edam

Educational Sciences: Theory & Practice - 13(4) • 2217-2228

www.edam.com.tr/estp
DOI: 10.12738/estp.2013.4.1721

A Study on Development of an Instrument to Determine Turkish Kindergarten Students' Understandings of Scientific Concepts and Scientific Inquiry Processes

Erdal SENOCAK[a]
Gaziosmanpasa University

Ala SAMARAPUNGAVAN[b]
Purdue University

Pınar AKSOY[c]
Gaziosmanpasa University

Cemal TOSUN[d]
Bartin University

Abstract

The aim of this study was to develop a valid and reliable instrument to measure Turkish kindergarten students' understandings of some science concepts and scientific inquiry processes which are grounded in the Turkish Preschool Curriculum. The sample of the study was 371 kindergarten students, 12 Subject Area Experts (SAE), and 7 Turkish Language Experts (TLE). Six stages were followed in the development process of the instrument: (i) item formulation, (ii) content validity, (iii) language validity, (iv) item difficulty and discrimination index, (v) factor analysis, and (vi) reliability. First, an item pool was constituted with 42 items. Second, SAEs and TLEs rated these items in respect to the degree to which they reflected the content and their understandability and grammar accuracy in Turkish. Third, all items were implemented kindergarten students, and 26 items were eliminated according to their item difficulty and discrimination index values. Last, factor analysis and reliability were studied by means of the data belonging to the rest of items. Results revealed that the instrument with 16 items had two factor structure and acceptable reliability.

Key Words

Instrument Development, Kindergarten Students, Preschool Education, Science Concepts, Scientific Inquiry Processes.

Early childhood education has a long history. There are many scholars who have studied childhood learning such as Martin Luther, John Comenius, John Dewey, Maria Montessori, and Jean Piaget (Brewer, 1998). They proposed diverse ideas to explain how children learn. Their ideas made important contributions to contemporary early childhood education programs. Friedrich Froebel is known as a pioneer of the kindergarten movement (Bryant & Clifford, 1992). Froebel believed that young children should be placed under the influence of a qualified program to foster their inherent

a **Erdal SENOCAK, Ph.D.,** is currently an associate professor of Science Education. His research areas are chemistry education, problem-based learning, science education in the early childhood years, and public understandings of nanoscience. *Correspondence:* Assoc. Prof. Erdal SENOCAK, Department of Science Education, College of Education, Gaziosmanpasa University, Tokat 60140, Turkey. Email: e_senocak_2000@yahoo.com.

b Ala SAMARAPUNGAVAN, Ph.D., is a professor of Educational Studies. Contact: Purdue University, Department of Educational Studies, College of Education, West Lafayette, IN 47907, US. Email: ala@purdue.edu.

c Pınar AKSOY is a research assistant of Department of Early Childhood Education. Contact: Gaziosmanpasa University, Department of Early Childhood Education, College of Education, Tokat 60140, Turkey. Email: aksoypnr@gmail.com

d Cemal TOSUN, Ph.D., is an assistant professor of Science Education. Contact: Bartin University, Department of Science Education, College of Education, Bartin, Turkey. Email: cemaltosun22@gmail.com.

curiosity by means of self-directed activities (Bryant & Clifford, 1992; Olsen & Zigler, 1989). He developed instructional activities to teach children who were between 3 and 7 years old in his school, which he named Kindergarten: *"children's garden"* (Bryant & Clifford, 1992). This movement gained wide acceptance and spread to other countries in later years (Bryant & Clifford, 1992; Shapiro, 1983). Today, kindergarten is used as a word to describe the education process at the beginning of primary school (generally at age 5), and it is compulsory in some countries. Nowadays, kindergarten classes, as a part of preschool education, try to provide some opportunities for children to improve their social, emotional, and cognitive development. Also, some countries' national preschool programs endeavor to encourage young children to learn the basic skills and knowledge of diverse learning areas such as science, math, and health (National Research Council, 1996; Ojala & Talts, 2007).

In Turkey, preschool education is an optional educational process for children from 36 to 60 months of age (Ministry of National Education, 2012). A substantial number of the population of Turkey is in the age group of 0-14 (Turkish Statistical Institute, 2010). Therefore, the Ministry of National Education has undertaken important efforts to improve the quality of Turkish pre- and primary schools in recent years. Although there is a growing interest in preschool education in Turkey, no strong agreement exists about science education in early childhood. The reason for this conflict may be the question: *"Is it possible to cope with scientific concepts in the early years?"* Particularly preschool teachers resist exposing children to science in the preschool years (Ayvacı, Devecioğlu, & Yiğit, 2002) and allocate little time for science activities in class (Akman, Ustun, & Guler, 2003). In fact, this question has been argued in the literature on the subject (Eshach & Fried, 2005), and there are some answers to the question by researchers. For example, Tu (2006, p. 247) stated: "*Science education has been strongly advocated in the primary school curriculum for its importance to young children.*" Mantzicopoulos, Samarapungavan, and Patrick (2009, p. 364) said, "*Our results strengthen the claim that science instruction should begin by the early school years…*" Smith (2001) pointed out, "*There are many strategies and techniques which can be used to enhance early childhood science learning.*" Additionally, there are many other studies suggesting that science education should begin in the early years (Eshach, 2011; Eshach & Fried, 2005; French, 2004; Watters, Diezmann, Grieshaber, & Davis, 2000).

The studies on the early childhood education in Turkey are generally focused on in-service and pre-service teachers' perceptions (Bedel, 2008; Durmusoglu, 2008; Erden & Sonmez, 2011; Kabadayi, 2010; Secer, 2010) or preschool curriculum (Atalay-Turhan, Koc, Isikal, & Isikal, 2009). There are few studies focusing on science learning at the kindergarten level (Akman et al., 2003; Ayvaci, 2010; Menekse, Clark, Ozdemir, D'angelo, & Scheligh, 2009; Sackes, Flevares, & Trundle, 2009). And there is limited number of studies on developing an instrument to measure Turkish kindergarten students' understanding of science concepts or scientific inquiry processes. Therefore, in this study, the aim was to develop an instrument to determine Turkish kindergarten students' (60-72 months) understanding of both science concepts grounded in the Turkish preschool curriculum and scientific inquiry processes. Another aim was to produce a practical tool so that a program planner can assess the output of the Turkish preschool program, or a researcher studying science in preschool education can gather empirical data through this tool.

