Adaptation of scale "Working Like Scientists" (WLS) A Turkish-language version: validation and reliability

Güntay Taşçi*[1], Maike Looss[2], Halil Yurdugül[3], Dagmar Hilfert-Rüppell[2], Ali Sülün[1], Dagmar Hinrichs[2], Sedat Aydogdu[1], Konstantin Klingenberg[2], Fehime Şeyma Taş[1]

1 Faculty of Education, Erzincan University, Erzincan, Turkey
2 Institute of Science Education, TU Braunschweig, Braunschweig, Germany
3 Faculty of Education, Hacettepe University, Ankara, Turkey

The aim of this study was to validate an instrument implemented by Höner et al. (2010) called "Working Like Scientists", which was designed according to German culture. The translation of the items from German to Turkish and back to German again was performed by Turkish and German working groups in an iterative process, eliminating incomprehensibilities and misfits of specific words. In a first step, the Turkish and German questionnaires were completed in a week by a total of 44 Turkish prospective teachers. So in this way, the data of statistical evidence of equivalence in both languages were collected (r=0. 74; p=0. 00). To establish the factorial construct for the Turkish instrument, the data set obtained from 304 Turkish prospective teachers was analysed by the CFA using a Lisrel 8.8 program. To prove the evidence of discriminant and convergent validity of the established model of the study, AVE (average variance extracted) values and the square roots of AVE value were calculated. In addition, the scale's prediction status of epistemological beliefs as a higher construct was examined by the second order CFA. Furthermore, the evidence for the scale's reliability was examined by item total correlations (> .30); internal consistency (Cronbach's Alpha: between 0.694 and 0.805), and construct reliability (McDonald´s Omega: between 0.694 and 0.806), and later the suitable values for both sub-dimensions and the whole scale.

Introduction

Conceptual framework: a historical aspect for relevant concepts

Knowledge, one of the basic subjects of philosophy, has been often researched by natural scientists for the last two decades. Therefore, in science education, we have frequently dealt with epistemology with respect to the nature of knowledge, the source of knowledge and acquiring knowledge. Epistemology consists of three basic areas investigating the limits, the source and the features of human knowledge (Welch, Roy, & Ray, 2012). Schraw (2013), who points out the difference between ontology (which studies reality and being) and

