Effects of an SWH Approach and Self-Evaluation on Sixth Grade Students' Learning and Retention of an Electricity Unit

Esra Kabataş Memiş*
Kastamonu University

Sabriye Seven**
Atatürk University

Abstract
The purpose of this study is to explore the effects of guided, inquiry-based laboratory activities using the Science Writing Heuristic (SWH) approach and self-evaluation on students' science achievement. The study involved three sixth grade classes studying an electricity unit taught by the same primary school teacher. Before the study began, one class was randomly selected to be the control group, and the other two classes were selected to be treatment groups. In the control group, students were instructed using a traditional didactic approach. Treatment groups engaged in guided, inquiry-based activities via the SWH approach. One treatment group was randomly selected to complete a self-evaluation of their SWH reports. Data collection tools included a baseline test at the beginning of the study to establish three skill-based groups and unit-based pretests, posttests, and retention tests. The Cronbach's alpha reliability of the electricity test was .91. Results indicated no significant mean differences among groups on pretest measures for the unit. Analysis of post and retention tests indicated that students in the SWH and self-evaluation SWH groups scored significantly higher than the students in the control group.

Keywords*: science writing heuristic approach (SWH), guided inquiry, self-evaluation

* Esra Kabataş Memiş is an assistant professor at the department of Science Education, Kastamonu University, Turkey.

** Sabriye Seven is a professor at the department of Science Education, Atatürk University, Turkey,

Correspondence: ekmemis@kastamonu.edu.tr

Introduction

One of the purposes of science education is to raise individuals who are aware of their own cognitive processes and learn in a conscious style (National Research Council, 1996; Ministry of National Education, 2006); thus, metacognition plays an important role. Metacognition is described by Flavell (1979) as knowledge and regulation of cognitive activities by an individual during the learning process. Hewsen, Beeth, and Thorley (1998), on the other hand, defined metacognition as the possession of control over an individual's own knowledge and conceptual development. Viewed from this perspective, metacognition has a significant relationship with conceptual development, and both factors are integral to the nature of learning. Concrete metacognition skills may be revealed through a learner's statements and behaviors, such as activating prior knowledge, goal setting, time management, expressing non-understanding, commenting on one's own activities, relating answers to questions, recapitulating, and drawing conclusions (van der Stel & Veenman, 2010). Further, metacognition, which contains the processes of planning how to approach a learning task, then monitoring (Bağ, Uşak, & Caner, 2006), understanding, and evaluating it (Çakıroğlu, 2007), is a significant determiner of learning performance. Developing metacognitive skills is an important step forward in educational environments (Sungur & Senler, 2009; van der Stel & Veenman, 2010), where learners currently have few related opportunities.

Learning environments must be created where students are responsible for their own learning and deal with open-ended tasks (NRC, 1999; Sungur & Senler, 2009; Eryaman et al., 2010). This concept was explicitly stated when the Turkish education system was modified in 2005 and science was described as a manner of research and thinking based on experimental criteria, logical thinking, reasoning, and constant inquiry; it was also emphasized that providing various learning environments and experiences was essential (MNE, 2006). This study applied the science writing heuristic (SWH) approach of argumentation and inquiry, where students are responsible for their own learning, with the goals of making them more aware of their own cognitive processes and improving their metacognitive skills through self-evaluation.

The purpose of science education is not only to provide learners with scientific concepts; argumentation is also significant, for it explains how scientific discourse should proceed (Kuhn, 2010). Lack of argumentation leads to perceiving science concepts as a sum of static events (Zohar & Nemet, 2002). Cavagnetto (2010) suggested that argumentation is a significant language application of science, emphasizing how it plays a key role in the understanding and safe production of new knowledge. The role of argumentation is also major in acquiring socially constructed knowledge (Baker, 2009; Driver, Newton, & Osborne, 2000), and it can help learners to understand the process of socially constructing scientific knowledge better, since they are in constant interaction (Baker, 2009; Schwarz, 2009). In order for students to participate in scientific argumentations and make correct decisions, they need to understand the nature of such discussions and practice valid methods on scientific content (Schwarz, 2009). Argumentation is part of the inquiry process and becomes a center for learning science by providing interaction for both the individual and the group. Driver et al. (2000) have strongly argued that argumentation is a central component of science education that will help students make decisions now and in the future through three fundamental formats: developing conceptual understanding (Dawson & Venville, 2010), understanding scientific epistemology (Nussbaum, Sinatra, & Poliquin, 2008), and enhancing research capacity (Kelly, Druker, & Chen, 1998; Kim & Song, 2006). Research is essential to the nature of science, and it is crucial for students to think and reason during the research and argumentation process (Newton, Driver, & Osborne, 1999). Scientific argumentation can be taught using certain templates (Cavagnetto, 2010) that promote actively researching concepts, such as in the 3, 4, 5, 7-E models (Barman, 1989; Ramsey, 1993) or Kolb's (1984) Learning Cycle. One of the templates used to achieve conceptual learning is the Science Writing Heuristic (SWH) approach. This study aims to

investigate whether embedding self-evaluation into the argument-based SWH approach makes it more effective on student achievement or permanence of learning.

The Science Writing Heuristic Approach

Hand and Keys (1999) developed this approach to define the structure of scientific arguments in order to improve them in education. It consists of a framework to guide activities as well as a metacognitive support to prompt reasoning and writing about data. Further, the activities and metacognitive scaffolds seek to provide authentic, meaningful opportunities for learners (Hand, 2008). The SWH approach serves as a bridge between formal and informal knowledge in science and enables students to consider structure and collaboration in scientific activities, discussions, and concepts within the framework of explanation and interpretation (Akkuş et al., 2007). It forms an effective, learner-centered environment enhanced by written and oral argumentation (Hand & Keys, 1999; Keys, Hand, Prain, & Collins, 1999).

The SWH approach consists of two parts (Hand, 2008): a teacher template and a student template (see Figure 1). The teacher template includes suggested activities for teachers when using inquiry activities and emphasizes phases of negotiation to facilitate for students. The template for students directs them to generate questions, claims, and evidence and to compare findings with other sources, including peers, the Internet, or the textbook. The student template also encourages learners to reflect on how their ideas change during an activity, asking questions to prompt scientific thinking. Beginning questions are crucially significant in developing a scientific argument, shaping relationships among questions, claims, and evidence (Choi, Notebaert, Diaz, & Hand, 2007).