Assessment in the Childhood Years

Assessment and evaluation start at the moment when a child is born. Thus, they play an important role in our lives. During the first minutes of life, babies are assessed such as heart rate per minute, weight, length, or other things. If a baby receives good scores on these assessments, she/he is evaluated as in good condition. This process occurs at nearly every stage of human life.

In recent years, there has been a debate on what types of assessment and evaluation are appropriate for the early years of life (Wortham, 2008). Researchers are especially concerned about the misuse of testing and results from the outcomes of the measurements. This is particularly important in early childhood education because it is important to understand the principles underlying assessment and evaluation in the early years to be able to make appropriate decisions about children.

The interest in early childhood education has rapidly increased during the last decade. Growing concern over education for the early childhood years has resulted in some outstanding education programs and new measurement tools to assess children's progress and the effectiveness of these programs. The Head Start on Science and Communication Program (Klein, Hammrich, Bloom, & Ragins, 2000), the

ScienceStart! Program (French, 2004), the Preschool Pathways to Science (Gelman & Brenneman, 2004), and the Scientific Literacy project (Samarapungavan, Mantzicopoulos, & Patrick, 2008) are outstanding examples of these programs. Furthermore, a variety of instruments have been used to assess children's progress and the effectiveness of these programs. Examples of these instruments are standardized tests of cognitive achievement such as the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997), the Woodcock Johnson III (Woodcock, McGrew, & Mather, 2001), and researcher developed instruments such as the Science Learning Assessment (Samarapungavan et al., 2008; Samarapungavan, Mantzicopoulos, Patrick, & French, 2009).

Measuring the cognitive achievements of young children is more difficult than older ones because young children experience vast variations in the different personal, developmental and environmental factors affecting their behaviors (Gullo, 2005). To address these difficulties, in 1994, the U.S. Department of Education organized a commission to meet and constitute general principles guiding assessment and evaluation practices for young children. The commission made some decisions on early childhood assessment (Shepard, Kagan, & Wurtz, 1998). A brief summary of these principles: the assessments should result in benefits for children, be used only in accordance with their specific designed purposes, be cautious about the limitations of young age, be age appropriate, be linguistically appropriate, and consider parents as a valued information source. Moreover, early childhood educators were advised to take into account these principles when assessing children.

As previously discussed in this study, an instrument will be developed to assess the extent to which Turkish preschool children have achieved some objectives of the Turkish Preschool Curriculum. Crucially, we would like to emphasize that the aim of this study is not to develop an achievement test to declare children as successful or not. The main aim is to develop a valid and reliable instrument to derive information from children about the effectiveness of the Turkish Preschool Program and child progress in respect to certain science concepts and scientific inquiry processes. Also, this instrument may help program developers and researchers gather useful information about the program to understand what children learn, what is working well, or what types of enhancements may be needed to improve the effectiveness of the program.

What are Inquiry and Scientific Inquiry Processes?

There are some studies on the limitations of empowering children's understandings of scientific inquiry processes (Metz, 2004; Samarapungavan et al., 2008; Zhang, Parker, Eberhardt, & Passalacqua, 2011). Many indicators reveal that children show interest in inquiry-based programs, and that they acquire the basic skills of the scientific process. Metz (2004) reported that parents whose children participated in an inquiry-based instruction stated that their children were interested in becoming a scientist. Eshach (2011, p. 442) said, "*Children have sufficient cognitive capabilities to engage in scientific inquiry.*" Samarapungavan et al. (2008) administered a science learning program that was grounded in scientific inquiry and literacy activities. They collected data from both invention and comparison group kindergarten children by using different data collecting tools. Results showed that intervention group children had better scores than comparison group children with respect to acquiring the key aspects of scientific inquiry processes and science concepts taught in this program.

Inquiry is a process that involves wondering, asking questions, collecting data, and answering questions in order to learn what is taking place around us. Inquiry is provoked by a sense of curiosity and is the powerful sense driving human beings to discover something of the world around them. *Curiosity is the desire to learn or know about something* (Harlan & Rivkin, 2008, p. 4). Barell (2008) stated that children are born with a sense of curiosity, and students at an early age can ask interesting and challenging questions to solve problems. He also suggested some effective activities to foster and keep children's curiosity such as keeping wonder journals, developing problematic scenarios, and garnering parental support. Furthermore, Rankin (2011) expressed that students' curiosity can be sustained with an effective pedagogical structure which can provide inquiry experiences.

Inquiry requires some skills such as observation, questioning, measurement, classification, and prediction. These are known as the basic science processes (Bentley, Ebert, & Ebert, 2007; Martin, Sexton, Franklin, Gerlovich, & McElroy, 2009), and these are appropriate for kindergarten (Martin et al., 2009). In fact, we use these processes in our daily lives when trying to solve problems, even if we generally are not aware of using them. Nowadays, educators endeavor to develop new ways to keep children's scientific curiosity alive or take them much further. "A Nation at Risk Report" is one of the most obvious examples of these endeavors. In 1981, a commission was constituted to present a report on the quality of education in the United States. Almost two years later, the commission released a report named, "A Nation at Risk: The Imperative for Educational Reform". They examined the quality of teaching and learning in schools. In this report, there were some emphasizes on science education. For example, it was proposed to revise the science courses and recommended using the methods of scientific inquiry and reasoning to raise citizens who are literate in science and technology (A Nation at Risk, 1983). In the early 2000s, a strong effort was started in Turkey to revise and develop new curriculum to parallel the new trends in education. With this effort, the main idea in the curriculum has changed from the content-centered to the student-centered (Bulut, 2007), and it aimed to raise scientifically literate citizens (Bahar, 2006). Therefore, we included a second dimension in the instrument to reveal some clues to understanding the extent to which Turkish preschool students are aware of scientific inquiry processes.