---

*Correspondence: gtasci@erzincan.edu.tr

epistemology, says that epistemology deals with theory of knowledge and rationality, and he extends the definition of epistemology with the methods and justification of human knowledge. On the other hand, the papers which focus on a limited part of epistemology covering a large body of content use the term epistemological beliefs (Schraw, 2013; DeBacker, Crowson, Beesley, Thoma, & Hestevold, 2008). Schraw (2013) explains the terms epistemological worldview and personal epistemology, or epistemological stances, which is a synonym of the former in literature. An epistemological worldview is defined as a set of beliefs or theories of knowledge about acquiring and applying information. Neumann and Kremer (2013) renamed these terms as "developmental psychological approaches". Apart from these, cognitive psychological approaches are studied. According to the classification suggested by Hofer and Pintrich (1997), epistemological beliefs consist of the "nature of knowledge" and the "nature of knowing". These areas, which are considered the core of epistemological beliefs, are again divided into two dimensions based on empirical studies: nature of knowledge (certainty of knowledge) and simplicity of knowledge. Nature of knowing consists of two dimensions as "source of knowledge" and "justification of knowing" (Neumann & Kremer, 2013). It is known that the debate by Neumann and Kremer (2013) as to whether epistemological beliefs should be considered solely in the domain or independently, free from the domain, has introduced the nature of science approach strongly based on the philosophy of science. The studies in the philosophy of science trying to understand science have helped to discover various approaches. These approaches are the studies of the historical development of science and looking into science in terms of rationale and philosophy. When science is considered in terms of rationale and philosophy, science turns out to be both a result and a process (Yıldırım, 2007, pp. 11). According to Yıldırım (2007), when science is looked at as a result, it is thought of as a systematic set of knowledge, but when it is looked at as a process, it is considered to be an actual intellectual pattern. Science, philosophy and history all discuss the process of producing scientific knowledge in terms of different factors such as methods, values, assumptions, and conceptual inventions. These scientific methods to construct scientific knowledge have been studied throughout the history of science by prominent scientists. Meanwhile, from all these conceptual discussions, two philosophical views called empiricism and rationalism have emerged. This conceptual approach is considered to have produced the term, scientific epistemological beliefs, which is called learners philosophical views of science by Tsai (2000). Therefore, learning under the effect of a constructivist approach as a new paradigm in psychology and philosophy and constructing this knowledge in both science and learners are considered to be similar processes. Lederman (1992) states that the nature of science can be defined in many ways, but most of the time it is defined as the development of valuable scientific knowledge. After all, these conceptual processes are related to teaching science at schools in their developing curricula. Lederman (1992) points out that the desired outcome of teaching science is to try to help understand science as an efficient way of knowing. In this respect the National Research Council (NRC, 1996) refers to the concept of scientific literacy by methods of scientific research and the knowledge of scientific concepts. "Scientific inquiry" and the "nature of science" are the core elements of scientific literacy (Schwartz, Lederman, & Crawford, 2004). Also, scientific inquiry is defined as a method and activities of producing scientific knowledge such as observations, determining the problem, investigating the sources (books, journals, etc.), raising hypothesis, planning and conducting the experiments, finding out answers, estimating, and explaining. According to Bybee, (2000), scientific inquiry in the context of science education is made up of three elements: skills, knowledge and pedagogical approaches for teaching.

Available measurement tools: an overview

The so-called conceptual process has helped to produce a wide range of literature that consists of various measurement approaches to acquire all these features mentioned above. There are six different strategies of measurement used by Schraw (2013). Of these, the most used are scales, interviews, questionnaires and concept maps. Essential data collection tools to determine individuals' epistemological beliefs and views of the nature of science are shown in Table 1.

Of these scales, EQ, developed by Schommer (1990), EBQ, by Hofer (2001), and EBI, by Schraw, Bendixen and Dunkle (2002) are all scales developed in similar theoretical frameworks and item pools. EQ (Schommer, 1990), developed as a precursor, consists of 63 items and five dimensions shown in the table. A modified version of EQ, EBI (Schraw, Bendixen & Dunkle, 2002) consists of 32 items and 4 dimensions. As an alternative to these scales, EBQ (Hofer, 2001) suggests a pattern of 4 factors with two subsumed terms. These scales are compared in a number of studies in terms of their psychometric features, and tried to be reconstructed (Welch et al., 2012; Ordonez, Ponsoda, Abad, & Romero, 2008). Although these scales consist of elements developed on a basis of a strong theoretical framework, in the studies on psychometric analysis of them those with similar frames, such as Epistemological Beliefs Survey (EBS; Wood & Kardash, 2002) and Connotative Aspects of Epistemological Beliefs (CAEB: Stahl and Bromme, 2007), with low reliability, low explained variances and some problems in the factorial constructs are stated (Schraw, 2013; DeBacker et al., 2008). The scales consist of different number of factors with some ambiguity in the factorial constructs, which is a subject of discussion. Another problem is that the same factor constructs are not obtained when the same scales are applied to different samplings. History of science and scientific procedures which have become more important as a result of considering epistemological beliefs in the framework of science philosophy have made the epistemological views of science a subject to be studied.