The Science Writing Heuristic, Part I: A template for teacher-designed activities to promote laboratory understanding	The Science Writing Heuristic, Part II: A template for students
Exploration of pre-instructional understanding (e.g., individual or group concept mapping).	Beginning ideas – What are my questions?
Pre-instructional activities (e.g., informal writing, making observations, brainstorming, and posing questions).	Tests – What did I do?
Participation in science activity.	Observations – What did I see?
Negotiation phase I – assigning personal meanings for science activity (e.g., writing journals).	Claims – What can I claim?
Negotiation phase II – sharing and comparing data interpretations in small groups (e.g., making a group chart).	Evidence – How do I know? Why am I making these claims?
Negotiation phase III – comparing science ideas to textbooks or other printed resources (e.g., writing group notes in response to focus questions).	Reading – How do my ideas compare with others?
Negotiation phase IV – individual reflection and writing (e.g., creating a presentation such as a poster or report for a larger audience).	Reflection – How have my ideas changed?
Exploration of post-instructional understanding through concept mapping.	

Figure 1. The two templates for the SWH: the teacher template and the student template

Language applications are the center of the SWH approach. Teachers support the learning of students at all levels by applying language applications to scientific inquiry. Norton-Meier (2008) has noted how this basic idea suggests that there is no science without language. SWH applications require language presentations via different modes (speaking, listening, writing, connecting, and visuals) and forms (fairy tales, letters, poems) to different audiences (Norton-Meier, Nelson, Hockenberry, & Wise, 2008). Norton-Meier (2008) also pointed out that individual and social thinking and writing take place within inquiry activities in the class environment, yet scientific argument is required for students to understand these activities by explaining their claims with scientific evidence. Being involved in scientific argument helps students to understand meaningful concepts and how science and people can work together to improve natural world development (Hofstein & Lunetta, 2004). As the SWH approach embeds science argument within inquiries undertaken by students (Hand, 2008), students need to understand argumentation to conduct this process. Therefore, the SWH approach is designed to provide scaffolding for purposeful thinking about relationships between question, evidence, and claims (Hand, 2008). Recent studies support this view (Hand, Norton-Meier, Jay, & Bintz, 2009; Nam, Choi, & Hand, 2010; Norton-Meir, Hand, Hockenberry, & Wise, 2008).

The SWH approach is a template used by learners who apply the procedures of inquiry, writing, critical thinking, conceptual understanding, and thinking about skills (Hand, Wallace, & Yang, 2004; Hohenshell, 2008). In addition, it helps teachers organize laboratory sessions and inform students how to write lab reports (Omar, 2004; Mohammed, 2007; Hand, 2008). Within this context, the SWH approach could also be called an alternative writing style that enables learners to think about and discuss science concepts. The approach contains various cognitive and metacognitive writing activities for science learning and enables students to connect data, methods, evidence, and claims; to formulate claims; and to blend processes such as support in writing (Hand et al., 2004; Hohenshell, 2008). Students will better understand concepts addressed in class when this approach is well planned and implemented (Keys et al., 1999; Omar, 2004).

Comparisons of the SWH approach with traditional styles have resulted in a better understanding of scientific concepts (Burke, Hand., Poack., & Greenbowe, T., 2005; Keys et al., 1999; Hand et al., 2004; Hsieh, 2005; Günel, Kabataş Memiş, & Büyükkasap, 2010; Nam et al., 2010; Schroeder, 2008; Kıngır, Geban, & Günel, 2012, 2013). Implementation of the SWH approach has helped students to improve conceptual understanding, meaning-making, and reasoning abilities as well as critical thinking.

Self-Evaluation

Assessment in education should feature a formative style that measures the process, rather than a summative style focusing only on the outcome. Formative assessment can define potential in new education programs or when seeking to enhance or improve applications. It includes traditional teacher evaluation, self-evaluation, and peer evaluation. In the center of individual learning or in learner-centered cultures, learners are supposed to want to learn, be aware of their learning, and take responsibility for it (Sebba, Deakin Crick, Yu, Lawson, & Harlen, 2008). Student self-evaluation has been found to increase learning significantly (Özoğul, Olina, & Sullivan, 2008). Self-evaluation includes judging one's own achievements and learning process (Sebba et al., 2008). Taras (2001) has suggested that self-evaluation is crucial, as it promotes life-long learning, professional development, and effective learning. It should always be put forward and to ensure that students take it seriously, they should be informed how its fundamental purpose is to develop learning; evaluations should also be considered within this context (Davies, 2002).

Günel, Hohenshell, and Hand (2006) stated that metacognition is the center of self-evaluation. From a metacognitive perspective, students formatively deal with self-evaluating in accordance with

defined rules to indicate mastery in expressing their interests and revealing their inadequacies (Özoğul et al., 2008). They take responsibility for their own learning and personal success. In addition, self-evaluation helps students exhibit their own development (Olina & Sulivan, 2004), which is a potential to improving performance (Özoğul et al., 2008). In their quasi-experimental study, Andrade, Du, and Wang (2008) investigated the effects of writing a model, creating a list of criteria, and self-evaluation in accordance with a rubric using that criteria. They determined that self-evaluation based on a graded rubric related to high scores; it helped make sense of writing and, when used actively by students, created significant quality.

The Study

Building from the literature, the main purpose of this study was to explore the effect on student achievement levels of the SWH approach with embedded self-evaluation compared to traditional teaching approaches. In the SWH approach, students engage in metacognitive activities (Hand et al., 2004). Under the Turkish education system, science lessons allow for students to experience an active process where they personally construct knowledge. Students are interested in hand-on activities and experiments, and they reach results rapidly in lessons employing the active SWH approach. They experience processes such as questioning, explaining, analyzing, recognizing support or opposition, reasoning, and mounting an argument. In addition, the SWH student template presents a non-traditional writing technique and reflects teaching by conducting reasoning processes. Thus, within the scope of the present study, a self-evaluated SWH group was formed in addition to the non-self-evaluated SWH group and control group in order to increase certain students' awareness of the process. Consequently, we explored the impact of each of the three implementations on student achievement.

