiments and activities for the purpose of showing children the world (Choi, 2016).

The study results reveal that children have preliminary information mainly about the necessity for measuring time with chronometer, length with meter and temperature with thermometer, which is associated with the fact that they encounter all three measuring instruments very frequently in daily life. On the other hand, children have minimum preliminary information about the necessity for measuring volume with graduated cylinder and weight with dynamometer, which is associated with their inability to recognize these measuring instruments or perceive what they measure.

Table 2. Opinions of Preschool Children (Aged 48-60 Months) about Their State of Recognizing and Selecting Measuring Instruments That are Used in Measuring Various Magnitudes Encountered in Daily Life

Concepts	**Code**	**Pre-test**	**Post-test**	**Permanence test**
Mass	K_4	**Answer:** Equal arm balance (✓) It will lift as it can scale anything.	**Answer:** Equal arm balance (✓) It is called balance. It has two plates; the light one rises whereas the heavy one descends just like a teeter totter.	**Answer:** Equal arm balance (✓) It measures with balance.
	E_1	**Answer:** Thermometer (X) Because it has numbers on it.	**Answer:** Equal arm balance (✓) Because it contains two sacs.	**Answer:** Equal arm balance (✓) Because it is balance.
	E_8	**Answer:** N/A (X) I do not know.	**Answer:** Equal arm balance (✓) Because it is balance.	**Answer:** Equal arm balance (✓) It puts them in sacs.
Weight	E_2	**Answer:** Meter (X) To measure the length of the basket.	**Answer:** Dynamometer (✓) The weighbridge rises when a weight is put on the hook. The hook perceives the heavy one.	**Answer:** Dynamometer (✓) Because it measures the weight of anything.
	K6	**Answer:** N/A (X) I do not know.	**Answer:** Dynamometer (✓) It measures the weights with a hook.	**Answer:** Dynamometer (✓) I forgot what it is called, but it is measured with it. Because others can not measure it.
	E5	**Answer:** Equal arm balance (X) It puts the apples in the basket and measures them.	**Answer:** Dynamometer (✓) It squeezes the apples in its hook. The basket rises when it is hung.	**Answer:** Dynamometer (✓) Because weight is measured with it. The hook is put on the basket.
	E8	**Answer:** N/A (X) I do not know	**Answer:** Dynamometer (✓)	**Answer:** Dynamometer (✓)
Length	K11	**Answer:** Meter (✓) Because that's what measures the length.	**Answer:** Meter (✓) Because it is meter. It measures lengths.	**Answer:** Meter (✓) My mother always measures my length with it.
	E4	**Answer:** Meter (✓) Because it's a length scale.	**Answer:** Meter (✓) We can measure it with meter. We used to measure length with it.	**Answer:** Meter (✓) Meter needs to be used. Because we measured with it.
	E6	**Answer:** Meter (✓) It puts them on and measures them.	**Answer:** Meter (✓) I forgot what it is called. It steps on your foot and remains like that.	**Answer:** Meter (✓) Because it measures according to your length.
	K6	**Answer:** N/A (X) I do not know.	**Answer:** Meter (✓) We used it for measuring length.	**Answer:** Meter (✓) It measures with meter.
Temperature	E4	**Answer:** Thermometer (✓) Because it can measure the temperature of water.	**Answer:** Thermometer (✓) Thermometer measures. Because the red sign moves in the heat.	**Answer:** Thermometer (✓) Because temperature of water can be measured with thermometer. The red sign either rises or descends.
	K11	**Answer:** Thermometer (✓) Because it feels the temperature.	**Answer:** Thermometer (✓) It is called thermometer. We put it in water.	**Answer:** Thermometer (✓) I forgot what it is called, but it puts things in water and measures..
Volume	K9	**Answer:** N/A (X) I do not know.	**Answer:** Thermometer (✓) It had risen when we put it in hot water and descended when we put it in cold water.	**Answer:** Thermometer (✓) Because hot water rises the red sign.
	K6	**Answer:** N/A (X) I do not know.	**Answer:** Graduated cylinder (✓) It throws the rocks inside and measures them.	**Answer:** Graduated cylinder (✓) I forgot what it is called, but that's it.
	E4	**Answer:** Equal arm balance (X) It can be measured with a scale	**Answer:** Graduated cylinder (✓) Because it is called scale.	**Answer:** Graduated cylinder (✓) We have one in our kitchen at home and we measure with it.
	E6	**Answer:** Equal arm balance (X) It puts one marble in one sac and another marble in another sac and measures them.	**Answer:** N/A (X) I do not know.	**Answer:** Microscope (X) It throws the marble in it and observes via its glass.
Time	E1	**Answer:** Chronometer (✓) Because there are inscriptions.	**Answer:** Chronometer (✓) Because there are inscriptions.	**Answer:** Chronometer (✓) Because there are numbers in it.
	K6	**Answer:** N/A (X) I do not know.	**Answer:** Chronometer (✓) It could be the other one as the two are not the answer.	**Answer:** Chronometer (✓) It is measured with it as the others can not.
	E4	**Answer:** Chronometer (✓) Only the one that resembles a watch can measure.	**Answer:** Chronometer (✓) Others measure temperature and another one is used in repairing.	**Answer:** Chronometer (✓) It is held in one hand to count the time.