Table 1. Overview of selected instruments of epistemological beliefs and of nature of science, their developers, and dimensions

Instrument	Developers	Dimensions/Structure
Views of Science-Technology-Society (VOSTS)	Aikenhead, Ryan and Fleming (1989)	114 Items multiple – choice questions
Epistemological Questionnaire (EQ)	Schommer (1990)	Certainty of knowledge Source of knowledge Structure of knowledge Control of knowledge acquisition Speed of knowledge acquisition
Views on Science-Technology-Society" (VOSTS)	Aikenhead and Ryan (1992)	Multiple – choice questions
Scientific Epistemological Beliefs (SEB)	Pomeroy (1993)	Bipolar Structure: Traditional view of science (empiricist view) Non-traditional view of science (constructivist views)
Epistemological Beliefs Questionary (EBQ)	Hofer (2001)	Certainty of Knowledge Simplicity of Knowledge Source of Knowledge Justification of Knowledge
Epistemological Beliefs Inventory (EBI)	Schraw, Bendixen and Dunkle (2002)	Certain Knowledge Innate Ability Quick Learning Simple Knowledge

		Omniscient Authority
Views on Nature of Science Questionnaire, Form C (VNOS-C)	Lederman, Abd-El-Khalick, Bell, & Schwartz (2002)	10 Open - ended questions and interviews
Views of scientific inquiry (VOSI)	Schwartz, Lederman, & Lederman (2008)	Open – ended question

These epistemological views of science involve the nature of science and the nature of scientific inquiry (Schwartz, Lederman, & Lederman, 2008). SEB (Pomeroy, 1993), which was developed in this respect, is a data collection tool of 17 items and of a 5-point Likert type with one dimension. It is stated that this scale is bipolar, which means that it reveals whether students' views of science philosophy are closer to the empirical view (called traditional) or to the constructivist view (called non-traditional), according to the points they gathered from their agreement with the items. VOSTS (Aikenhead, Ryan, & Fleming, 1989), VNOS-C (Lederman et al., 2002), and VOSI (Schwartz, Lederman, & Lederman, 2008) are the data collection tools which are not of a Likert type, and which consist of open items. Aikenhead and Ryan (1992) state that the ambiguity in the Likert-type scales is a factor that weakens the validity of the tests. Instead, they say, the studies carried out with open-ended questions and semi-structured interviews can reflect the understanding of the learners in a better way. These three so-called tools of measurement were developed by constructing their theoretical frame (NOS Aspect; Characteristics of VOSTS, etc.) with open-ended or multiple-choice questions.

This study, whose aim is not to suggest a new approach of measurement or a new scale, addresses developing a data collection tool that will be used in a project to be carried out in Turkey and Germany. The original version of this scale used in this study was applied to prospective teachers in Germany by Höner, Strahl, Müller, Eghtessad, Pietzner, Looß, Klingenberg and Gläser (2010). This study of adapting this scale, aiming to measure prospective teachers' epistemological stances on science, seems to be significant regarding the discussions about the measurement of the nature of science and the nature of scientific research based on epistemological beliefs discussed in the literature.

Method

Participants

Data were collected from a total of 348 prospective teachers to test the construct validity of the scale. The participants in the first stage consisted of 44 (24 female, 20 male) prospective teachers who attended the department of German Education Program at Hacettepe University, Ankara, and all of the participants were in their fourth semester. This stage was performed through a correlation study to confirm the accordance of Turkish and German items for language equivalency.

The participants in the second stage were group of 304 (54% female, 46% male) Turkish prospective teachers in the department of Primary Education in Erzincan University. The majority of the participants (42.4%) were in their first semester, 10.2% in their third semester, 27.6% in their fifth semester and 19.4% in their seventh semester. The academic average score of samples was 2.80, and their ages were between 18 and 25 years. In this stage the confirmation factorial construct and reliability of the scale were examined.

Instrument

The original scale, "Working Like Scientists" (WLS), consisted of 13 items that are of a 5-point Likert type, and the item response categories of scale were ordered from "totally agree" to "strongly disagree". The scale, first developed by Gromadecki and Mikelskis-Seifert (2006) for sixth-grade students, was applied to German prospective teachers by Höner, Strahl, Müller, Eghtessad, Pietzner, Looß, Klingenberg and Gläser (2010). The research was carried out with exploratory factor analysis (principal component analysis and varimax rotation) to test the construct validity of the WLS scale. According to the results of this analysis, constructs with four factors were formed and named as uniqueness of knowledge (UK), importance of experiments (IE), durability of knowledge (DK) and procedure for acquisition of knowledge (PAK). The internal consistency of the scale whose factor loads range from 0.410 to 0.740 was calculated as 0.57 by using Cronbach's Alpha.