Method

A mixed method design was applied to this study. The study—in which a quasi-experimental pretest and posttest model was used—was implemented in three different sixth grade classes. Researchers randomly selected two of these classes as treatment groups (one SWH and one self-evaluated SWH) and one as the control group.

Participants

This study was implemented in three different sixth grade classes taught by a single teacher at a public primary school in the eastern part of Turkey. The students were generally of middle class socio-economic status and consisted of 51 females and 57 males, totaling 108 students. One class was randomly selected as the control group (35 students, 18 males and 17 females) and the other two as treatment groups (37 students, 22 males and 15 females; 36 students, 20 males and 16 females). The teacher taught the same content in 40-minute classes four times a week. The teacher participating in the research had five years of professional experience but was applying the SWH approach for the first time. The teacher also had a master's degree in science education and was pursuing a doctorate at the time of the study.

Treatment and Control Groups

Treatment Groups

Students in both treatment groups dealt with explorative, inquiry-based activities in a school laboratory environment and reported on their activities individually using the SWH student template. One treatment group also conducted self-evaluations according to a rubric prepared by the researcher. The implementation was carried out on an electricity unit, which was taught over six weeks (four course hours

a week). Students completed five guided-inquiry activities in small groups (four or five people) and took part in classroom discussions. In the introduction, the teacher attracted students' attention with an activity and provoked the question "Why?" Later, a class discussion was initiated that included all students. Students were provided with the opportunity to act independently within enhanced limits, to identify what they were curious about, and to reflect within their groups and as a class.

Control Group

Control group students received education via the traditional approach, where teachers gave information directly, as students listened and answered questions from time to time. A subject or theme was followed from a text book, and chapter questions were answered as individual activities. Control group students only watched activities carried out by their teachers in the classroom and conducted no individual or group experiments or discussion. When necessary, the teacher demonstrated experiments to the whole group.

Application Topic

Electricity was selected as a research topic since it is a basic science topic across all levels of schooling (MNE, 2006) with large coverage and common misconceptions (Akdeniz, Bektaş, & Yiğit, 2000; Çepni & Keleş, 2006; Sencar & Eryılmaz, 2004). This study is limited to a single electricity unit. Basic subjects included concepts of conductors and non-conductors, simple circuits and series/parallel circuits, electrical resistance, and bulb brightness and resistance.

Data Collection Tools

Baseline Test

Questions taken from the NEAPS (National Assessment of Educational Progress) and TIMMS (Trends in International Mathematics and Science Study) tests were used to investigate differences among the students' science achievement levels. The baseline test asked 4 chemistry, 6 biology, and 10 physics questions. This test was constructed by the SWH research team at Iowa State University, and its reliability was found to be .75 (Gunel, Akkuş, Hohenshell, & Hand, 2004). After the translation of the test, the researcher examined whether the questions were suitable for the curriculum content established by the Ministry of National Education, and two questions were removed. The test was examined by four specialized researchers in terms of suitability for curriculum and content validity, and corrections were made. Finally, another researcher in Turkish education was asked to examine the text in terms of semantics and orthography, and corrections were made once more. The final test consisted of 18 questions. It was administered to sixth grade students at a different 54 –people in the elementary school as a pilot study. The test's Cronbach's alpha reliability coefficient was .71. The final test was used to determine differences between students in terms of science achievement and to categorize students into groups (low, medium, and high science achievement). See Figure 2 for sample baseline test items.

Figure 2. Baseline test sample items.

Electricity Pretest, Posttest, and Retention Test

The content planned for the sixth grade unit (simple electrical circuits, conductor and non-conductor materials, resistance, bulb brightness) was taken into consideration as the test was prepared. A total of 20 questions were selected from an appropriate pool of multiple choice questions released by the National Standardized Tests. Conceptual questions were also prepared where students were asked to write sentences to demonstrate justification, developing arguments, reasoning, and explaining processes. Because this type of question was not common within the teacher's or school's current practices, the researchers balanced the demands of the research with the practical nature of working within the school system and limited the test to the following four conceptual questions:

1. The power fails while you are doing your homework at home. You need light in order to do your homework. Your mother says that there is only a wire, a battery, and a lamp that you can use. Can you provide light with these materials? How? (Explain the electrical circuit you make by drawing it.)
2. Your teacher wants you to create an electrical circuit on a wood ground. You bring it to school to show it to your teacher after you complete it. While you are preparing to demonstrate the circuit, you notice that the switch has been broken. What material(s) would you use to complete the circuit until you could insert a new switch? Why have you chosen those materials? Explain.
3. People around you consider you an electrical expert. How would you explain to them how a lamp produces light?
4. Ali calls the police and tells them that his child at home has been shocked as a result of an accident. When the policemen arrive, the child is lying on the floor next to a power outlet. He has a screwdriver in his hand and is wearing slippers. One of the policemen glances around, but he does not see anything metal near the child. Policeman: Is anyone else at home?
Ali: No.
Policeman: Then arrest this man!
How did the policeman realize that Ali is the murderer?

Content validity (Balcı, 2004; Çepni, 2005; Karasar, 2004) of the electricity test was considered by five experts, and modifications were made. After this process, the test was administered to a group of 72 sixth grade students from an inner city middle school in the eastern part of Turkey. The inter-rater reliability was found to be .91. This finalized test was used as a pretest, posttest, and retention test, which was administered eight months after the original application. The answer key for the open-ended questions was prepared by a researcher with teaching experience in the subject area who also scored all responses. When randomly selected answer sheets were scored by another researcher and the teacher, inter-rater reliability was observed to be 90%.

SWH reports and Self-Evaluation Rubric

When completing SWH applications, students in the experiment group prepared reports in accordance with the student template. This report had several parts: beginning ideas, where students wrote initial questions to research; test, where they experimented with their questions; observations, where they wrote about their findings; claims, where they asserted their opinions; evidences, where they cited resources; readings, where they compared and contrasted results; and reflections, where they indicated personal changes they experienced during the process. These reports were submitted after each activity.