Table 1.
Main Areas and Concepts Derived from the Indicators in the Program

Main Areas	Concepts	Indicators
Life science	Living things	Understand living and non-living concepts.
		Understand living things have life cycle (they are born, develop into adult, and die).
		Understand the parts of the plants (such as seed, stem, root, leaf and flower)
Physical Science	Properties of the Objects, Heat and Temperature, Sound	Understand that there is not always a linear relationship between sizes and masses of the objects.
		Understand what the objects are made of.
		Understand the objects magnets interact with.
		Understand the concepts of hot and cold.
		Understand people hear the sound with their ears.
		Understand the properties of sound such as high and low volume.
Earth/ Space Science	Day and Night	Understand the concepts of daytime and nighttime.

Also, the indicators in the Turkish Preschool Program that referred to scientific inquiry processes were classified in accordance with the Scientific Inquiry Subtest developed by Samarapungavan et al. (2009). Although there was no explicit stress on scientific inquiry processes in the program, it was hardly detected some indicators which correspond with the Scientific Inquiry Subtest developed by Samarapungavan et al. (2009). This integration is presented in Table 2.

After analyzing the Turkish Preschool Program and determining some indicators, the research team started to write items for each subtest of the instrument. In this process, researchers considered some directions in accordance with scale development and young children's learning literature, concluding that (DeVellis, 2003; Gullo, 2005; Puckett & Black, 2008): *i*) the content of each item should reflect the construct measured, *ii*) each item should include short scenarios as much as possible to avoid breaking children's attention, *iii*) each item should be supported with a picture to capture children's attention, *iv*) each item should include clear sentences, *v*) at least three items should be written for each indicator for the first subtest of the instrument to generate a rich item pool.

Method

Sample

There were three groups of participants in this study. Twelve Subject Area Experts (SAE) who were science educators constituted the first participant group to rate the items according to the degree to which the items reflected the content. Seven Turkish language experts (TLE) were the second participant group who rated the understandability of the items. The last participant group of the study comprised 371 Turkish kindergarten students. These students, who were attending 13 different urban public preschools located in the north of Turkey, took all items during the end of the school year in May and June of 2012.

Instrument Development Steps

The development process of the instrument constituted six stages.

Item Formulation: Before the item writing process, researchers analyzed the Turkish Preschool Program to determine the target understandings that referred to science concepts and scientific inquiry processes. The results of the analysis showed that the program did not explicitly emphasize both science concepts and scientific inquiry processes. Nevertheless, there were a set of indicators addressing some science concepts, which are presented in Table1.

Content Validity: In this study, the items were reviewed by a group of subject area experts (SAEs) who were knowledgeable in the area of science learning during the early years. They rated the items by assigning a number between 0 and 10, in which 0 was "not reflect the content at all" and 10 was "reflects the content." The content validity was calculated using Content Validity Ration (CVR) based on Lawshe's formula (Lawshe, 1975).

Table 2.
The Integration of the Indicators to Scientific Inquiry Processes

Targeted understandings in the Scientific Inquiry Processes	Indicators in the Turkish Preschool Program
Understand science as a process of inquiry is based on asking questions and making predictions about the natural world.	Tell the possible reasons of an event. Tell the possible results of an event.
Understand the empirical basis of science: Scientific ideas are evaluated by their correspondence or fit to empirical evidence.	State the problem. Make some suggestions for the solution of the problem. Choose the most suitable suggestions for the solution of the problem. Test the chosen suggestions. Make a decision the most suitable suggestion for the solution of the problem. Explain the reasons of the decision.
Understand simple tools used to gather, record, analyze, and share data.	Predict the result of the measurement. Measure using non-standard units. Compare the measurement results with predicted results. Explain the functions of the tools that measure time. Use the concepts with regard to time in accordance with their function.

Language Validity: There are some factors that might influence the performance of the test taker or the content validity of a test (Kaplan & Saccuzzo, 1997), such as understandability of an item. Therefore, Turkish language experts (TLEs) rated the items in the item pool according to the accurate usage of the Turkish language and kindergarten children's understanding. They rated each item with a number between 0 and 10, in which 0 was when "a kindergarten child can't understand at all" and 10 was when "a kindergarten child is able to understand".

Item Difficulty and Discrimination: Item difficulty and discrimination indexes were calculated for each item in the item pool. The aim of determining item difficulties was to show how hard or easy the items of the instrument. We also calculated discrimination indexes of the items because we desired to reveal how well an item serves to discriminate between students higher and lower levels of knowledge. Moreover, Devellis (2003) stressed that item difficulty and discrimination are two important item characterization indexes showing an item's performance.

Factor Analysis: In the study, confirmatory factor analysis (CFA) was used to evaluate construct validity of the instrument. CFA is described as a theory testing procedure opposed to the CFA (Roberts, 1999), and it is used to test to what extent the observed variable(s) fits predefined variable(s). CFA was used to test whether the observed variables were consistent with the theoretical variables. The theoretical variables of this study were formed from a two factor model: understanding of the basic science concepts taught in the Turkish Preschool Program and understanding of the scientific inquiry processes.

Reliability: Reliability is another main topic in psychological measurement, and it represents the extent of instrument consistency or stability (DeVellis, 2003; Gullo, 2005). There are different ways to estimate the reliability of a test such as the test-retest method, the parallel forms method, and internal-consistency methods. Internal-consistency methods require only one administration of a test (Tekindal, 2008). The split-half method, the Kudher-Richardson ($KR_{20 \text{ and } 21}$) formula, and the Spearman-Brown formula are used to calculate the internal-consistency coefficient. Because we administered the items only once, internal-consistency methods are appropriate ways to estimate reliability for this study. Also, our instrument included dichotomous items with multiple response options which can be classified as right/wrong. Therefore, KR_{20} and Cronbach's alpha can be preferred to test the reliability of the instrument. However, Gronlund and Waugh (2009) emphasized that the KR-20 is useful with a homogeneous test but can be misleading when it is used with a test that includes heterogeneous content. Although our instrument is a binary one, it has two contents, science concepts and scientific inquiry processes. Therefore, Cronbach's alpha was used to determine the reliability coefficient of the instrument.

Results

Item Formulation

The item formulation process is composed of two parts: writing items for the science concepts subtest (SCS) and the scientific inquiry subtest (SIS). Items were written based on the indicators (see Table 1) derived from the Turkish Preschool Program and the target understandings that was referred to by Samarapungavan et al. 2009 (see Table 2). Three items with three choices were written and pictured for each indicator in the SCS. That is, a total of

30 items were written for the first sub-scale. One example of the items is shown below.