Procedure: Scale construction

The scale construction was conducted in three phases: scale translation, language identification (with statistical application), and adaptation.

Firstly, the items of the WLS scale were translated into Turkish within the research team. In this process two bilingual translators were involved. Another subject-matter expert revised this translation on the basis of the technical terms and their comprehensibility and consistency. In this context, all of the items were translated into Turkish. Some of these items were formed as alternate items in Turkish to express clearly in compliance with Turkish language. Thereafter, the entire list of items was translated back into German, and it was sent to working group at the Technical University of Braunschweig in Germany. At this level, the German working group compared the items gained from back translation with the ones in the original scale. Thus, the conceptual situation and the clarity of the items were especially maintained. In this way, two identical forms were obtained: Working Like Scientists, Turkish language (WLS-TR) and Working Like Scientists, German language (WLS- DE).

Data analyses

In the process of the development of a scale regarding a certain theoretical framework, confirmatory factor analysis was performed to show the accordance of measurement models and data models (Watkins, 1989). In order to confirm the factor structure determined through EFA, CFA was performed in the second sample, and so the factor structure of the scale was tested. In this phase of adaptation, the correlation of data and sampling were tested for the data collected from Turkish samples by CFA. The convergent and discriminant validities of the scale obtained were analysed by calculating AVE values and their square roots (Fornell & Larcker, 1981). Construct reliability (McDonald's Omega), which fitted the data set for congeneric measurements and which was calculated with the standard path coefficients of items and error terms, and Cronbach's Alpha coefficients to represent internal consistency were calculated (Nunnally & Bernstein, 1994). The goodness of fit of the theoretical model structure to the empirical data was assessed by various fit indices. It was recommended to attest some of these indices for the model fit and thus make a decision (Steiger, 1990). In this study Chi Square / df; RMSEA; NNFI; NFI; CFI; GFI were the values used for the model fit indices. Furthermore, at the beginning, the cohesion between the Turkish and German language scales were tested using the Pearson correlation.

Results

Adaptation

Determination of measurement models

The 4-factor structure was obtained with EFA, which was tested through CFA. Three different models were formed for this. Of these models, the first one was the single factor model, assuming that 13 items had one factor; the second one was the uncorrelated model based on the assumption that the 4 factors formed through EFA results were separate structures, and the third one was the correlated model based on the assumption that the items formed a 4-factor structure. The connection diagram for these three models was analysed on the Lisrel 8.8 programme. It is clear that the goodness of fit indices is not at acceptable levels for one-factor model and uncorrelated model, whereas the indices are acceptable for the correlated model [($\chi2$ (62, N = 304) = 158.66 p < .000, $\chi2$ / df= 2.56, RMSEA = 0.077, GFI = 0.92, CFI = 0.98, NNFI = 0.98, NFI = 0.97]. Therefore, the four-factor structure of the scale was accepted and the modifications to the model were checked afterwards. Accordingly, m9 and m2 items yield high correlations with other factors and with error variances. The model data fit indices found with these models are shown in Table 2.