A 4-point Likert scale rubric was prepared for students to evaluate their own SWH reports, named the self-evaluation rubric. The main purpose of this rubric was to help students improve their SWH reports. The researcher selected 12 criteria to define the rubric questions, and relevance was checked by another researcher with applications in this field. Criteria were revised in light of this review, and the self-evaluation rubric was finalized. Only one of the three groups applied this rubric to their SWH reports, and these evaluation scores were not used as data in the study.

Results

Baseline Test

Analyses indicated that group performances on the independent variable of baseline score were not statistically different ($F(2, 98) = 2.980$, $p = 0.056$; see Table 1 for M and SD distribution of groups).

Table 1
Group Distribution and Baseline Test Scores

Groups	n	M	SD
Control	28	21.41	8.32
SWH	34	25.85	8.38
Self-evaluated SWH	35	22.37	7.95
Total	97	23.30	8.35

The researchers used this baseline test to establish groups based on achievement. Students were categorized as having low, medium, or high skill levels in relation to test scores. Scores one-half standard deviation around the mean ($\bar{X} - \frac{1}{2}$SD, $\bar{X} + \frac{1}{2}$SD) represented the medium achievement level, a score one-half standard deviation below the mean or lower ($\bar{X} - \frac{1}{2}$SD and below) indicated low achievement level, and a score one-half standard deviation above the mean ($\bar{X} + \frac{1}{2}$SD) indicated high achievement (see Table 2). The overall total distribution of students was approximately equal for each achievement level.

Table 2
Achievement Level Distribution of Groups

Group	Low (n)	Medium (n)	High (n)
		Achievement Level	
Control	10	14	5
SWH	6	14	14
Self-evaluated SWH	15	10	10
Total	31	38	29

Pretest

Analyses indicated that group performances on the independent variable of pretest total score were not statistically different ($F(2, 106) = 0.310$, $p = 0.734$). No other significant differences among classes were found (see Table 3 for *M* and SD distribution of groups).

Table 3
Group Distribution and Pretest Scores

Group	n	M	SD
Control	33	19.82	7.11
SWH	32	18.84	5.24
Self-evaluated SWH	34	19.32	6.00

Posttest

Results showed significant differences among the groups on the posttest total score ($F(2, 95) = 6.873$, $p < .01$), where students in SWH group ($M = 53.76$, $SD = 15.09$) scored significantly higher than students in the control group ($M = 39.59$, $SD = 14.33$), $t(63) = 3.88$, $p < .01$; and students in the self-evaluated SWH group ($M = 47.67$, $SD = 16.76$) scored significantly higher than students in the control group ($M = 39.59$, $SD = 14.33$), $t(63) = 2.09$, $p < .01$). Significant differences were also found on the posttest total conceptual questions ($F(2, 95) = 10.169$, $p < .01$),where students in the SWH group ($M = 22.30$, $SD = 9.22$) scored significantly higher than students in the control group ($M = 12.96$, $SD = 6.85$), $t(63) = 4.64$, $p < .01$; and students in the self-evaluated SWH group ($M = 20.12$, $SD = 9.74$) scored significantly higher than students in the control group ($M = 12.96$, $SD = 6.85$), $t(63) = 3.44$, $p < .01$.

Analysis of the conceptual questions indicated significant differences among the groups on question 1 ($F(2, 95) = 6.744$, $p < .01$), where students in the SWH group ($M = 8.67$, $SD = 2.39$) scored significantly higher than students in the control group ($M = 6.00$, $SD = 3.22$), $t(63) = 3.79$, $p < .01$ and significantly higher than students in the self-evaluated SWH group ($M = 7.18$, $SD = 3.13$), $t(64) = 2.17$, $p < .01$. Significant differences were also determined for scores of conceptual question 3 ($F(2, 95) = 7.966$, $p < .01$), where students in the SWH group ($M = 4.85$, $SD = 3.60$) scored significantly higher than students in the control group ($M = 1.91$, $SD = 2.22$), $t(63) = 3.98$, $p < .01$; and students in the self-evaluated SWH group ($M = 6.09$, $SD = 4.40$) scored significantly higher than students in the control group ($M = 1.91$, $SD = 2.22$), $t(63) = 4.86$, $p < .01$. For conceptual question 4 ($F(2, 95) = 3.471$, $p < .01$), students in the non-self-evaluated SWH group ($M = 2.76$, $SD = 2.90$) scored significantly higher than students in the control group ($M = 1.06$, $SD = 2.37$), $t(63) = 2.59$, $p < .01$; and students in the self-evaluated SWH group ($M = 2.73$, $SD = 3.49$) scored significantly higher than students in the control group ($M = 1.06$, $SD = 2.37$), $t(63) = 2.26$, $p < .01$.

Students were categorized into low, medium, and high groups in relation to their general achievement levels at the beginning of the study. Mean scores on the electricity unit posttest for these

three achievement groups are displayed in Figure 3. In the graph, the lowest mean among students in the low achievement group is within the control group, whereas mean scores of the treatment groups are similar and higher than those of the control group. In the medium achievement level group, the control group again has the lowest scores, and the treatment groups have higher mean scores. Among the two treatment groups, the mean score of the SWH group was higher than the mean score of the self-evaluated SWH group. Finally, scores among the high achievement level group were, from lowest to highest, control group, SWH group, and self-evaluated SWH group.
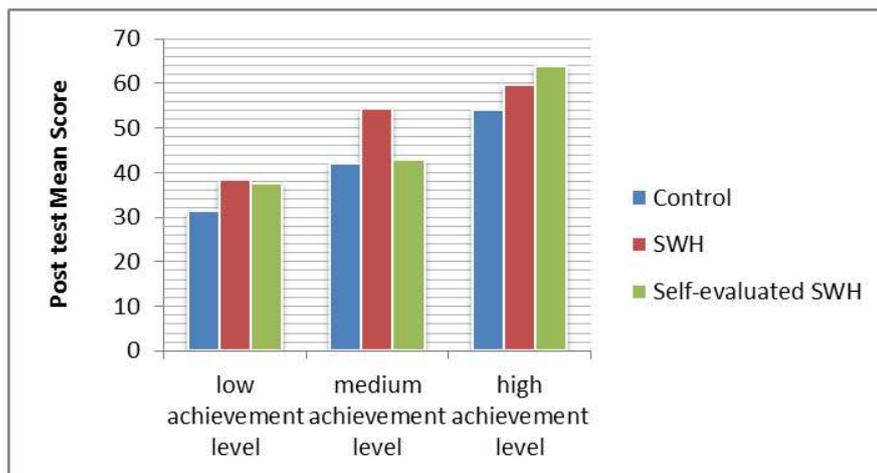
Figure 3. Distribution of Posttest Scores According to Achievement Levels

