Educational Sciences: Theory & Practice • 14(5) • 1988-1994

www.edam.com.tr/estp
DOI: 10.12738/estp.2014.5.1671

Conceptual Understanding of Science Teacher Candidates regarding the Reason for Measurement and Evaluation and for Determining Learning Change

Mehmet Altan KURNAZ[a]
Kastamonu University

Abstract

Determining the learning quality and the role of measurement and evaluation are accepted as part of the duties and responsibilities of teachers and operators in structured teaching programs. This qualitative case study research examined teacher candidates' conceptual understanding of the reasons for measurement and evaluation and for determining of the learning change to ascertain if trained teacher candidates met the required expectations. The study participants were determined using criteria sampling and consisted of 101 science teacher candidates. The data was collected using open ended questions, which were analyzed using content analysis. The research found that the teacher candidates' conceptual understanding as to the reasons for evaluation and determining the learning change were inadequate. It was found that the participants were unable to relate the measurement and evaluation purposes with learning change, and that the understanding of the measurement and evaluation concepts aimed at determining the need for learning change did not coincide. The data suggested that there could be teacher training insufficiencies and that the duties and responsibilities of the new teaching programs encumbered the teachers. In this context, it was suggested that determining the learning changes by measurement and evaluation practices be emphasized during teacher training.

Keywords

Determining the Learning Change, Teacher Candidates, Measurement, Evaluation, Conceptual Understanding.

Over the last five years the teaching programs in our country, which had been developed within the frame of the European Union adaptation standards, have been restructured with regard to learning areas and approaches. Newer teaching programs, therefore, have improved natures/structures (Kurnaz & Çepni, 2012; Milli Eğitim Bakanlığı [MEB], 2007; Üstün, Eryılmaz, & Gülyurdu, 2008). Therefore, teachers and the operators of teaching programs have new important responsibilities with regards to the areas of learning and approaches to the program (Ateş et al., 2009; Gunes & Tasar, 2006; Serin & Kanlı, 2008). This implies that teachers are expected to assimilate these renewed learning areas and approaches into their teaching, such as the constructivist learning approach which requires that learners take responsibility for their own learning progress while the teachers act as facilitators or guides (Cobern, 1993; Karamustafaoğlu, Coştu, & Ayas, 2005; Kurnaz & Çalık, 2008, 2009; Mathews, 2000; Osborne & Wittrock, 1983).

a **Mehmet Altan KURNAZ**, Ph.D., is currently an associate professor of physics education. His research interests include student mental models on fundamental concepts of science, the use of multiple representations in physics and science, and teacher education. *Correspondence:* Kastamonu University, Faculty of Education, Department of Primary Education, Kastamonu, Turkey. Email: makurnaz@kastamonu.edu.tr & altan.kurnaz@gmail.com

Table 1
*Comparison of Assessment Types**

	Diagnostic	Formative	Summative
Aim	Determining learning requirements	Deepening learning	Determining of and judging learning outcomes
Focus	Result (What's known?)	Result and process (What's known?)	Last result (How much is known?)
Measurement Method	Subsidiary techniques and success tests	Subsidiary techniques	Success tests
Using the measurement results	Performance improvement (structuring the learning environment)	Performance improvement	Decisions on performance
Feedback's nature	Performance	Performance	Grade
Timing	Beginning of the learning process	During the learning process	End of the learning process
Responsibility	Teacher	Student and teacher	Teacher

* Table is structured from the viewpoint of Harlen (2003), Harlen and James (1997), Jones (2006), Kim (2003), Parker, Fleming, Beyerlein, Apple, & Krumsieg (2001), and Odabaşı Çimer (2008) studies.

For science and technology teaching programs, teachers are required to carry out diagnostic, formative, and summative measurement and evaluation (Kurnaz & Çepni, 2012). Therefore, there needs to be a change in the way teachers perceive the measurement and evaluation process in parallel with the new teaching programs, a process which is clearly stated in the teaching programs (see also MEB, 2006, 2013). The key features of the diagnostic, formative, and summative assessments are shown in Table 1.

As shown in Figure 1, a teacher needs to first conduct a diagnostic assessment to determine a student's present knowledge. From the data obtained, a formative assessment is conducted to enable a deepening of the learning. Finally, a summative assessment is conducted to determine the student's acquisition, recognize the learning change, and identify learning gaps. What is new here is the need for the diagnostic assessment at the start of the learning process and determining the learning changes. As stated above, the diagnostic assessment is related to understanding the extent of a student's prior knowledge.