Table 2. CFA Fit indices for the models

Fit indices	Perfect fit	Acceptable fit	One-factor model	Uncorrelated model	Correlated model
$\chi2$/df	$\chi2$/df <3	3< $\chi2$/df <5	3,14	16,69	2,19
RMSEA	0<RMSEA<,05	,05<RMSEA<,08	0,084	0,228	0,063
NNFI	0,97≤NNFI≤1	0,95<NNFI<0,97	0,97	0,78	0,98
NFI	0,97≤NFI≤1	0,95<NFI<0,97	0,97	0,81	0,98
CFI	0,97≤CFI≤1	0,95<CFI<0,97	0,98	0,82	0,99
GFI	0,95≤GFI≤1	0,90<GFI<0,95	0,91	0,64	0,94

As a result of repeating the analyses after correlating the error terms of those items, it was found that the fit indices were improved and that they reached acceptable or perfect fit levels adapting [($\chi2$ (58, N = 304) = 126.71 p < .000, $\chi2$ / df= 2.19, RMSEA = 0.063, GFI = 0.94, CFI = 0.99, NNFI = 0.98, NFI = 0.98]. The connection diagram prepared as a result of the analyses is shown in Figure 1 along with standard coefficients.
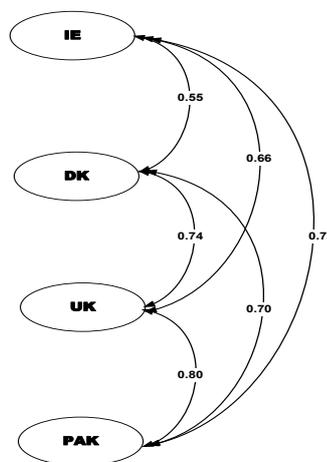

Figure 1. First order CFA connection diagram

Examining the correlations between factors, it was found that the lowest level of relation was between the importance of experiments (IE) and the durability of knowledge (DK; 0.55), and that the highest relation was between the uniqueness of knowledge (UK) and procedure for the acquisition of knowledge (PAK; 0.80). In Table 4, it can be seen that the standardized path coefficient or factor loadings are between 0.48 and 0.84. Thus, the adapted four factorial models were confirmed.

Validity

According to the definition of the validity, the validity between the item level and the level of the overall test connections with one or the other constructs was examined. Therefore, the detection of convergent and discriminant validity should be required as essential indicators of the presence of construct validity in the causal-analytical validity. Convergent validity means that valid measurement with different factors is highly correlated with each other. In contrast, the goal of discriminant validity is to ensure that the measurements of a factor do not correlate so much with the measurements of another factor. AVE is proposed for an empirical confirmation of convergence as criteria. The confirmation of the discriminant validity occurs through the assessment of the correlation between the factors if the square roots of AVE are greater than the correlation of the two factors. Consequently, the AVE values and the square roots of the AVE values were calculated. The findings are shown in Table 3. The square roots of the AVE values are bigger than the values of correlations with other dimensions. This situation is thought to give evidence for discriminant validity. On the other hand, it is clear that the average variances explained (AVE) for latent variables as evidence for convergent validity are bigger than 0.50. This is considered to be evidence for convergent validity.

Table 3. Descriptive Statistics for sub-dimensions, correlations between dimensions, and square roots of AVE

	IE	DK	UK	PAK	Mean	Std. Deviation	AVE
IE	**0,77***				11,22	3,31	0,58
DK	0,552	**0,76***			11,27	3,69	0,63
UK	0,664	0,739	**0,80***		15,26	4,16	0,60
PAK	0,730	0,705	0,798	**0,87***	11,85	3,57	0,76

*:Square roots of AVE values

To demonstrate that the four sub-dimensions revealed through confirmatory factor analysis represent the epistemological beliefs which were announced in the original form of the scale, the structural relations of the four dimensions with upper order variables were tested with the second order confirmatory factor analysis. For these purposes, the second order *epis* latent variable was added to the connection diagram, and thus analyses were done. As a result of the analyses, it was found that the fit indices were improved and that they reached acceptable or perfect fit levels adapting [($\chi2$ (58, N=304) = 160.98 p< 0.000, $\chi2$/df = 2,93, RMSEA = 0.080, GFI = 0.92, CFI = 0.97, NNFI = 0.96, NFI = 0.96]. On examining the factor loads between the first order latent variables and the epistemological beliefs, it was found that the uniqueness of knowledge (UK) (β = 0.98, t = 2.69, p < 0.05), the importance of experiments (IE) (β = 0.90, t = 4.83, p < 0.05) were the strongest components. Durability of knowledge (DK) (β = 0.86, t = 6.50, p < 0.05) and procedure for acquisition of knowledge (PAK) (β = 0.71, t = 15.57, p < 0.05) were, on the other hand, lower, yet strong components.