Retention Test

For the retention test, significant mean differences were found among the groups on total score ($F_{(2, 86)} = 3.264$, $p < .01$), where students in the self-evaluated SWH group (M = 44.46, SD = 16.09) scored significantly higher than students in both the non-self-evaluated SWH group (M = 35.97, SD = 12.89), $t(59) = 2.29$, $p < .01$, and the control group (M = 36.39, SD = 14.81), $t(58) = 2.02$, $p < .01$. Significant mean differences were also found on the conceptual questions' total score ($F_{(2, 86)} = 7.31$, $p < .01$), where students in the self-evaluated SWH group (M = 20.19, SD = 9.64) scored significantly higher than students in the SWH group (M = 11.65, SD = 8.06), $t(59) = 3.76$, $p < .01$, and the control group (M = 14.32, SD = 9.54), $t(58) = 2.37$, $p < .01$.

Analyses of each conceptual question indicated significant mean differences among the groups on conceptual question 1 ($F_{(2, 86)} = 3.098$, $p < .01$), where students in the self-evaluated SWH group (M = 7.13, SD = 3.40) scored significantly higher than students in the SWH group (M = 5.34, SD = 3.76), $t(59) = 1.94$, $p < .01$, and the control group (M = 5.11, SD = 3.21), $t(58) = 2.37$, $p < .01$. For conceptual question 2 ($F_{(2, 86)} = 4.801$, $p < .01$), students in the self-evaluated SWH group (M = 7.5, SD = 4.02) scored significantly higher than students in the SWH group (M = 4.38, SD = 4.45), $t(59) = 2.86$, $p < .01$, and the control group (M = 4.64, SD = 4.70), $t(58) = 2.60$, $p < .01$. On conceptual question 3 ($F_{(2, 86)} = 5.110$, $p < .01$), students in the self-evaluated SWH group (M = 3.53, SD = 3.10) scored significantly higher than students in the SWH group (M = 1.55, SD = 2.37), $t(59) = 2.82$, $p < .01$, and the control group (M = 1.68, SD = 2.50), $t(58) = 2.56$, $p < .01$.

Figure 4 shows mean scores of the retention test compared to general achievement levels of students in each group. Self-evaluating SWH students had the highest mean scores among the groups at all achievement levels.
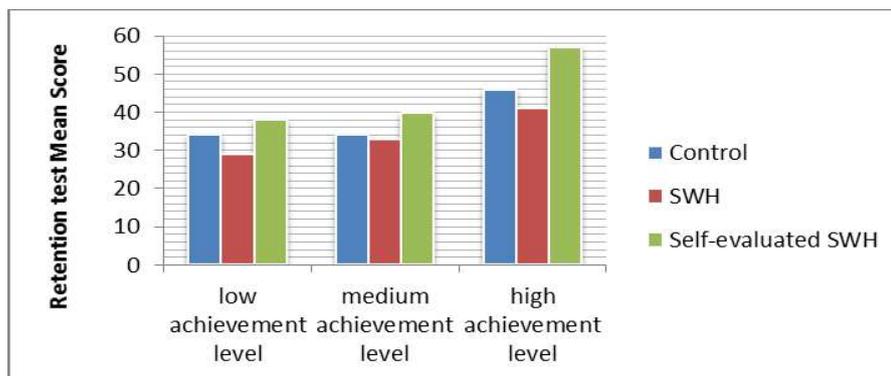
Figure 4. Distribution of Retention Test Scores According to Achievement Levels

Cohen's d Effect Sizes on Posttest and Retention Test

Cohen's d index is widely used in the social sciences. The use of d is not only a necessity demanded by the practical requirements of table making, but it proves salutary in those areas of the behavioral sciences where raw units are arbitrary, lack meaning outside the investigation, or both (Cohen, 1998). Cohen defined effect sizes as small ($d = .2$), medium ($d = .5$), and large ($d = .8$). We applied this analysis to better describe the differences among means. The effect sizes of the groups are given in Table 4. In the posttest, a small effect size was found on multiple choice questions among the control and SWH groups, and a large effect size was found on concept questions and total test score in favor of the SWH group. Between the control and self-evaluated SWH groups, a large effect size was found on concept questions in favor of the control group, whereas this effect was medium-sized for the total test score. Between the two treatment groups, a small effect size was found in favor of the non-self-evaluated SWH group on all three categories for the posttest. Similarly, on the retention test, between the control and SWH groups on multiple choice questions, the SWH treatment group had a small effect size, whereas the opposite was true for the concept questions. A small effect size was found for multiple choice questions between the control and self-evaluated SWH groups, whereas a medium effect size was found in favor of the self-evaluated SWH group on concept questions and total scores. Between treatment groups, greater effects were seen for the self-evaluated SWH group on concept questions and total score.

Table 4. *Cohen's d Effect Size on Posttest and Retention Test*

Measure	Groups	Cohen's d	Scale
Posttest MCQ	Control/SWH	.36	Small (SWH)
	SWH/Self-evaluated SWH	.35	Small (SWH)
Posttest CQ	Control/SWH	1.15	High (SWH)
	Control/Self-evaluated SWH	.85	High (Self-evaluated SWH)
	SWH/Self-evaluated SWH	.23	Small (SWH)
Posttest Total	Control/SWH	.96	High (SWH)
	Control/Self-evaluated SWH	.52	Medium (Self-evaluated SWH)
	SWH/Self-evaluated SWH	.38	Small (SWH)
Retention MCQ	Control/SWH	.32	Small (SWH)
	Control/Self-evaluated SWH	.31	Small (Self-evaluated SWH)
Retention CQ	Control/SWH	.30	Small (control)
	Control/Self-evaluated SWH	.61	Medium (Self-evaluated SWH)
	SWH/Self-evaluated SWH	.90	High (Self-evaluated SWH)
Retention Total	Control/Self-evaluated SWH	.52	Medium (Self-evaluated SWH)
	SWH/Self-evaluated SWH	.58	Medium (Self-evaluated SWH)

Discussion and Conclusion

According to the results of the electricity pretest and baseline test, each group began with relatively equal achievement levels, and the low, medium, and high achievement students were distributed within each experiment group. When considering the posttest conducted at the end of the study, a significant difference was found between the SWH and control groups in favor of the SWH group for the conceptual questions total score and on the first, third, and fourth questions. On the retention test, on the other hand, which was conducted eight months later, no statistically significant difference was found between the control and SWH groups. However, during effect size analysis, the SWH group was found to be more effective with a small effect size compared to the control group on multiple choice questions. The open-ended conceptual questions were evaluated separately.