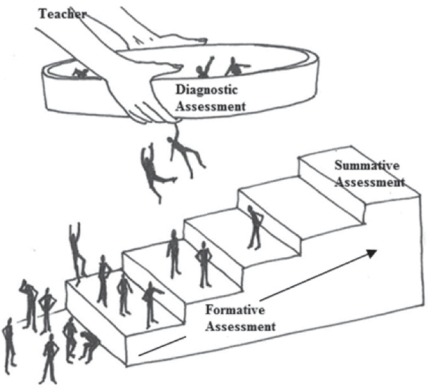

Figure 1: Application of diagnostic, formative, and summative assessments to learning environments.

According to Hailikari (2009), prior knowledge is used for (I) determining the student's level, (II) determining the class aims, and (III) designing the student learning environment. In recent research, it has been observed that students' prior knowledge level should be considered for determining learning change (Hake, 2002a, 2002b; Ladd & Walsh, 2012; Meltzer, 2002; Thompson, 2008; Zeilik, 2012). The effect of prior knowledge has been observed as an obstacle in determining learning change if only a single test is held at the completion of the subject area, so an evaluation technique which is able to reveal the real effects of the learning environment through the comparison of pretest and posttest results is preferred. Hake (1998) suggested a normalized gain formula [g = (posttest score − pretest score) / (maximum score − pretest score)] for this process. According to this formula, normalized gain values fall between 0 (acquisition/no acquisition) and 1(all possible acquisitions/acquisitions gained). By determining the normalized gain, the real change in education can be revealed (Meltzer, 2002), a technique that has been discussed in a number of previous studies (see Hake, 1998, 2002a, 2002b; Kalkan & Kıroğlu, 2007; Ladd & Walsh, 2012; Thompson, 2008; Zeilik, Schau, & Mattern, 1998). With reference to this diagnostic assessment, the conditions suitable for determining learning change are identified in this paper.

From the literature review, teacher candidate knowledge regarding new assessment and evaluation approaches activated within the curriculum (e.g., Birgin & Gürbüz, 2008; Sağlam-Arslan, Avcı, & İyibil, 2008; Taşlıdere & Özsevgeç, 2012; Yeşilyurt & Yaraş, 2011), developed through training in assessment and evaluation (e.g., Arslan & Özpınar, 2008; Uğurlu & Akkoç, 2011) and promoted through self-sufficiency(e.g., Kılınç, 2011; Mert Uyangör & Kobak, 2012 Yaman & Karamustafaoğlu, 2011)

Table 2
Comparing the Reason of Measurement and the Intended Use of Measurement Results with the Reason of Evaluation

	Reason of evaluation	Determining the student acquisition	Determining change	Grading	Teaching method	Determining the learning deficiency	Deciding on the success	Other
Reason for measurement	Determining the change	29	31	1	-	1	12	3
	Recognizing the effectiveness of the learning environment	12	8	-	-	-	5	1
	By regulations	6	1	1	-	-	-	-
	Need for objective evaluation	1	1	1	-	-	2	-
	Grading	1	1	1	1	-	-	-
	Teaching method	1	1	1	-	-	-	-
Use of measurement result	Determining the success level	19	21	2	-	-	5	2
	Giving feedback	14	11	-	-	-	4	-
	Supervising teaching activities	12	6	1	-	1	7	1
	Grading	5	2	1	1	-	3	-
	Student evaluation	-	3	1	-	-	1	1

are analyzed. However, there have been few studies which that have examined the determination of learning change and the reasons for certain assessment or evaluation procedures.

At the end of their training, teacher candidates are expected to know how to determine learning changes and how to interpret the measurement and evaluation measures. This research analyzed science teacher candidates conceptual interpretation of the reasons for the measurement and evaluation measures and the need for determining the learning change.

Method

The research was carried out from a qualitative perspective and a case study was used as the research method (Yin, 2003). The research focus was an assessment of the science teacher candidates' conceptual understanding of the measurement and evaluation measures.

Study Group

The participants were 101 final grade science teacher candidates. The study group was determined using criteria sampling, one of the purposeful sampling methods. The criteria used to select this study group were; (I) being in the final grade and (II) having successfully passed the measurement-evaluation course related to this research subject.

Data Collection Tool

6 open ended questions were used as the data collection tool to reveal the teacher candidates' understanding of the research subject. The reason for the use of open ended questions was to gain deeper information about the subject without needing to lead the teacher candidates (White & Gustone, 1992). The questions were given to 15 teacher candidates and the deficiencies were examined in parallel with an expert. The teacher candidates were given 30 minutes to answer the questions.

