Critical Thinking Level of Biology Classroom Survey: Ctlobics

Atilla Çimer[1], Melih Timuçin [2], Mehmet Kokoç [3]

[1] Assistant Professor, Department of Science Education
Fatih Faculty of Education, Karadeniz Technical University
cimeratilla@yahoo.com

[2] PhD Candidate, Department of Science Education
Fatih Faculty of Education, Karadeniz Technical University
meltimucin@yahoo.com

[3] PhD Candidate, Department of Computer Education and Instructional Technology,
Fatih Faculty of Education & Distance Education Research and Application Center,
Karadeniz Technical University
mkokoc@ktu.edu.tr

ABSTRACT

Within this study, Critical Thinking Level of Biology Classroom Survey (CTLOBICS) was developed to determine to what extend Biology classroom environment supports Critical Thinking Culture. The survey was developed by following the scale development model by McMillan & Schumacher (2010). The theoretical framework of the study was adapted from Critical Thinking Strategies (Paul et al., 1990). The item pool of the survey was composed of these strategies expressed in sentences and a 59-item- scale was prepared along with the collected expert opinions. The pilot study was conducted with 387 9th, 10th and 11th grade students from two different secondary schools located in Trabzon in Turkey. The survey was finalized after Exploratory and Confirmatory Factor Analyses. The Survey has 32 items and 5 factors (Affective Critical Thinking Skills, General Thinking Skills, Basic Critical Thinking Skills, Associating with Real Life, Reasoning Skills). The inner consistency coefficient α of the scale was calculated as 0,92.

Keywords: *Critical thinking, biology education, survey, higher order thinking*

INTRODUCTION

"*Everyone thinks. But much of our thinking, left to itself, is biased, distorted, partial, uninformed or down-right prejudiced*." (Paul & Elder, 2002, p.15). The quality of our life is correlated with the quality of our thinking. We should be aware of our thinking process and we should systematically improve it (Paul *et.al.*, 1990).

Different cognitive skills have been referred as Higher Order Thinking Skills (HOTS) by different resources (Lewis & Smith, 1993; Zohar & Dori, 2003; Paul, 2005; Tilestone, 2005 p.47-58, Mulnix, 2011). Critical Thinking (CriT), Creative Thinking (CreT), Problem Solving and Reflective Thinking have been the most frequently mentioned HOTS. These HOTS overlap and coincide at times. Still, we think that CriT and CreT Skills are fundamental and it is possible to express the latter two by various combinations of CriT and CreT.

Even the basic literacy and calculation skills were redefined with the interpretation from the point of view of HOTS (OECD, 2005, p. 16; Kirsch, 2001, p. 1-5). Some scientists associate HOTS with the upper three levels of Bloom's Taxonomy of Educational Objectives (Bloom *et.al*., 1956; Krathwohl & Anderson, 2001). So, they associate lower three levels and all other irregular thinking styles with Lower Order Thinking (Zohar & Dori, 2003; Duron *et al*., 2006). Classifying HOT and Lower Order Thinking by referencing Bloom's Taxonomy is acceptable. However, it must be kept in mind that the Bloom's Taxonomy is only a classification to categorize objectives whereas HOTS like CriT or CreT are larger scale processes substantially reflecting on human behaviour.

CRITICAL THINKING

De Bono (1995) argues that "*many highly intelligent people are **bad thinkers***". He resembles the intelligence to horsepower of a car. One may still drive the car very badly even she/he has a very powerful engine. CriT Skills are like one's driving skills. After reviewing widely accepted CriT definitions, Huitt (1998) evaluated the definition "*reasonable reflective thinking focused on deciding what to believe or do*" by Ennis (1992) as the best definition and defined CriT himself as; "*the disciplined mental activity of evaluating arguments or propositions and making judgments that can guide the development of beliefs and taking action*." Unlike intelligence, CriT Skills are improvable (Walsh& Paul 1988, p13) and they do not depend on age so they should be thought at all ages (Lipman *et.al*., 1980)

Conceptualization of CriT would be incomplete without quoting the staggering annotation added by Kuhn (1993) arguing that almost all the things we teach as CriT skills are nothing but some meta-cognitive strategies whereas CtiT is rather congenital and developmental.