The purpose here is, in case of no difference in total scores, to determine if intervention elicited differences on a conceptual level. When findings were examined, conceptual differences were found strictly in favor of the treatment groups for each test. Thus, this approach develops conceptual understanding. The fact that there had been little difference on the retention test but a noticeable difference on the posttest indicates that this approach supports students' learning on a conceptual level (Keys et al., 1999). The short duration of the experiment may explain why these differences did not achieve high levels. The SWH approach requires students to think during the process, establish relationships among data, interpret data, and communicate conclusions upon connecting claim and evidence; all of these operations improve higher level thinking skills. While it is not appropriate to expect metacognition to be overly effective in short-term applications, for longer applications, significant differences such as those noticed on the posttest may also be observed on the retention test. The literature cites many studies where SWH applications made significant differences on posttest scores (Akkuş et al., 2007; Hand & Keys, 1999; Hohenshell & Hand, 2006; Kabataş, Günel, Büyükkasap, Uzoğlu, & Hand, 2008; Kabataş Memiş, Günel, & Büyükkasap, 2009, Günel et al., 2010).The results of these studies are parallel to the results of the current study. Once students are able to pursue their interests, they are empowered; their motivation towards science comes alive through experiments they design and conduct. Science is no longer a hard subject, and as interest in learning increases, retention is successful. During the SWH approach, students become more cognitively and metacognitively engaged in the learning process. Not only did students indicate willingness to complete the heuristic, a cognitive activity, they also recognized that they were required to make connections between various elements of the heuristic, a metacognitive activity (Hand et al., 2004). Students engaged in asking questions, examining evidence, making claims, and comparing their claims and evidence with current scientific knowledge and peers' opinions. Thus, students used decision-making strategies.

During their self-evaluation process, after the completion of their reports, students conducted self-evaluations of their SWH applications using the rubric prepared by the researcher. These evaluations are not included in the quantitative data. The researcher did help students to ensure consistency and impartiality across the entire SWH process. Considering the posttest's one way ANOVA and effect size findings, as well as the retention test, the self-evaluated SWH group was more successful than the control group. On the retention test, the success of the self-evaluated SWH group compared to SWH group also became clear, and the long-term effect of self-evaluation was revealed. Self-evaluation helps students take responsibility for their own learning and improves individual achievement (Olina & Sulivan, 2004). In addition, self-evaluation is important for learners to be a part of the process, and it provides learners with an outside perspective on themselves (MNE, 2006). A learner notices personal inadequacies through self-evaluation and can correct them before they become permanent. When self-evaluation and SWH were conducted together, even after eight months, almost all learned knowledge was remembered, demonstrating that knowledge in this process is constructed to be more permanent. A similar study was conducted by Günel, Hohenshell, and Hand (2006). They investigated the effect of students' self-

evaluation on non-traditional writing activities as part of the SWH approach, and the self-evaluated group was more successful on posttest scores.

The number of students in each achievement group across all three experiment groups was approximately equal, denoting that classes were homogeneous on general science achievement. Detailed information based on mean scores has been presented in figures for each achievement level. These graphs show that students in the non-self-evaluated SWH group were more successful on the electricity posttest at all achievement levels; compared to the control group, they had higher means at medium and high achievement levels. This result was not repeated for the retention test. Among low and medium achievement level students, non-self-evaluated SWH students received the highest scores, whereas the self-evaluated SWH group received the highest scores for the high achievement level. This result indicates that metacognitive assessments are demanding and explicit instruction might be needed with regard to metacognition. Among students from all three achievement levels, the self-evaluated SWH group was the most successful on mean scores of the posttest. Therefore, self-evaluation for SWH reports was more effective at the high achievement level, since self-evaluation requires using metacognition and students on the high achievement level are better at using metacognition (Rivard, 2004). On the other hand, at the medium achievement level, for students to experience the process on a one-on-one basis considerably improves learning (Figure 3). Students at this achievement level are those who need support when in difficulty and succeed when helped. Therefore, students at this achievement level strongly benefit from SWH applications. The result in our study is parallel to Hsieh (2005), who observed that this application makes the maximum difference for students at the medium achievement level.

Writing is considered to be a tool in creating an argument by organizing information into an effective presentation from a basic theoretical perspective. Many studies have shown that writing activities can improve students' conceptual level understanding. In our study, we asked students to complete the SWH template, including the non-traditional writing activity. For all three achievement levels, self-evaluating students noticed their inadequacies earlier and were able to correct mistakes. High achieving students are generally more successful applying metacognition, which may explain the differences among the self-evaluated SWH group. A similar result was found by Rivard (2004), who stated that students with low achievement levels are better at peer discussions and therefore understand concepts better, whereas students with high achievement levels are better at writing than talking activities. This finding indicates that writing requires using metacognition. In addition, Wallace (2007) has emphasized that one of the most salient features of writing that may promote conceptual understanding is its potential to generate metacognitive thought. When learners write, they discover what they know and what remains a gap. Thus, metacognition is a form of learning produced by writing. On the other hand, Grimberg (2008) observed that high achievement level students use more alternative cognitive operations, although all cognitive categories (perception, making sense, and generalization) were utilized in reports by students from all three achievement levels in the current study. Students dealing with activities targeting the use of metacognition report increased success.