The study results show that children make information obtained from active-participant experiments meaningful by associating them with daily life. For instance, the child coded E4 responded “equal arm balance” in the pretest and “graduated cylinder” in the posttest and the permanence test to the question, “What measuring instrument should measure volume”. The child explained his reason for selecting the option in the permanence test as, “We have one in our kitchen at home and we measure with it.” This answer proves that children make information obtained from experiments meaningful by associating them with daily life. Studies suggest that active-participant experiments conducted by children become more meaningful when they are associated with daily life. In their experiments, children may make generalizations on the basis of simple results (Şahin, 2000). Experiments help young children distinguish certain and uncertain information (Piekny and Maehler, 2013). Preschool teachers believe that activities like experiment, observation and estimation develop children’s scientific skills (Maier et al., 2013). Experimental activities at playschools allow children to comprehend oppositions and differences between various situations (Andersson and Gullberg, 2014).

The study results also reveal that the greatest increase in correct answers is observed in the necessity for measuring weight with dynamometer. While the number of children who gave correct answers in the pretest was 9; this number rose to 15 in the posttest. The minimum increase was observed in the necessity for measuring length with meter. In this context, while the number of children who gave correct answers to this question in the pretest was 14; this number rose to 15 in the posttest.

Recommendations

As a consequence, this study reveals that active-participant experiments conducted by preschool children are effective upon the skills of children to recognize measuring instruments and produce permanent results. In this context, the following recommendations are made.

- Science activities play an important role for preschool children to explore their world. Thus, preschool educational environments should be supported by science materials.
- Competence of preschool teachers concerning science education activities should be increased and vocational training opportunities should be given to teachers on this subject.
- Science education activities should be involved more in preschool curriculums.
- Parents should be supported for home-centered science education activities and for acquiring knowledge and skills on this subject.
- Social awareness should be increased by organizing scientific activities like conference, congress and workshop aimed at preschool science activities.
- Other scientific process skills of children should be investigated via active-participant experiments.

References

Akman, B., Üstün, E., & Güler, T. (2003). Altı yaş çocuklarının bilim süreçlerini kullanma yetenekleri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, (24), s. 11-14.

Aktaş Arnas, Y. (2002). Okul öncesi dönemde fen eğitiminin amaçları. *Çocuk Gelişimi ve Eğitim Dergisi*, 6-7, 1–6.

Alisinanoğlu, F. Özbey, S. & Kahveci, G. (2011). *Okul öncesinde fen eğitimi*. Ankara: Maya Akademi.

Andersson, K. & Gullberg, A. (2014). What is science in preschool and what do teachers have to know to empower children? *Cultural Study of Science Education*, 9, 275–296.

Ayvacı, H. Ş., Devecioğlu, Y. & Yiğit, N. (2002). Okul öncesi öğretmenlerin fen ve doğa etkinliklerindeki yeterliliklerinin belirlenmesi. *V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi Bildiri Kitabı*, ODTÜ, Ankara.

Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In W. Damon & R. Lerner (Eds.), *Handbook of child psychology*. NJ: Wiley.