These components, which are described as scientific epistemological beliefs, were also statistically verified in this sample.

Reliability

For the full scale, Cronbach's Alpha 0.918 and McDonald´s Omega 0.920 were calculated. The value for the factors was described in Table 4.

Table 4. Reliability for the factors, and descriptive statistics of the items

Factors	Items	Averages	Factor Loads	Std. deviations	Item total correlations	McDonalds Omega	Cronbach's Alpha
IE	m3	3,89	0,79	1,44	0,700		
	m4	3,72	0,71	1,43	0,630	0,705	0,694
	m10	3,6	0,48	1,33	0,410		
DK	m5	4,00	0,83	1,48	0,762		
	m9	3,64	0,68	1,62	0,624	0,737	0,758
	m11	3,64	0,56	1,4	0,587		
UK	m2	3,75	0,60	1,42	0,527		
	m6	3,97	0,77	1,38	0,729	0,714	0,779
	m8	3,62	0,60	1,22	0,549		
	m12	3,92	0,81	1,34	0,760		
PAK	m1	3,93	0,70	1,32	0,664		
	m7	4,08	0,84	1,43	0,810	0,805	0,806
	m13	3,83	0,74	1,45	0,715		

As indicated in Table 4, the item total correlations of the items on the scale are greater than 0.30, and vary between 0.361 and 0.685. The factors on the scale indicate the values between 0.694 and 0.806 for Cronbach's Alfa. The Cronbach's Alpha values and McDonald´s Omega values are very close to each other, and also all values are above 0.70.

Discussion and Conclusions

In the studies of the scale adaptation, firstly the criteria of scale adaptation proposed by Hambelton and Patsula (1999) were tried to establish the scale developed in Germany. It should be suitable for the sociocultural structure of Turkey. The questionnaires developed in this way were administered to 44 Turkish students, and the correlation findings obtained (0.74; p<0.05) were at very good levels. Then, whether the findings of validity and reliability of the original scale's 4-factor construct were established in the Turkish version was investigated. In this respect, the goodness of fit indices for correlated, uncorrelated and one-dimensional models are at the higher acceptable level of the correlated model. Also, it was recorded that to test the convergence and discriminant validity of the scale, the variance and square root values analysed for the dimensions are the same level as indicated in the literature. In addition, the reliability data of the scale (Cronbach's alpha: 0.918; McDonald´s omega: 0.920) are as high as expected. Finally, the second order CFA indicates that these developed factors can predict a significantly higher level.

The statistical findings obtained when adapting the scale provide evidence that the scale is valid and reliable, and that it can be administered in both cultures. The factorial construct of

the original scale, which was confirmed after the adaptation, consists of uniqueness of knowledge (UK), importance of experiments (IE), durability of knowledge (DK) and procedure for acquisition of scientific knowledge (PAK). Also, it is indicated that the UK and DK dimensions reflect "science and certainty" in the literature. By the same token, BE reflects "methods and hypothesis" while PAK reflects "hypothesis and production" (Höner et al., 2010). Apart from this, the comparison of the scale's dimensions and items with the ones in the literature is summarized in Table 5. There is consistency between VNOS – C, which was developed on the grounds of NOS aspect, and VOSI measurement tools and scale dimensions and item statements developed with SEB. These tools of measurement contain more items or provide more data related to more dimensions with NOS aspect, but also it can be said that the adapted scale contains less about NOS.