One of the components of the SWH approach is inquiry, which is crucial in reaching new knowledge. For students, being active has significant results, such as improving research skills and providing better understanding about the nature of science (Bilgin, 2009). Learning that is based on inquiry allows students to use scientific process skills and methods (Tatar, 2006). Thus, students learn science best by hand-on experiences when they not only conduct experiments but also work actively in planning, implementing, and evaluating the process. This approach, which allows students to construct knowledge built on previously learned information and express connections meaningfully, is quite appropriate for science classes. Survival of this inquiry process depends on language practices such as verbal expression and writing (Prain, 2007). In light of this study as well as existing literature, learning environments should continue to be analyzed through investigation of verbal and written language

practices and arguments mentally recorded by students. In addition, alternative evaluations should be conducted beyond self-evaluation of students.

References

Akdeniz, A., Bektaş, U., & Yiğit, N. (2000). The 8th grade students' levels of understanding of the introductory physics concepts. *Journal of Hacettepe University Education Faculty*, *19*, 5–14.

Akkuş, R., Günel, M., & Hand, B. (2007). Comparing an inquiry-based approach known as the science writing heuristic to traditional science teaching practices: Are there differences? *International Journal of Science Education*, 1–21.

Andrade, H.L., Du, Y. & Wang, X.(2008). Putting rubrics to the test: the efect of a model, criteria generation, and rubric- referenced self- assesment on elementary school students' writing. E*ducational Measurement: Issue and Practice,* 27(2), 3-13.

Bağ, H., Uşak, M., & Caner, F. (2006). *Metacognition*. In M. Bahar (Ed.), *Teaching of science and technology* (pp. 249–276). Ankara, Turkey: Pegem.

Baker, M. (2009). Argumentative interactions and the social construction of knowledge. In N. M. Mirza and A.-N. Perret-Clermont (Eds.), *Argumentation and learning, theoretical foundations and practices* (pp. 127–144). New York, NY: Springer.

Balcı, A. (2004). *Research in social science: Methods, techniques and principles* (4th ed.). Ankara, Turkey: Pegem.

Barman, C. (1989). Making it work. *Science Scope, 12*(5), 28–31.

Burke, K., A., Hand., P., Poack., J., and Greenbowe, T., (2005). Using the science writing heuristic. *Journal of College Science Teaching*, 35 (1), 36-41.

Çakıroğlu, A. (2007). Metacognition. *Journal of Turkish Social Research*, *11*(2), 21–27.

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

Çepni, S. (2005). *Introduction to research and project work* (7th ed.). Trabzon, Turkey: Pegem.

Çepni, S., & Keleş, E. (2006). Turkish students' conceptions about the simple electric circuits. *International Journal of Science and Mathematics Education*, *4*, 269–291.

Choi, A., Notebaert, A., Diaz, J., & Hand, B. (2007). *Examining structure of science argument generated from the science writing heuristic approach across year 7 and 10 students.* Paper presented at International Conference of European Science Education Research Association, Malmö, Sweden.

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

Davies, P. (2002). Using students reflective self- assessment for awarding degree classifications. *Innovations in education and teaching ınternational*, 39 (4), 307-319.

Dawson, V., & Venville, G. (2010).Teaching strategies for developing students' argumentation skills about socio scientific issues in high school genetics. *Science Education*, *40*, 133–148.

Driver, R., Newton, P., and Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*, 287–312.

Eryaman, M. Y., Ozdilek, S. Y., Okur, E., Cetinkaya, Z. & Uygun, S. (2010). A participatory action research study of nature education in nature: Towards community-based eco-pedagogy. *International Journal of Progressive Education*, *6*(3), 53-70.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*, 906–911.

Grimber, B., I. (2008). *Promoting* high-order thinking through the use of the science writing heuristic. In B. Hand (Ed.), Science *Inquiry, Argument and Language* (pp. 87-98). Rotterdam: Sense Publisher.

Günel, M., Akkuş, R., Hohenshell, H., & Hand, B. (2004). *Improving student performance on higher order cognitive questions through the use of the science writing heuristic.* Paper presented at the National Association for Research in Science Teaching, Vancouver, BC, Canada.

Günel, M., Hohenshell, L., & Hand, B. (2006). *The impact of students' self-evaluations of the science writing heuristic: Closing the achievement gap.* Paper presented at the National Association for Research in Science Teaching, San Francisco, CA.

Günel, M., Kabataş Memiş, E., & Büyükkasap, E. (2010). Effects of the science writing heuristic approach on primary school students' science achievement and attitude toward science course, *Education & Science*, *35*(155), 49–62.

Hand B. M., (2008). Introducing the science writing heuristic approach. B. M., Hand (Ed.). *Science Inquiry*, Argument and Language.(pp.1-11).Sense Publisher.

Hand, B., & Keys, C. (1999). Inquiry investigation: A new approach to laboratory reports. *The Science Teacher*, *66*, 27–29.

Hand, B., Norton-Meier, L., Jay, S., & Bintz, J. (2009). *Negotiating science: The critical role of argument in student inquiry, grades 5-10*. Portsmouth, NH: Heinemann.

Hand, B., Wallace, C., & Yang, E. (2004). Using the science writing heuristic to enhance learning outcomes from laboratory activities in seventh grade science: Quantitative and qualitative aspects. *International Journal of Science Education, 26*, 131–149.

Hewsen, P. W., Beeth, M. E., & Thorley, N. R. (1998). Teaching for conceptual change. *International Handbook of Science Education*, *1*, 199–218.

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*, 28–54.

Hohenshell, L. M. (2008). Secondary students' perceptions of the SWH approach to nonconventional writing: Features that support learning of biology concepts and elements of scientific

argumentation. In B. Hand (Ed.), *Science inquiry, argument and language: A case for the science writing heuristic* (pp. 99–110). Rotterdam, The Netherlands: Sense Publishers.

Hohenshell, M. L. & Hand, B., (2006). Writing-to-learn strategies in secondary school cell biology: A mixed method study. *International Journal of Science Education.* 28(2), 261-289.

Hsieh, J. K. (2005). *Promoting students' ability and disposition toward critical thinking through using a science writing heuristic in elementary science.* Paper presented at the International Conference of European Science Education Research Association, Barcelona, Spain.