Analyzing the Data

A content analysis method was used to analyze the research data. During the analyses, another researcher with experience in qualitative research was consulted. Both researchers working independently conducted the analyses over three phases (*i*-naming phase, *ii*-classification/elimination phase, *iii*-category development phase) Then, working together, the researchers discussed the coding for the "agreed" and "dissented" items. Then, the reliability calculation formula suggested by Miles and Huberman (1994) was used to analyze the researchers correspondence percentage [$R_{reliability}$ = agreement / (agreement + dissidence)], which was calculated to be 88%.

Table 3
Comparison of Interpretations regarding the Diagnostic Measurement and Evaluation and the Determination of Learning Changes

	Quality of pretest	Determining foreknowledge	Designing learning environment	Determining Change	Informing about the subject
Pre-post test result comparison	39	9	3	2	
Process observation	29	5	3	1	
With grade/test	8	5	4	1	
Alternative techniques	6	-	2	-	
Other	4	3	1	1	

Findings

Conceptual Understanding Related to the Reason for the Measurement and Evaluation measures

The teacher candidates' understanding for the reason for the measurement and evaluation measures and the intended use of the measurement results were compared, as shown in Table 2.

Determining the Student Change and Conceptual Understanding regarding the Diagnostic Assessment

The teacher candidates' understanding regarding the diagnostic assessment and the determination of learning changes were compared, as shown in Table 3.

Teacher candidates' answers about the relationship between the diagnostic assessment and the determination and recognition of learning change were then tabulated, as shown in Table 4.

Table 4
The Case of Teacher Candidates' Evaluating Sample Assessment Results

Top theme	Sub Theme	Code	*f*	%
Determining learning Change	Successful evaluation	Considering the pre and post test result comparison	88	88
	Unsuccessful evaluation	Considering level difference	13	13

Discussion

Determining the features or qualifications of the target object or individual is classified as assessment, while making judgments or comments about that object or individual is classified as evaluation (Semerci, 2008; Yaşar, 2008). The teacher candidates in the study were able to judge the quality of assessment, but did not make judgments or decisions about the evaluation, and instead relied on measuring the change in the students' learning through assessment only. Further, understanding regarding the use of the evaluation results and the conceptual understanding of the reason for the evaluation did not coincide. The conceptual understanding of some teacher candidates about the reason for the measurement and evaluation did coincide, but these tended to be interpreted as having the same concepts. These findings highlighted that the science teacher candidates' understanding of the conceptual relationship between measurement and evaluation was not scientific because of a deficiency in understanding the reason for structured evaluation. These findings are similar to the results of previous studies (Arslan & Özpınar, 2008; Birgin & Gürbüz, 2008; Mert Uyangör & Kobak, 2012; Usta, Dikyol & İnce, 2010; Yaman & Karamustafaoğlu, 2011).

Some candidates were found to attribute the same meaning to both measurement and evaluation. These findings indicate that the teacher candidates participating in this research were expected to start their careers without a qualified understanding of the measurement and evaluation measures being used. Because of the effect of assessment and evaluation on learning performance (Black & Wiliam, 1998; Struyven, Dochy, & Janssens, 2005), it is clear that teacher candidates should be well trained in the assessment and evaluation procedures, and so undergraduate teaching practices should be revised to reflect the aims of the new teaching programs.

The aim of a diagnostic assessment activity is to gather information about the students' foreknowledge and, in this research, it was found that the teacher candidates understood this aim well. However, the fact that only one fifth of the teacher candidates related the diagnostic measurement and evaluation measures with the structuring of the learning environment highlighted a deficiency in this area. Another surprising aspect was that the features of the diagnostic measurement and evaluation which allowed for a determination of the student's learning change (Hake, 1998, 2002a, 2002b; Ladd & Wals, 2012; Meltzer, 2002; Thompson, 2008; Zeilik, 2012) was understood by only one tenth of the research participants. After

analyzing the conceptual understanding of the determination of student learning change, nearly half the research participants considered that the pretest-posttest data was sufficient to determine learning change, so the reason for the diagnostic measurement and the evaluation were not clear to them. Therefore, the candidates appeared to have no clear conceptual understanding about the measurement and evaluation measures required for determining and defining a learning change.

When the candidates' understanding related to the reason for the role of measurement and evaluation and for determining the learning change were considered together, the candidates focused only on the measurement activity. Therefore, the candidates did not appear to have a conceptual understanding as to how to create coherence between the diagnostic, formative, and summative assessments, which are a necessary part of the new teaching programs.

Therefore, from these findings, there are obvious deficiencies in the teacher training for new candidates with regards to the expectations in the new teaching program. Accordingly, it is suggested that the findings of this research be used to advise present teacher training to attach more importance to the measurement of learning change.