CRITICAL THINKING & EDUCATION

"*Both Piaget and Vygotsky thought learning is what leads to the development of higher order thinking"* (Blake & Pope, 2008). Since the sixties governments have had tendency towards educational policies to adopt teaching /learning approaches and curricula teaching how to learn, expecting students to apply academic knowledge on daily life and to analyse situations and problems when they face for the first time. This situation so escaladed the significance of concerning HOTS in education that there were even some mass campaigns launched directly aiming to increase the CriT level of all stakeholders of education (King *et al.*, n.d.). What we end up today is all contemporary teaching/learning theories require altering the instructional paradigm in a way to ensure teaching and assessing HOTS with proper instruments (Brooks & Brooks, 1993; Tileston, 2005; Paul, 2005; Keles & Cepni, 2006; Jensen, 2008; Radin, 2009; MoNET, 2007, 2008a, 2008b, 2009, 2010).

There are standardized tests to assess CriT (Ennis, 1993; Stein, 2003), CreT (Torrance, 1966) and Problem Solving (Ross & Ross, 1976). Ennis (1993) evaluated certain standardized tests assessing personal CriT skills and dispositions. It would not be unfair to mention California Critical Thinking Dispositions Inventory (CCTDI) and Watson-Glaser Critical Thinking Appraisal (WGCTA) as most popular CriT tests. CCTDI is composed of 75-Likert items and assess only dispositions towards CriT (Stein, 2003; Facione & Facione, 1994). It has translated into several languages and applied on very large samples in different countries. WGCTA has 80 items (Watson& Glaser, 1980). Although WGCTA has some inspiring CriT texts and covers different daily-life bound issues deeply, it is a multiple choice test. In the same way, Stein (2003) reported; possible test bias, lack of cross-validation studies, and low item correlations for this appraisal.

Critical Thinking Level of Biology Classroom Survey (CTLOBICS) aims to evaluate CriT Culture *of* the classroom. Although the literature review yielded variety of standardized tests measuring CriT skills and dispositions of the individuals (Ennis, 1993; Stein, 2003; Insight Assessment, 2011) as individual scores by directing problems to be solved to reveal the attainment of individuals over determined HOTS objectives, no scale was found so as to assess the state of CriT environment in the classroom *holistically* by inquiring about classroom environment and teacher habits based on student opinions.

Extensively applied, Constructive Learning Environment Survey (Taylor & Fraser, 1991) is a good example of such tests measuring the state of learning environment holistically in terms of certain aspects. One of the best ways of assessing a thinking skill is to define the sub-skills and seeking them in related settings. The CriT sub-skills were listed by different researchers (Ennis, 1985; Paul *et al*., 1990; Facione, 1990). The CriT Strategies or 35-Dimensions of Critical Thought, put forward by Paul and his team (1990), have been one of the most widely accepted CriT Sub-skill Classifications.

Dr. Richard Paul and the Critical Thinking Community (CTC) are widely accepted contributers of CriT knowledge (The Critical Thinking Community, 2011). Based on 35 Dimensions of Critical Though, CTC remodelled various level lesson plans (K-3, 4-6, 6-9 and High-school) for the USA curricula (The Critical Thinking Community-Strategy List…, 2011). Some researchers cited and applied these CriT Dimensions in their studies (Greenockle & Purvis, 1995; Allen, 2003; Harrigan & Vincenti, 2004; Dolapci, 2009; Cimer & Timucin, 2010).

Turkey, where CTLOBICS piloted, has revised all primary and secondary level curricula referencing to contemporary learning theories since the year 2000 (Ministry of National Education Turkey MoNET, 2007; MoNET, 2008a; MoNET, 2008b; MoNET, 2009). The change was sudden but all the stakeholders of education have somehow internalized the new system by time. Now the curricula, approaches and teachers are partially capable of conducting instruction minding HOTS. However, the national university entrance examination poses multiple choice questions of Application or lower level of Bloom's Taxonomy. Teachers, students and even parents use the pressure put by this examination as an excuse and they favor a very uncritical way of learning by aiming marking the right choice of the questions in the shortest time possible, generally with little reasoning (Cimer, 2004; Icbay, 2005; Timucin, 2008; Azar, 2010).

The situation for Biology course is even worse. The number of Biology course hours in most secondary school types is quite low, which makes it difficult to apply an innovation effectively or to create a continuous course culture. Thus, for Turkey case, it can be argued that the pressure caused by university enterance exam and the state of Biology course are two of major elements fundamentally shaping CriT and other HOT habits in classrooms (Ozden, 2007; Cimer & Timucin 2010).

In Turkey, Science and Anatolian Secondary Schools are popular with students with higher academic profile. Supposing the state of Biology course instruction is better and factorization would be clearer, one from both types was chosen for the pilot study.