Kabataş, E., Günel, M., Büyükkasap, E., Uzoğlu, M., & Hand, B. (2008). *Effect of the implementation of the science writing heuristic on students' understanding of electricity unit in sixth grade setting in Turkey.* Paper presented at the conference of the National Association for Research in Science Teaching, Baltimore, MD.

Kabataş Memiş, E., Günel, M., & Büyükkasap, E. (2009). Isı ünitesinin ilköğretim 6.Sınıf seviyesi öğretiminde yaparak yazarak bilim öğrenimi-YYBÖ yaklaşımının kullanılmasının öğrenci akademik başarı ve tutumuna etkisi [English translation]. Paper presented at the 18th National Congress of Educational Sciences, Izmir, Turkey.

Karasar, N. (2004). *Scientific research method.* Ankara, Turkey: Nobel.

Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. *International Journal of Science Education*, *20*, 849–871.

Keys, C. W., Hand, B., Prain, V., & Collins, S. (1999). Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science. *Journal of Research in Science Teaching, 36*(10), 1065–1084.

Kim, H., & Song, J. (2006).The features of peer argumentation in middle school students' scientific inquiry. *Research in Science Education*, *36*(3), 211–233.

Kıngır, S., Geban, Ö., & Günel, M. (2012). How does the science writing heuristic approach affect students' performances of different academic achievement levels? A case for high school chemistry. *Chemistry Education Research and Practice*, *13*, 428–436.

Kıngır, S., Geban, Ö., & Günel, M. (2013). Using the science writing heuristic approach to enhance student understanding in chemical change and mixture. *Res. Sci. Educ., 43*, 1645–1663.

Kolb, D. A. (1984). *Experiential learning experience as a source of learning and development.* Englewood Cliffs, NJ: Prentice Hall.

Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, *94*(5), 810–824.

Ministry of National Education. (2006). *Elementary science and technology curriculum (6, 7 and 8 classes).* Ankara, Turkey: Author.

Mohammed, E. G. (2007). *Using the science writing heuristic approach as a tool for assessing and promoting students' conceptual understanding and perceptions in the general chemistry laboratory* (Unpublished doctoral dissertation). Iowa State University, Ames.

Nam, J., Choi, A., & Hand, B. (2010). Implementation of the science writing heuristic (SWH) approach in 8th grade science classrooms. *Int. J. Sci. Math. Educ., 9*(5), 1111–1133.

National Research Council (NRC). (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academies Press.

Newton, P., Driver, R., & Osborne, J. (1999).The place of argumentation in the pedagogy of school science. *International Journal of Science Education, 21*(5), 553–576.

Norton-Meier, L. (2008). Creating border convergence between science and language: A case for the science writing heuristic. In B. Hand (Ed.), *Science inquiry, argument and language* (pp. 13–24). Rotterdam, The Netherlands: Sense Publishers.

Norton-Meier, L., Nelson, S., Hockenberry, L. & Wise, K. (2008).Transforming pedagogy: Embedding language practices in elementary classrooms.In Hand, B. Science Inquiry, Argument and Language: The case for the Science Writing Heuristic (SWH). Rotterdam, the Netherlands:Sense Publishers.

Nussbaum, E. M., Sinatra, G. M., & Poliquin, A. M. (2008). The role of epistemic beliefs and scientific argumentation in science learning. *International Journal of Science Education, 30,* 1977–1999.

Olina, Z., & Sullivan, H. J. (2004). Student self-evaluation, teacher evaluation, and learner performance. *Educational Technology Research and Development*, 52(3), 5-22.

Omar, S. (2004). *Inservice teachers' implementation of the science writing heuristic as a tool for professional growth* (Unpublished doctoral dissertation). Iowa State University, Ames.

Ozogul, G., Olina, Z. & Sullivan, H. (2008). Teacher, self and peer evaluation of lesson plans written by preservice teacher. *Education Tech Research Dev*, 56, 181-201.

Prain, V. (2007). The role of language in science learning and literacy. In C.S. Wallece, B. Hand & V. Prain (Eds.), *Writing and learning in the science classroom.* (pp. 33-46). Science & Techonology Education Library, Springer.

Ramsey, J. (1993). Developing conceptual storylines with the learning cycle. *Journal of Elementary Science Education, 5*(2), 1–20.

Rivard, L.P. (2004). Are language- based activities in science effective for all students, including low achievers?. *Science Education*, 88, 420-442.

Schroeder, J. D. (2008). The laboratory-lecture correlation: From the science writing heuristic to the traditional organic chemistry laboratory. In B. Hand (Ed.), *Science Inquiry, Argument and Language* (pp. 165–176). Rotterdam, The Netherlands: Sense Publishers.

Schwarz, B. B.(2009). Argumentation and learning. In N. M. Mirza and A.-N. Perret-Clermont (Eds.), *Argumentation and learning, theoretical foundations and practices* (pp. 91–126). New York, NY: Springer.

Sebba, J., Deakin Crick, R., Yu, G., Lawson, H. & Harlen, W.(2008). Impact of self and peer assessment on students in secondary schools. *Research Brief.*

Sencar, S., & Eryılmaz, A. (2004). Factors mediating the effect of gender on ninth-grade Turkish students' misconceptions concerning electric circuits. *Journal of Research in Science Teaching*, *41*(6), 603–616.

Sungur, S., & Senler, B. (2009). An analysis of Turkish high school students' metacognition and motivation. *Educational Research and Evaluation*, *15*(1), 45–62.

Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment task: Towars transparency for students and for tutors. *Assessmnet & Evaluation in Higher Education*, 26 (6).

Tatar, N. (2006). The effect of inquiry-based learning approaches in the education of science in primary school on the science process skills, academic achivement and attitude. (Unpublished doctoral dissertation). Gazi University, Ankara.

van der Stel, M., & Veenman, M. V. J. (2010). Development of metacognitive skillfulness: A longitudinal study. *Learning and Individual Differences*, *20*, 220–224.

Wallece, C.S. (2007). Evidence from the literature for writing as a mode of science learning. In C.S. Wallece, B. Hand & V. Prain (Eds.), *Writing and learning in the science classroom.* (pp. 9-21). Science & Techonology Education Library, Springer.

Zohar, A., & Nemet, F.(2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, *39*(1), 35–62.