> These pictures show the stages of human life (show the pictures which has following order empty slot, a baby, a teenage boy and a young girl). Now look at these pictures (show pictures) and tell me which of these pictures should go up here (point to missing picture) to complete the stages of human life?

Another subtest of the instrument aimed to reveal children's understanding of scientific inquiry processes. As seen in Table 2, we tried to determine children's understanding based on 3 targeted understandings in scientific inquiry. In this scope, we wrote a total of 12 items with three choices: 4 items for the first targeted understanding, 3 items for the second targeted understanding, and 5 items for the last targeted understanding. One example of the items is shown below.

> (Show pictures) In the first picture, Ali's mother is picking up the needles dropped on the ground through a magnet. In the second picture, Ali is trying to pick up dried beans dropped on the ground through the same magnet, but it doesn't work. Here are three boys (show pictures). I will tell you what each boy is saying about. Now think about what Ali and her mother did and tell me which one of these children talked about their work?

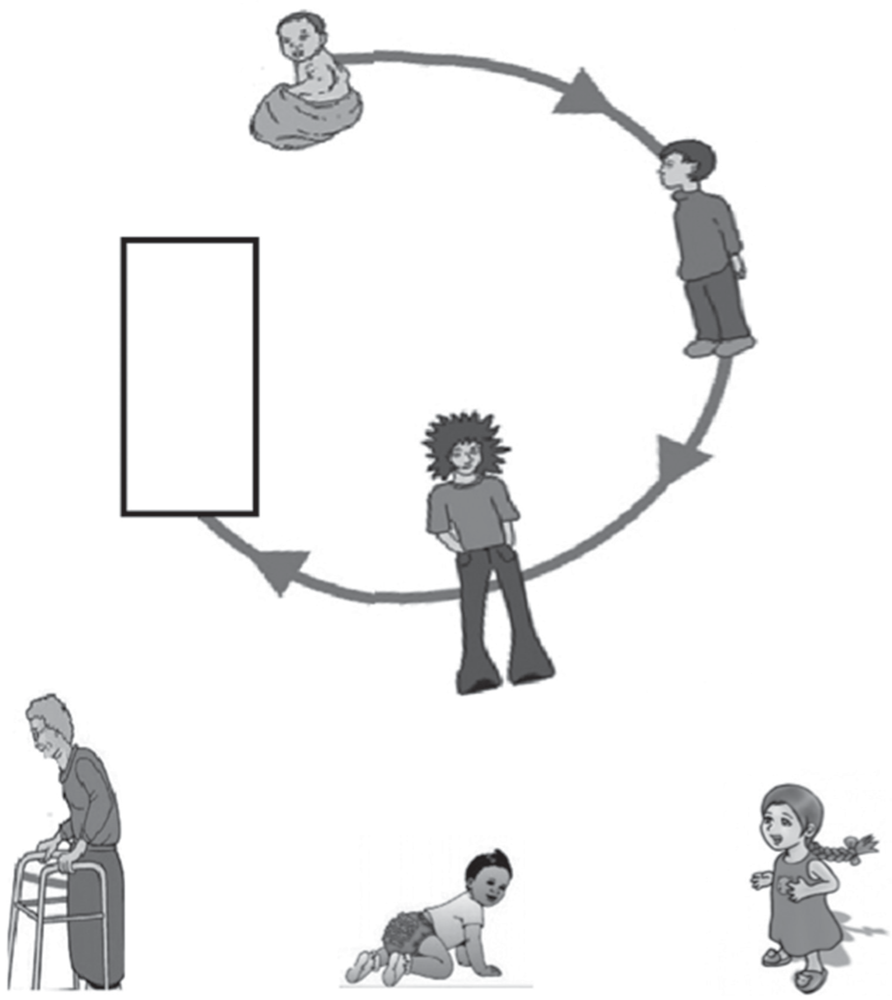

Figure 1.
Schematic Presentation of a SCS Item

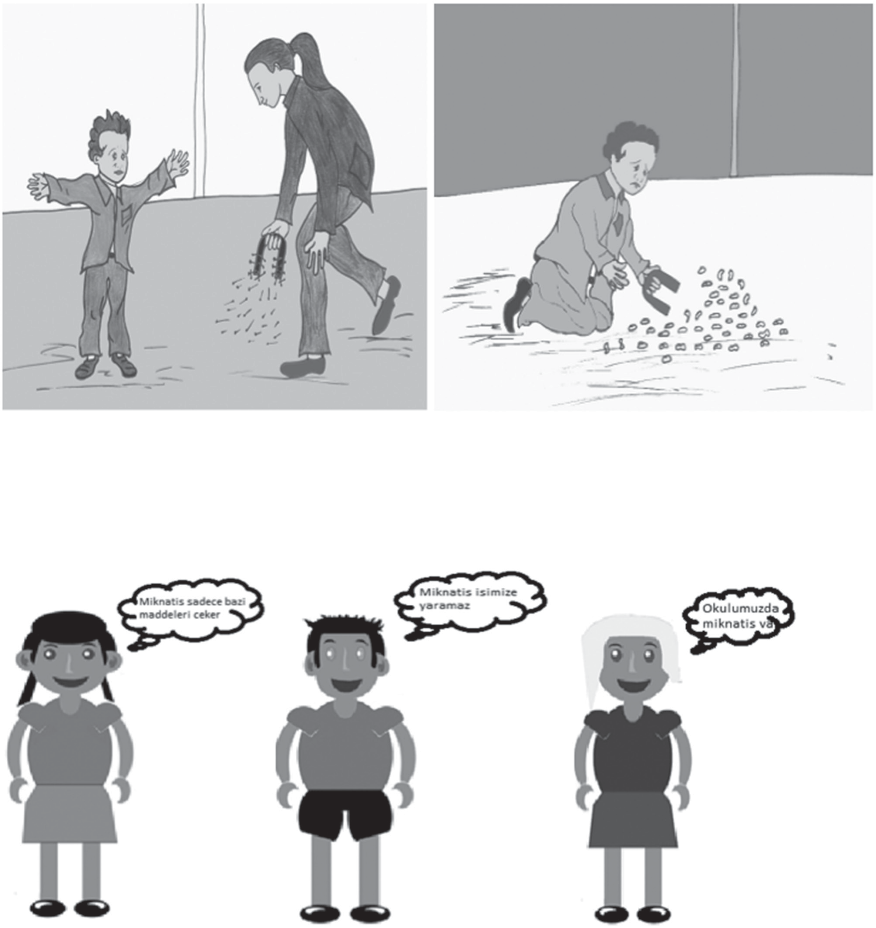

Figure 2.
Schematic Presentation of a SIS Item

Content Validity

To determine the content validation, 12 SAEs rated each item in respect to the degree to which it reflected the content of the related indicator. Also, they were asked to check the quality of each item and suggest necessary item revisions. After this procedure, some items were revised in accordance with the SAEs' suggestions, and data derived from the SAEs' rates was analyzed (see Table 3). As seen in Table 3, all items were rated with higher scores. Ratings ranged from 7.33 to 9.83 (M=8.80; SD=1.58). These findings showed that there is strong agreement among SAEs on the items' power to measure the content of the instrument.

Furthermore, the Lawshe (1975) content validity ratios (CVR) was used to measure the content validation. These ratios use a technique developed by Lawshe to gauge the content validity of items on an empirical measurement. In this method, a particular item is rated as "essential," "useful, but not essential," or "not necessary" by experts who are knowledgeable about the content of the item. Lawshe developed a formula to calculate the content validity ratio (CVR) which helps determine whether an item is retained in an instrument. This formula is $CVR = [n_e - (N / 2)] / (N / 2)$, where n_e = number of SMEs who rated the item as "essential", and N = total number of SMEs. CVR can take values between -1.0 and 1.0, where CVR=0.00 and positive values show that at least half the SMEs rated the item as essential. Lawshe has determined the minimum CVR values, which vary according to the number of SAEs. In this study, 12 SAEs rated