Brunton, P. & Thornton, L. (2010). *Science in the early years: building firm foundations from birth to five*. UK: Sage Publications

Büyüköztürk, Ş., Çakmak E. K., Akgün Ö. E., Karadeniz, Ş. & Demirel, F. (2010). *Bilimsel araştirma yöntemleri.* Ankara: Pegema Yayıncılık.

California Department of Education (2013). *Preschool learning foundations*. Volume 3. Sacramento: CA.

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Chang, N. (2012). The role of drawing in young children's construction of science concepts. *Early Childhood Education Journal,* 40, 187–193.

Charlesworth, R.& Lind, K. (2003). *Math and science for young children* (Fourt Ed.) NY: Delmar.

Choi, B. (2016). *Early science learning among low-income Latino preschool children: The role of parent and teacher values, beliefs, and practices*. Doctoral Thesis, Unıversıty Of Calıfornıa, San Dıego: USA.

Curriculum For Excellence. (2004). Education Scotland. 15 Kasım 2016 tarihinde http://www.educationscotland.gov.uk/Images/all_experiences_outcomes_tcm4-539562.pdf sayfasından erişilmiştir.

Dere, H. & Ömeroğlu, E. (2001). *Okul öncesi eğitimde fen, doğa, matematik çalışmaları*. Ankara: Anı Yayıncılık.

Early, D., Iruka, I., Ritchie, S., Barbarin, O., Winn, D., Crawford, G. & Pianta, R. (2010). How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, 25(2), 177–193. doi:10.1016/j.ecresq.2009.10.003

Erden, F., & Sönmez, S. (2011). Study of Turkish preschool teachers' attitudes toward science teaching. *International Journal of Science Education*, 33(8), 1149-1168.doi:10.1080/09500693.2010.511295.

Erten, S., Kiray, S.A., & Sen-Gumus, B. (2013). Influence of scientific stories on students ideas about science and scientists. *International Journal of Education in Mathematics, Science and Technology*, 1 (2), 122-137.

Fleer, M. (2013). *Play in the early years*. Port Melbourne: Cambridge University Press.

Gelman, R., Brenneman, K., MacDonald, G., & Román, M., (2010). *Preschool pathways to science: ways of doing, thinking, communicating, and knowing about science.* Baltimore: Brookes Publishing.

Güler, D. & Bıkmaz, F. (2002). Anasınıflarında fen etkinliklerinin gerçekleştirilmesine ilişkin öğretmen görüşleri. *Eğitim Bilimleri ve Uygulamaları*, 1(2), 249-267.

Kabadayı, A. (2006) Analyzing the preschoolers' overgeneralizations of object labeling in the process of mother tongue acquisition in Turkey. *Early Child Development and Care*, 176 (7), 723-734.

Kabadayı, A. (2012). Okulöncesi çocuklarının türkçe ediniminde yaptıkları kurallaştırma hatalarının incelenmesi. *Turkish Studies,* 1561-73.

Kallery, M. (2004). Early years teachers' late concerns and perceived needs in science: An Exploratory Study. *European Journal of Teacher Educations*, 27(2) 147-164.

Kandır, A., Can Yaşar, M., İnal, G., Yazıcı, E., Uyanık, Ö. & Yazıcı, Z. (2012). *Etkinliklerle bilim eğitimi.* Ankara: Efil Yayınevi.

Karaer, H. & Kösterelioğlu, M. (2005). Amasya ve Sinop illerinde çalışan okulöncesi öğretmenlerin fen kavramlarının öğretilmesinde kullandıkları yöntemlerin belirlenmesi. *Kastamonu Eğitim Dergisi*, 13 (2), 447-454.

Karamustafaoğlu, S. & Üstün, A. (2006). Okul öncesi öğretmen adaylarının fen ve doğa etkinliklerini uygulayabilme düzeylerinin belirlenmesi. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 21, 15-23.

Keil, C., Haney, J. & Zoffel, J. (2009). Improvements in student achievement and science process skills using environmental health science problem-based learning curricula. *Electronic Journal of Science Education*, 13 (1), 1-18.