Table 5. Comparison of the scale's dimensions and items with the statements of other data collection tools

Factors	Items*	SEB	VNOS – C	VOSI
Importance of experiments (BE)	4.Results of experiments are used to check predictions.	30. Science is based on experiments which any other competent scientist in the field should be able to repeat at will.	(2) What is an experiment? (3) Does the development of scientific knowledge require experiments?	What do you think a scientific experiment is? What types of activities do scientists do to learn about the natural world? Discuss how scientists do their work.
Durability of knowledge (BW)	9.Scientific knowledge will not change in the future. 11.The results of experiments are always true.	13. Insofar as a theory cannot be tested by experience, it ought to be revised so that its predictions are restricted to observable phenomena.	(4) After scientist have developed a scientific theory, does the theory ever change?	
Uniqueness of knowledge (EW)	10. Scientists often have different hypotheses about a science subject. 12. Scientists can draw different conclusions about the same problem.		(8)…. How are different conclusion possible if scientist in both groups have access to and use the same set of data to draw conclusions?	If several scientists, working independently, ask the same question and follow the same procedures to collect data, will they necessarily come to the same conclusions? Explain why or why not. And (b), (c)
Procedure for acquisition of knowledge (VW)	7. Before the experiment, it must be thought of how it will be done. 13.If the hypothesis and the experiment results do not match, the hypothesis is changed and it is confirmed by a new experiment.	04. The acquisition of new knowledge moves from observation to hypothesis to testing to generalizing theory.	(10) … Do scientists use their creativity and imagination during their investigations?	The "scientific method" is often described as involving the steps of making a hypothesis, identifying variables, designing an experiment, collecting data, reporting results. Do you agree that to do good science, scientists must follow the scientific method?

* They were translated into English without any statistical procedure.

The data collection tool, SEB (Pomeroy, 1993), and the Turkish version of this scale

(Deryakulu & Bıkmaz, 2003) share a number of similar items. These comparisons indicate that the original scale and its adapted version in Turkish can measure the construct related. When the relations among the four dimensions are considered, this measurement tool can measure students' views about the nature of science and the nature of scientific inquiry. In this study, which was carried out with prospective teachers from two different cultures, a tool of measurement was developed whose factorial construct and reliability criteria are convenient. Also, this study helped provide a 4-dimensional measurement tool that is valid and reliable. Moreover, it is a valid and reliable tool of measurement because this scale consists of sub-dimensions that can provide scores between 1-5 for each item and also for the scale as a whole to provide a total score. Another important feature of this scale is that it can measure prospective teachers' epistemological views on the nature of science. Ultimately, the total score obtained through this method indicates how much closer the participants are to the empirical scientific methods.

Acknowledgments

References

Aikenhead, G. S., & Ryan, A. G. (1992). The Development of a New Instrument: "Views on Science- Technology-Society" (VOSTS). *Science Education*, *76*(5), 477–491. http://doi.org/10.1002/sce.3730760503

Aikenhead, G. S., Ryan, A. G., & Fleming, R. W. (1989). Views on science-technology society. Retrieved from http://www.usask.ca/education/profiles /aikenhead/webpage/vosts.pdf

Bybee, R. (2000). Teaching science as inquiry. In J. Minstrell & E. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 20–46). Washington: American Association for the Advancement of Science.

DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The Challenge of Measuring Epistemic Beliefs: An Analysis of Three Self-Report Instruments. *The Journal of Experimental Education*, *76*(3), 281–312. http://doi.org/10.3200/JEXE.76.3.281-314

Deryakulu, D., & Bıkmaz, H. F. (2003). Bilimsel Epistemolojik İnançlar Ölçeğinin Geçerlik ve Güvenirlik Çalışması [The Validity And Reliability Study Of The Scientific Epistemological Beliefs Survey]. Educational Sciences and Practice, *2*(4), 243–257.

Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research,* 18(1), 39-50.