Table 3.
Understandability Rates, Content Validity Rates and CVR Values of the Items

Item No.	Understandability Rates		Content Validity Rates		CVR Values	Item No.	Understandability Rates		Content Validity Rates		CVR Values
	M	SD	M	SD			M	SD	M	SD	
Item 01	9.43	1.13	9.25	1.13	1.00	Item 22	9.86	0.37	9.75	0.62	1.00
Item 02	10.00	0.00	8.67	1.61	0.83	Item 23	9.57	0.78	8.75	1.42	1.00
Item 03	9.29	1.89	9.00	1.27	1.00	Item 24	9.57	1.13	9.42	1.24	0.83
Item 04	8.57	2.54	8.42	2.28	0.67	Item 25	9.71	0.48	7.50	3.20	0.67
Item 05	9.29	1.25	8.92	1.16	0.83	Item 26	10.00	0.00	8.83	1.11	1.00
Item 06	7.86	2.85	7.83	2.29	0.67	Item 27	9.43	1.51	7.33	2.34	0.83
Item 07	9.71	0.48	9.08	1.50	0.83	Item 28	9.29	0.95	9.08	0.99	1.00
Item 08	8.29	2.87	8.67	1.07	1.00	Item 29	9.71	0.75	9.75	0.62	1.00
Item 09	9.71	0.75	7.83	1.58	0.67	Item 30	9.71	0.75	9.00	1.20	1.00
Item 10	9.57	0.78	8.00	3.04	0.67	Item 31	9.86	0.37	9.25	1.13	1.00
Item 11	9.71	0.48	7.58	3.26	0.67	Item 32	9.43	1.13	9.33	1.30	0.83
Item 12	9.86	0.37	7.67	2.99	0.67	Item 33	9.14	1.46	7.50	2.64	0.67
Item 13	9.29	1.49	9.58	0.79	1.00	Item 34	9.14	1.86	8.67	1.50	0.83
Item 14	9.43	1.13	9.42	1.08	1.00	Item 35	9.29	1.25	8.08	2.53	0.67
Item 15	9.29	1.11	8.50	2.87	0.83	Item 36	9.57	0.78	8.50	1.56	0.67
Item 16	9.43	1.13	9.25	1.28	0.83	Item 37	9.71	0.48	8.58	1.16	1.00
Item 17	9.57	0.78	9.67	0.49	1.00	Item 38	10.00	0.00	9.58	1.00	1.00
Item 18	9.71	0.78	9.17	1.03	1.00	Item 39	10.00	0.00	9.33	1.61	0.83
Item 19	9.86	0.37	9.33	1.07	1.00	Item 40	10.00	0.00	9.75	0.86	1.00
Item 20	9.00	1.18	8.75	1.48	0.83	Item 41	9.71	0.48	8.58	2.87	0.83
Item 21	9.86	0.37	9.83	0.39	1.00	Item 42	9.86	0.37	8.67	1.77	0.83

Note: *The first 30 items belong to the BSC subtest and the rest of the items belong to the SIPS subtest*

all of the items; in this way, an item requires .56 or more CVS values to be accepted as "essential" (Lawshe, 1975). In this study, the items rated 7 and over were deemed "essential." As seen in Table 3, all of the items reached higher CVR values than 0.56. Namely, SMEs agreed that all items successfully measured the content.

Language Validity

In the third step of the instrument development process, seven TLEs rated each item in terms of the items' grammatical accuracy and understandability by an average Turkish kindergarten child. The TLEs rated the items and made some suggestions on some items. Then, these items were revised in accordance with the experts' suggestions. Data derived from the TLEs were also analyzed (see Table 3). Ratings ranged from 7.58 to 10.00 (M=9.50, SD=0.91). These findings revealed that there was an important consensus among TLEs on the items' understandability. According to these findings, we can say that all items are understandable for kindergarten children, and that they have language accuracy.