This research study had some limitations. The two main limitations are that the research was conducted on only a small group of participants, and there was only one data collection process used. By supporting this study with the results of other studies carried out over a longer time with a wider range of data collecting tools, these findings could be more generalizable, thus contributing more to the field.

References/Kaynakça

Arslan, S. ve Özpınar, İ. (2008). Öğretmen nitelikleri: İlköğretim programlarının beklentileri ve eğitim fakültelerinin kazandırdıkları. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED), 2*(1), 38-63.

Ateş, S., Serin, G., Karaömer, M., Güneş, B., Eryılmaz, A., Kanlı, U., Gülyurdu, T., & Arslan, A. (2009, August-September). *Philosophy, foundations, and vision of the new high school physics curriculum in Turkey*. Paper presented at the ESERA Conference, İstanbul.

Birgin, O. ve Gürbüz, R. (2008). Sınıf öğretmeni adaylarının ölçme ve değerlendirme konusundaki bilgi düzeylerinin incelenmesi. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 20,* 163-181.

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment.* London: Kings College School of Education.

Cobern, W. (1993). *The construction of knowledge: A radical constructivist view*. In K. Tobin (Ed.), *The practice of constructivism in science education* (pp. 51-69). Washington D.C.: AAAS Press.

Gunes, B., & Tasar, M. F. (2006, August). *An Examination of physics subjects in the new national curriculum for science and technology in Turkey*. Paper presented at the GIREP 2006 Conference: Modeling in Physics and Physics Education, Amsterdam, The Netherland.

Hailikari, T. (2009). *Assessing university students' prior knowledge implications for theory and practice.* Finland: Helsinki University Print.

Hake, R. R. (1998). Interactive engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*, 64-74.

Hake, R. R. (2002a). Lessons from the physics education reform effort. *Ecology and Society,* 2, 28. Retrieved from http://www.ecologyandsociety.org/vol5/iss2/art28/

Hake, R. R. (2002b, December). *Assessment of physics teaching methods. Proceedings of the UNESCOASPEN. Workshop on Active Learning in Physics*, Univ. of Peradeniya, Sri Lanka. Retrieved from http://www.physics.indiana.edu/~hake/Hake-SriLanka-Assessb.pdf

Harlen, W. (2003). Enhancing inquiry through formative assessment. San Francisco, CA: Institute for Inquiry Exploratorium. Retrieved from www.exploratorium.edu/IFI

Harlen, W., & James, M. (1997). Assessment and Learning: differences and relationships between formative and summative assessment. *Assessment in Education, 4*(3), 365-379.

Jones, C. A. (2006). *Assessment for learning, vocational learning support programme: 16–19*. Published by the Learning and Skills Development Agency. Retrieved from www.LSDA.org.uk.

Kalkan, H., & Kıroğlu, K. (2007). Science and nonscience students' ideas about basic astronomy concepts in pre-service training for elementary school teachers. *Astronomy Education Review*, *6*(1), 15-24.

Karamustafaoğlu, S., Coştu, B., & Ayas, A. (2005). Efficiencies of periodical table material developing with simple tools. *Journal of Turkish Science Education*, *2*(1), 19-31.

Kılınç, M. (2011). Öğretmen Adaylarının Eğitimde Ölçme ve Değerlendirmeye Yönelik Özyeterlik Algı Ölçeği. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD), 12*(4), 81-93.

Kim, S. (2003). The effect of authentic assessment strategy on students achievement in a constructivist classroom. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education* (pp. 257-260). Chesapeake, VA: AACE.

Kurnaz, M. A., & Çalık, M. (2008). Using different conceptual change methods embedded within 5E Model: A Sample teaching for heat and temperature [Online]. *Journal of Physics Teacher Education, 5*(1), 3-10.

Kurnaz, M. A., & Çalık, M. (2009). A thematic review of 'energy' teaching studies: Focuses, needs, methods, general knowledge claims and implications. *Energy Education Science and Technology Part B: Social and Educational Studies*, *1*(1), 1-26.

Kurnaz, M. A., & Çepni, S. (2012). An evaluation of changes to the Turkish high school physics curriculum. *International Education Studies, 5*(5), 92-108.

Ladd, H. F., & Walsh, R. P. (2012). Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review, 21*, 1-17.

Mathews, M. R. (2000). Constructivism in science and mathematics education. In D. C. Phillips (Ed.), *National Society for the Study of Education*, *99th Yearbook* (pp. 161-192). Chicago: University of Chicago Press.

Milli Eğitim Bakanlığı. (2006). *İlköğretim fen ve teknoloji dersi (6, 7 ve 8. sınıflar) öğretim programı*. http://ttkb.meb.gov.tr/program.aspx?tur=ilkogretim adresinden edinilmiştir.