Gromadecki, U. & Mikelskis-Seifert, S. (2006). Naturwissenschaftliches Argumentieren [Scientific argue]. Schülerfragebogen, Klasse 6, Kiel: IPN

Hambleton, R.K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1(1), 1-30.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Journal of Educational Psychology Review*, *13*(4), 353–383.

Hofer, B. K., & Pintrich, P. R. (1997). The develop- ment of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*(1), 88 – 140.

Höner, K, Strahl, A., Müller, R., Eghtessad, A., Pietzner, V., Looß, M., Klingenberg, K. & E. Gläser (2010). Das Wissenschaftsverständnis von Lehramtsstudierenden [Epistemological Beliefs of Teacher Candidates], chim. & ct. did., 103 (36). Jg., 39-66.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, *29*(4), 331–359. http://doi.org/10.1002/tea.3660290404

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, *39* (6), 497–521. http://doi.org/10.1002/tea.10034

[National Research Council], NRC. (1996). *National Science Education Standards*. *National Academies Press*.

Neumann, I., & Kremer, K. (2013). Nature of Science und epistemologische Überzeugungen Ähnlichkeiten und Unterschiede [Nature of Science and Epistemologicial Beliefs – Similarities and Differences]. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *19*, 211–234.

Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric theory*. 3rd Edition. McGraw-Hill: New York.

Ordonez, X. G., Ponsoda, V., Abad, F. J., & Romero, S. J. (2008). Measurement of Epistemological Beliefs: Psychometric Properties of the EQEBI Test Scores. *Educational and Psychological Measurement*, *69* (2), 287–302. http://doi.org/10.1177/0013164408323226

Pomeroy, D. (1993). Implications of teachers' beliefs about the nature of science: Comparison of the beliefs of scientists, secondary science teachers, and elementary teachers. *Science Education,* 77 (3), 261-278.

Schommer, M. A. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82* (3), 498 – 504.

Schraw, G. (2013). Conceptual Integration and Measurement of Epistemological and Ontological Beliefs in Educational Research. *ISRN Education*, 19. http://doi.org/10.1155/2013/327680

Schraw, G., Bendixen L.D. & Dunkle M.E. (2002). Development and validation of the epistemic belief inventory (EBI). In: Hofer BK, Pintrich PR, Eds. Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing. 1st ed. New Jersey: Erlbaum, 261-75.

Schwartz, R. S., Lederman, N. G., & Lederman, J. S. (2008). An Instrument To Assess Views Of Scientific Inquiry: The VOSI Questionnaire. *Annual Meeting of the National Association for Research in Science Teaching*, 1–24. Retrieved from http://homepages.wmich.edu/~rschwart/

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, *88* (4), 610–645. http://doi.org/10.1002/sce.10128

Stahl, E. & R. Bromme, R. (2007). "CAEB: an instrument for measuring connotative aspects of epistemological beliefs," *Learning and Instruction,* 17 (6), 773 - 785.

Steiger, J. H. (1990). Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research*, 25 (2), 173-80.

Participatory Educational Research (PER)

Tsai, C. (2000). Relationships between student scientific epistemological beliefs and perceptions of constructivist learning environments. *Educational Research*, *42* (2), 193–205. http://doi.org/10.1080/001318800363836

Watkins, Christopher J.C.H. (1989). Learning from delayed rewards, PhD thesis, University of Cambridge, Psychology Department.

Welch, A. G., Roy, C. M., & Ray, C. M. (2012). A preliminary report of the psychometric properties of the Epistemic Beliefs Inventory. *The European Journal of Social & Behavioural Sciences*, 2301–2218. http://doi.org/10.15405/ FutureAcademy/ejsbs(2301-2218).2012.2.12

Wood, P., & Kardash, C. (2002). Critical elements in the design and analysis of studies of epistemology. In B. . Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 231–260). Mahwah, NJ: Erlbaum.

Yıldırım, C. (2007). *Bilim Felsefesi[Philosophy of Science]*. 11. Ed., İstanbul: Remzi Kitapevi.