Item Difficulty and Discrimination

After language validity, all items in the pool were administered to the kindergarten students, and then the findings were analyzed. Analysis results showed that all items had different ranges of difficulty and discrimination values (see Table 4). As expressed before, at least three items were written for targeted understanding. In this stage of the study, the items which ideally reflected optimized item difficulty and discrimination indexes were chosen for each targeted understanding. One item for each targeted understanding in the first sub-scale of the instrument and two items for each targeted understanding in the second sub-scale of the instrument were determined according to item difficulty and discrimination indexes values (see Table 4). According to Walsh and Betz (2000), the items of an instrument should have difficulty index values between 0.1 and 0.90. Also, Kaplan and Saccuzzo (1997) argued that the optimum difficulty level for four-choice items is about .62. According to this assumption, the optimum difficulty level for three-choice items is about .66. As for discrimination indexes of the items, if an item has a discrimination value between .0 and 1.0, it means that this item distinguishes between high achieving examinees and low achieving examinees (Kaplan & Saccuzzo, 1997). In addition, item discrimination is optimized when item difficulty is close to .50, where the values of .20 and above are considered to be desirable.

As seen in Table 4, the items which have appropriate item difficulty and discrimination values were chosen for the instrument. Namely, the instrument constituted 16 items in total, 10 of which (items 3, 4, 9, 11, 15, 17, 20, 3, 27, and 28) belong to the SCS and the rest (items 3, 34, 36, 37, 38, and 42) belongs to the SIS. While the item difficulty values of the items ranged from .40 to .85, the discrimination values for them ranked from .27 to .57. The instrument included both easy and difficult items. The easiest item was item 3, which was answered correctly by 85% of the examinees, and the most difficult item was item 32, which was answered correctly by 40% of the examinees. The average difficulty and discrimination indexes through all items were .63 and .44, respectively.

Table 5.
The average difficulty and discrimination values of the subscales

Subtest	Difficulty	Discrimination
Basic science concepts	.62	.41
Scientific inquiry	.63	.48

Table 5 represents the average difficulty and discrimination values of the subtests of the instrument. As seen in Table 5, the average difficulty values of both the SCS and the SIS were so close to the optimum difficulty level (.66) for the three-choice item. In addition, the average discrimination values of both sub-scales were close to the optimum discrimination level (.50).

Factor Analysis

To test the latent structure of the hypothesized two-factor model of the instrument, CFA was conducted using LISREL, which is suited for a binary data set (Simsek, 2007). Chi-square/df ratio, the root mean square error of approximation (RMSEA), the goodness fit index (GFI), and the standardized root mean square residual (SRMR) were considered as goodness-of-fit indexes for this study. Also, factor loadings of all items and fit indexes for hypothetical two-factor model are summarized in Table 6.

As seen in Table 6, the standardized factor loadings are acceptable; higher than .30, except for item 27 which belongs to the BSC sub-scale. By convention, factor loading greater than or equal

Table 4.
Item Difficulty and Discrimination Values of the Items

Target Understanding	Item No	Difficulty	SD	Discrimination	Target Understanding	Item No	Difficulty	SD	Discrimination
Understand living and non-living concepts.	1	.86	.35	.27	Understand people hear the sound with their ears.	22	.90	.30	.22
	2	.86	.35	.23		**23*	*.46*	*.50*	*.45*
	**3*	*.85*	*.36*	*.31*		24	.83	.38	.32
Understand living things have life cycle.	**4*	*.61*	*.49*	*.43*	Understand the properties of sound.	25	.88	.33	.15
	5	.60	.49	.36		26	.94	.23	.02
	6	.50	.50	.14		**27*	*.42*	*.49*	*.27*
Understand the parts of the plants.	7	.46	.50	.40	Understand the concepts of daytime and nighttime.	**28*	*.65*	*.48*	*.48*
	8	.33	.47	.45		29	.63	.48	.39
	**9*	*.55*	*.50*	*.55*		30	.86	.35	.28
Understand the relationship between sizes and masses of the objects.	**10*	*.80*	*.40*	*.30*	Understand science as a process of inquiry.	31	.90	.30	.24
	11	.57	.50	.23		**32*	*.40*	*.49*	*.47*
	12	.85	.36	.26		33	.50	.50	.32
Understand what the objects are made of.	13	.70	.46	.34		**34*	*.60*	*.49*	*.49*
	14	.88	.32	.22	Understand the empirical basis of science.	35	.80	.40	.34
	**15*	*.71*	*.45*	*.42*		**36*	*.74*	*.62*	*.44*
Understand the objects magnets interact with.	16	.87	.34	.20		**37*	*.68*	*.47*	*.52*
	**17*	*.58*	*.49*	*.51*	Understand simple tools.	**38*	*.84*	*.37*	*.41*
	18	.84	.37	.39		39	.84	.37	.36
Understand the concepts of hot and cold.	19	.96	.20	.06		40	.89	.31	.17
	**20*	*.60*	*.49*	*.40*		41	.47	.50	.34
	21	.92	.27	.12		**42*	*.54*	*.50*	*.57*

* Represent the item which has appropriate item difficulty and discrimination values

to .30 are considered to meet the minimal level (Hair, Anderson & Black, 1995). Chi-square/df (358.41/103) is 3.48, less than 5. By convention, Chi-square/df should be between 2 and 5 for an acceptable fit index (Simsek, 2007). The RMSEA is .08, which is acceptable. By convention, the RMSEA should be between .05 and .08 for an acceptable fit index (Chan, Lee, Lee, Kubota, & Allen, 2007). The GFI is .90, equal to the conventional criterion of .90 or greater for a good fit index (Jöreskog & Sörbom, 2003). The SRMR is .06, also acceptable, by falling below the conventional cutoff criterion of .08. By convention, an SRMR value which is between .05 and .08 is an acceptable fit index (Hu & Bentler, 1999; Simsek, 2007). There is no general consensus among authorities on how to determine a certain cutoff criterion for each index (Simsek, 2007). For example, Chen, Curran, Bollen, Kirby, and Paxton (2008) have suggested that there should not be a universal cutoff criterion for the RMSEA fit index.