Maier, M. F., Greenfield, D. B., & Bulotsky-Shearer, R. J. (2013). Development and validation of a preschool teachers' attitudes and beliefs toward science teaching questionnaire. *Early Childhood Research Quarterly,* 28 (2), 366-378. DOI: 10.1016/j.ecresq.2012.09.003

Martin, D. J., Jean-Sigur, R., & Schmidt, E. (2005). Process-oriented inquiry—a constructivist approach to early childhood science education: teaching teachers to do science. *Journal of Elementary Science Education*, 17(2), 13–26.

Martin, D. J. (2011). *Elementary science methods: a constructivist approach*. USA: Cengage Learning.

Merino, C., Olivares, C., Navarro, A., Ávalosb, K. & Quirogac, M. (2014). Characterization of the beliefs of preschool teachers about sciences. *Procedia - Social and Behavioral Sciences* 116 (2014) 4193 – 4198

NAEYC (2014). *Early Childhood Program Standards and Accreditation Criteria & Guidance for Measuring.* 10 Ekim 2016 tarihinde http://www.naeyc.org/files/academy/file/AllCriteriaDocument.pdf adresinden erişilmiştir.

National Research Council. (2007). *Taking science to school*. Washington, DC: National Academies Press.

Özbek, S. (2009). *Okul öncesi öğretmenlerinin fen eğitimine ilişkin görüşleri ve uygulamalarının incelenmesi*. Yüksek Lisans Tezi, Çukurova Üniversitesi Sosyal Bilimler Enstitüsü, Adana.

Özbey, S. (2006). *Okul öncesi eğitim kurumlarında görev yapan öğretmenlerin fen etkinliklerine ilişkin yeterliliklerinin belirlenmesi*. Yüksek Lisans Tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Saçkes, M., Trundle, K. C., & Flevares, L. M. (2009). Using children's literature to teach standard-based science concepts in early years. *Early Childhood Education Journal*, 36, 415–422.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, 43 (4), 1045-1050.

Sundberg, B., & Ottander, C. (2014). *"Science in Preschool – A Foundation for Education for Sustainability? A View from Swedish Preschool Teacher Education." In Research in Early Childhood Education for Sustainability.* International Perspectives and Provocations, edited by S. Elliott and J. Davis, 280–293. Routledge.

Şahin, F. (2000). *Okul öncesinde fen bilgisi öğretimi ve aktivite örnekleri*. İstanbul: Ya-Pa Yayınları.

Taş, I. (2010). *Etnografik bakış açısıyla kırsal kesimde okul öncesi fen eğitimine yönelik bir durum çalışması*. Yüksek Lisans Tezi, Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü, Eskişehir.

TDK. (2016). *Büyük Türkçe sözlüğü*. Ankara: TDK Yayınları.

Uyanık Balat, G. & Önkol, F.L. (2010). Okul öncesi dönemde fen eğitimi öğretim yöntemleri. B.Akman, G. Uyanık Balat & T. Güler (Editörler), *Okul öncesi dönemde fen eğitimi* (s.89-126). Ankara: Pegem Akademi.

Ünal, M. & Akman, B. (2006). Okul öncesi öğretmenlerinin fen eğitimine karşı gösterdikleri tutumlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30, 251-257.

Yıldırım, A. & Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin Yayıncılık.

Yıldız, E., Aydoğdu, B., Akpınar, E. & Ergin, Ö. (2006). Fen bilgisi öğretmenlerinin fen deneylerine yönelik tutumları. *Boğaziçi Üniversitesi Eğitim Dergisi*, 24 (2), 71-86.

Author Information

Özkan Sapsağlam
Gaziosmanpaşa University
Faculty of Education
Contact e-mail: *ozkaanim@gmail.com*

Aykut Emre Bozdoğan
Gaziosmanpaşa University
Faculty of Education

Appendix 1. Sample Questions about Measuring Instruments

Keloğlan pazardan aldığı elmaların ağırlığını ölçmek istiyor. Sizce aşağıdaki ölçme araçlarından hangisini kullanmalıdır?

Keloğlan wants to measure the weight of the apples he had bought from the market place. Which of the following measuring instruments should he use?

Tom ocaktaki suyun kaynaması için ocağı yakmış ve su kaynamaya başlamış. Sizce kaynayan bu suyun sıcaklığını hangi araç ile ölçebilir?

Tom burns the oven to boil water and the water starts boiling. Which of the following measuring instruments can measure the temperature of the boiling water?