Milli Eğitim Bakanlığı. (2007). *Ortaöğretim 9. sınıf fizik dersi öğretim programı*. http://ttkb.meb.gov.tr/www/ogretim-programlari adresinden edinilmiştir.

Milli Eğitim Bakanlığı. (2013). İlköğretim fen bilimleri dersi (3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı. http://ttkb.meb.gov.tr/www/ogretim-programlari adresinden edinilmiştir.

Meltzer, D. E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics*, *70*, 1259-1268.

Mert Uyangör, S. ve Kobak, M. (2012, Haziran). *Öğretmen adaylarının akademik başarıları ve sahip oldukları öğretmen yeterlilikleri arasındaki ilişki.* X. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi'nde sunulan bildiri, Niğde.

Miles, M. B., & Huberman, M. A. (1994). *An expanded sourcebook qualitative data analysis*. London: Sage.

Odabaşı Çimer, S. (2008). *7181 kodlu ölçme değerlendirmede yeni yaklaşımlar dersi notları.* Trabzon: Karadeniz Teknik Üniversitesi.

Osborne, R. J., & Wittrock, M. C. (1983). Learning science: A generative process. *Science Education*, *67*(4), 489-508.

Parker, P. E., Fleming, P. D., Beyerlein, S., Apple, D., & Krumsieg, K. (2001, October). *Differentiating assessment from evaluation as continuous improvement tools*. Paper presented at the 31st ASEE/IEEE Frontiers Education Conference, Reno.

Sağlam-Arslan, A., Avcı, N. ve İyibil, Ü. (2008). Fizik öğretmen adaylarının alternatif ölçme değerlendirme yöntemlerini algılama düzeyleri. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi, 11*, 115-128.

Semerci, Ç. (2008). Eğitimde ölçme ve değerlendirme. E. Karip (Ed.), Ölçme ve değerlendirme içinde (s. 2-15). Ankara: Pegem Akademi.

Serin, G., & Kanlı, U. (2008, July). *Strands in Turkish physics curriculum physics curriculum design development and validation*. Paper presented at the GIREP 2008 Conference, Cyprus.

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment* & *Evaluation in Higher Education, 30*(4), 325-341.

Taşlıdere, A. ve Özsevgeç. T. (2012, Haziran). *Fen ve teknoloji öğretmen adaylarının pedagojik alan bilgisi bağlamında strateji-yöntem-teknik ve ölçme-değerlendirme bilgilerinin incelenmesi*. X. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi'nde sunulan bildiri, Niğde.

Thompson, T. D. (2008). *Growth, precision, and CAT: An examination of gain score conditional SEM*. New York: National Council on Measurement in Education.

Uğurlu, R. ve Akkoç, H. (2011). Matematik öğretmen adaylarının ölçme-değerlendirme bilgilerinin gelişiminin tamamlayıcı-şekillendirici ölçme-değerlendirme bağlamında incelenmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, *30*, 155-167.

Usta, S., Dikyol, D., & İnce, E. (2010). The alternative evaluation tools chosen by social and science teacher candidates. *Procedia Social and Behavioral Sciences, 2*, 3457-3462.

Üstün, U., Eryılmaz, A., & Gülyurdu, T., (2008, July). *Needs assessment for Turkish high school physics curriculum design and development*. Paper presented at the GIREP 2008 Conference, Cyprus.

White, R., & Gustone, R. (1992). *Probing understanding*. London: The Falmer Press.

Yaman, S. ve Karamustafaoğlu, S. (2011). Öğretmen adaylarının ölçme ve değerlendirme alanına yönelik yeterlik algı düzeylerinin incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 44*(2), 53-72.

Yaşar, M. (2008). Eğitimde ölçme ve değerlendirmenin önemi. S. Tekindal (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (s. 2-7). Ankara: Pegem Akademi.

Yeşilyurt, E. ve Yaraş, Z. (2011). Sınıf öğretmeni adaylarının ölçme ve değerlendirme yöntemlerine ilişkin algıladıkları bilgi düzeylerinin değerlendirilmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD), 12*(4), 95-118.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). London: Sage.

Zeilik, M. (2012). *Classroom assessment techniques conceptual diagnostic tests*. Retrieved from http://www.wcer.wisc.edu/archive/cl1/flag/cat/diagnostic/diagnostic1.htm

Zeilik, M., Schau, C., & Mattern, N. (1998). Misconceptions and their change in university-level astronomy courses. *The Physics Teacher*, *36*(1), 104-107.