Table 6.
Confirmatory Factor Analysis Results

Item No	The basic science concepts subscale		The scientific inquiry subscale	
	Factor Loading	T-value	Factor Loading	T-value
3	.35	6.18		
4	.40	7.12		
9	.49	8.93		
10	.31	5.46		
15	.51	9.42		
17	.56	10.57		
20	.31	5.43		
23	.46	8.38		
27	.23	3.93		
28	.49	8.96		
32			.41	7.39
34			.52	9.64
36			.46	8.39
37			.60	11.47
38			.64	12.34
42			.43	7.72

Reliability

As argued in the methodology section of the article, the Cronbach's alpha, α, was used to determine the reliability coefficient of the instrument. The Cronbach's alpha was calculated as .67. α can range from .00 to 1.00. Although 1.00 is perfect reliability, .67 is considered an acceptable reliability coefficient for an achievement test (Shum, O'Gorman, & Myors, 2006). The reliability of the instrument may appear low at first glance, or it may be recommended to increase the reliability by lengthening the instrument. We did not, however, want to proceed in this way. This instrument was developed for kindergarten level students, and these students have short attention spans. Therefore, children can easily be bored with a lengthy instrument. That is, the lengthening of the instrument can prevent the applicability of this kind of instrument.

Discussion and Comments

The following steps were completed successfully in this study to develop an instrument for Turkish kindergarten students. First, the literature on Turkish children's understandings of science concepts was examined. Second, the indicators in the Turkish Preschool Program that referred to some science concepts and scientific inquiry processes were determined. Third, an item pool was constituted with 42 items, of which 30 items belonged to the SCS and the rest of them to the SIS. Content validity and language validity of the items were then studied by way of SAEs and TLEs. The results showed that all items reached acceptable CVR values and there was significant consensus among TLEs on the items' understandability. These findings encouraged us to start the next step of this study. In the next step, all items were administered to 371 Turkish kindergarten students, and 16 items were selected in light of the optimized item difficulty and discrimination indexes. Lastly, hypothetical factor structure and internal-consistency reliability were examined.

Our findings showed that the instrument has adequate psychometric properties and is an instructionally sensitive tool that may be used to measure Turkish kindergarten students' understandings of some science concepts and scientific inquiry processes. As noted at the outset, this instrument may serve as a useful tool for future program developers and researchers who are attempting to understand the effectiveness of the Turkish Preschool Program or to reveal children's understandings of some science concepts and scientific inquiry processes. Also, the instrument is well-organized, age and linguistically appropriate, and supported with pictures to make the items more understandable.

The instrument also has some remarkable characteristics which can make it attractive. First, the average difficulty and discrimination values are .63 and .44, respectively, and both of them are very close to the optimum values outlined in the literature (Kaplan & Saccuzzo, 1997). Second, both SAEs and

TLEs rated all items with high scores, 8.80 and 9.50 out of 10, respectively. To be precise, the instrument has items which are easily understandable by an average Turkish kindergarten child and reflect the content of targeted understandings. Last, it has acceptable goodness-of-fit indexes and internal consistency.

Although preliminary properties of the instrument indicated promising results, we are aware that there is still room to make the instrument stronger. For example, it can be administered to a larger sample and in different places in a future study. The results can then be analyzed based on the findings of the present study to check the psychometric properties of the instrument.

Acknowledgement

This work was supported by The Scientific and Technological Research Council of Turkey. The corresponding author is grateful to Prof. Dr. Ala Samarapungavan for hosting him at Purdue University. Also, the authors would like to special thanks to their collaborators and participants for their valuable feedback and participation during the instrument development process.

References

A Nation at Risk. (1983). *A nation at risk: The imperative for educational reform*. Retrieved from http://www2.ed.gov/pubs/NatAtRisk/recomm.html.

Akman, B., Ustun, E., & Guler, T. (2003). Using science process skills in six years old children. *Hacettepe University Journal of Education*, *24*, 11-14.

Atalay-Turhan, T., Koc, Y., Isikal, M., & Isikal, H. (2009). The new Turkish early childhood teacher education curriculum: A brief outlook. *Asia Pacific Education Review*, *10*, 345-356.

Ayvaci, H. S. (2010). A pilot survey to improve the use of scientific process skills of kindergarten children. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 4*(2), 1-24.

Ayvacı, H. S., Devecioğlu, Y., & Yiğit, N. (2002, September). *Okul oncesi ogretmenlerinin fen ve doga etkinliklerindeki yeterliliklerinin belirlenmesi* [An assessment on the proficiency of early childhood teachers in science and nature activities]. Paper presented at the 5th National Science and Mathematics Congress, Ankara.

Bahar, M. (2006). *Fen ve teknoloji ogretimi (Science and technology teaching).* Ankara: Pegema Yayincilik.

Barell, J. (2008). *Why are school buses always yellow?* Thousand Oaks, California: Corwin Press.

Bedel, E. F. (2008). Interactions among attitudes toward teaching and personality constructs in early childhood preservice teachers. *Journal of Theory and Practice in Education*, *4*(1), 31-48.

Bentley, M. L., Ebert, E. S., & Ebert, C. (2007). *Teaching constructivist science K-8.* Thousand Oaks, California: Corwin Press.

Brewer, J. A. (1998). *Introduction to early childhood education*. Needham Heights: Allyen & Bacon.

Bryant, D. M., & Clifford, R. M. (1992). 150 years of kindergarten: How far have we come? *Early Childhood Research Quarterly*, *7*, 147-154.

Bulut, M. (2007). Curriculum reform in Turkey: A case of primary school mathematics curriculum. *Eurasia Journal of Mathematics, Science & Technology Education*, *3*(3), 203-212.

Chan, F., Lee, G. K., Lee, E., Kubota, C., & Allen, C. A. (2007). Structural equation modeling in rehabilitation counseling research. *Rehabilitation Counseling Bulletin*, *51*, 44-57.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods and Research*, *36*, 462-494.

DeVellis, R. F. (2003). *Scale development: Theory and applications.* California: Sage.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3th ed.). Circle Pines, MN: American Guidance Service.

Durmusoglu, M. C. (2008). An examination of the opinions of preschool teachers about preschool learning settings in their schools. *Eurasian Journal of Educational Research, 8*(32), 39-54.

Erden, F. T., & Sonmez, S. (2011). Study of Turkish preschool teachers' attitudes toward science teaching, *International Journal of Science Education*, *33*(8), 1149-1168.

Eshach, H. (2011). Science for young children: A new frontier for science education. *Journal of Science and Technology Education*, *20*, 435-443.

Eshach, H., & Fried, M. N. (2005). Should science be taught in early childhood? *Journal of Science Education and Technology, 14*(3), 315-336.

French, L. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly*, *19*, 138-149.

Gelman, R., & Brenneman, K. (2004). Science learning pathways for young children. *Early Childhood Research Quarterly*, *19*, 150-158.

Gronlund, N. E., & Waugh, C. K. (2009). *Assessment of student achievement*. New Jersey: Pearson.

Gullo, D. F. (2005). *Understanding assessment and evaluation in early childhood education*. New York: Teachers College Press.

Hair, J. F., Anderson, R. E., & Black, W. C. (1995). *Multivariate data analysis: With readings*. Englewood Cliffs, NJ: Prentice Hall.

Harlan, J. D., & Rivkin, M. S. (2008). *Science experiences for the early childhood years: An integrated affective approach*. New Jersey, Ohio: Pearson.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional Criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Jöreskog, K. G., & Sörbom, D. (2003). *LISREL Version 8.54: User's reference guide* [electronic manual]. Chicago, Ill: Scientific Software International Inc.

Kabadayi, A. (2010). Investigating demographic characteristics and teaching perceptions of Turkish preschool teachers. *Early Child Development and Care, 180*(6), 809-822.

Kaplan, R. M., & Saccuzzo, D. P. (1997). *Psychological testing: Principles, applications, and issues* (4th ed.). Pacific Grove, California: Brooks/Cole Publishing Company.

Klein, E. R., Hammrich, P. L., Bloom, S., & Ragins, A. (2000). Language development and science inquiry: The head start on science and communication program. *Early Childhood Research and Practice*, *2*(2), 1-22.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*, 563-575.

Mantzicopoulos, P., Samarapungavan, A., & Patrick, H. (2009). "We learn how to predict and be a scientist": Early science experiences and kindergarten children's social meanings about science. *Cognition and Instruction*, *27*(4), 312-369.

Martin, R., Sexton, C., Franklin, T., Gerlovich, J., & McElroy, D. (2009). *Teaching science for all children: An inquiry approach.* Boston: Pearson.

Menekse, M., Clark, D. B., Ozdemir, G., D'angelo, C., & Scheligh, S. (2009, September). *Turkish students' force meanings.* A paper presented at the European Science Education Research Association Conference, Istanbul, Turkey

Metz, K. E. (2004). Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations of their own design. *Cognition and Instruction, 22*, 219-290.

Ministry of National Education. (2012). *Okul oncesi egitimi programi: 36-60 aylik cocuklar icin* [Preschool education program: For children 36–60 months old]. Retrieved from http://ttkb.meb.gov.tr.

National Research Council. (1996). *National Science Education Standards*. Washington, DC: National Academy Press.

Ojala, M., & Talts, L. (2007). Preschool achievement in Finland and Estonia: Cross-cultural comparison between the cities of Helsinki and Tallinn. *Scandinavian Journal of Educational Research*, *51*(2), 205-221.

Olsen, D., & Zigler, E. (1989). An assessment of the all-day kindergarten movement. *Early Childhood Research Quarterly*, *4*, 167-186.

Puckett, M. B., & Black, J. K. (2008). *Meaningful assessment of the young children: Celebrating, development and learning.* New Jersey: Pearson.

Rankin, L. (2011). Pathways to Inquiry. *Science and Children*, *48*(6), 8-9.

Roberts, J. K. (1999, February). *Basic concepts of confirmatory factor analysis.* Paper presented at an annual meeting of the southwest educational research association, San Antonio, TX, US.

Sackes, M., Flevares, L. M., & Trundle, K. C. (2009). Four-to six-year-old children's conceptions of the mechanism of rainfall. *Early Childhood Research Quarterly, 25*, 536-546.

Samarapungavan, A., Mantzicopoulos, P., & Patrick, H. (2008). Learning science through inquiry in kindergarten. *Science Education*, *92*(5), 868-908.

Samarapungavan, A., Mantzicopoulos, P., Patrick, H., & French, B. (2009). The development and validation of the science learning assessment (SLA): A measure kindergarten science learning. *Journal of Advanced Academics, 20*, 502-535.

Secer, Z. (2010). An analysis of the effects of in-service teacher training on Turkish preschool teachers' attitudes towards inclusion. *International Journal of Early Years Education, 18*(1), 43-53.

Shapiro, M. S. (1983). *Child's garden. The kindergarten movement from Froebel to Dewey*. University Park: The Pennsylvania State University Press.

Shepard, L., Kagan, S. L., & Wurtz, E. (1998). *Principles and recommendations for early childhood assessments* (National Education Goals Panel). Washington, DC: U.S. Government Printing Office.

Shum, D., O'Gorman, J., & Myors, B. (2006). *Psychological testing and assessment*. Oxford: Oxford University Press.

Simsek, O. F. (2007). *Introduction to structural equation modeling.* Ankara: Ekinoks Press.

Smith, A. (2001). Early childhood-A wonderful time for science learning. *Australian Primary & Junior Science Journal, 17*(2), 18-21.

Tekindal, S. (2008). *Assessment and evaluation in education*. Ankara: Pegema Press.

Tu, T. (2006). Preschool science environment: What is available in a preschool classroom? *Early Childhood Education Journal*, *33*(4), 245-251.

Turkish Statistical Institute. (2010). *Addressed based population registration system results of 2010* [Online]. Retrieved from http://www.turkstat.gov.tr/AltKategori.do?ust_id=11.

Watters, J. J., Diezmann, C. M., Grieshaber, S. J., & Davis, J. M. (2000). Enhancing science education for young children: A contemporary initiative. *Australian Journal of Early Childhood, 26*(2), 1-7.

Walsh, W. B., & Betz, N. E. (2000). *Tests and assessments* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement.* Itasca, IL: Riverside.

Wortham, S. C. (2008). *Assessment in early childhood education*. New Jersey: Pearson.

Zhang, M., Parker, J., Eberhardt, J., & Passalacqua, S. (2011). "What's so terrible about swallowing an apple seed?" Problem-based learning in kindergarten. *Journal of Science and Technology Education*, *20*, 461-481.