PURPOSE

Within the frame drawn above; the aim of this study is to develop a scale to be applied on students to detect the status of CriT culture of secondary school Biology classrooms. In the reviewed literature, there is a tendency to assess CriT entities of individuals *separately* and to interpret averages of the individual scores to reflect the state of whole group. However, opinions of the individuals about the social process on-going in the classroom are another *shared* dimension. What makes CTLOBICS significant is its focusing this frequently missed dimension, which offers an innovative perspective and data triangulation opportunities. We also believe that this attempt will provoke further research and considerations about students' perception of the CriT aspects in their learning environment.

METHOD

Although this scale was planned to be supported by qualitative data, "Critical Thinking Friendly Biology Classroom Environment Survey" itself is a complete quantitative instrument. Roots of a quantitative study can be said to anchor into positivistic paradigm (Alev, 2003). Survey method fits the requirements of the studies aiming quantitative scales (Cepni, 2001, p.40). CTLOBICS was developed by following the scale development steps determined by McMillan & Schumacher (2010).

Sample

The sample of the study was 387 students attending 9^{th}, 10^{th} and 11^{th} grades at two different secondary schools in Trabzon/Turkey, as students of 4 different biology teachers. Minding the considerations about university entrance examination preparations, 12^{th} grades were not included in the study. The participants were asked to mark the frequency of the given event in the classroom by thinking about their biology course and teacher.

Table 1 summarizes the gender and school year distribution of the students in the sample. Prior to the application, the participant students were informed about the aim of the survey. All of the students were participated voluntarily. The printed form of the survey was applied. The average completion duration of the survey was about 15 minutes.

Table 1. Gender and Educational Level Distribution of the Participants

		Frequency (f)	Percent (%)
Gender	Female	183	47,30
	Male	204	52,70
	Total	387	100
Level (Grade)	9	193	49,87
	10	99	25,58
	11	95	24,55
	Total	**387**	**100**

Step One: Reviewing Literature and Writing Items

After reviewing the literature, setting the goals and theoretical framework (see *Introduction*), Critical Thinking Basic Strategies and explanatory examples (Paul *et.al*., 1990) were translated into Turkish. In order to avoid translation mistakes, Turkish scripts retranslated into English by a language expert and compared with the original text. Then from this source the item pool of the pre-scale was formed with 62 Turkish items scripted from the strategies suitable to be expressed with statements of frequency. The items of the scale were 5-choice Likert type inquiring the frequency of the given teacher/classroom habit. They were graded as; (1) Less than 1 or 2 in a year / (2) 3-4 times in a month / (3) Once in 3-4 class hours / (4) A few times in each class hour / (5) A lot of times in each class hour.

Afterwards the scale was subjected to Validity Procedures, Item Analysis, Exploratory Factor Analysis, and Confirmatory Factor Analysis. Concerning the nature of the related data, the brief explanations and findings of these procedures are presented in *Results* section.

RESULTS

Content and logical validity procedures were applied to check whether the items in the pre-scale fit the defined aims and assessment aims (Fraenkel & Wallen, 2008). For this aim the pre-scale was subjected to the supervision of 3 related field and 1 assessment expert. Along with the feedback, the items of the scale were revised and altered; 3 items were removed for being almost coincident. 59-item pilot scale was formed. All the items were favourable things desired to happen more frequently for the sake of CriT. All the items were positive statements except for item 55 '*Never gets angry when we tell the truth*' but then it excluded in the factor analyses. Then two Turkish Language Education experts checked the pilot scale in terms of language comprehensibility and spelling. Finally, pilot scale was pre-applied to a 9th grade classroom in another school than the sample and the students were interviewed about whether/how they understood the items of the scale. The implementation was conducted as defined in *Sample* part.

Step Two: Item Analysis

Item analysis procedures were applied on the data set obtained from the sample by the 59-item scale. In the first place, item total correlation values were calculated to determine discrimination power index of the items. It was concluded that, item total correlation values varied between 0,41 and 0,69 and the values were greater than the acceptable limits (Buyukozturk, 2009).

Then, skewness, kurtosis, standard deviation and average score values were examined to test the data distribution. It was observed that, standard deviation values varied between 0,84 and 1,36 while average item scores were between 2,29 and 4,35. Skewnees index value were between -1,26 and 0,68 and kurtosis index value were between -1,16 and 1,40 and these values support the normality of the distribution of the data set.

Step Three: Exploratory Factor Analysis

This step tested the structural validity of the scale. The concept of structural validity was defined as; "*the degree of precise measurement of an abstract concept in the context of the behaviour which is to be measured*" (Buyukozturk, 2007). In this study, explanatory factor analysis and basic components analysis were applied to picture the factor structure of the scale (Tabachnick & Fidell, 2007). Varimax Rotation Technique, which is frequently applied and easy to interpret, was used as factor rotating method (Pallant, 2001; Brown, 2006). Prior to the factor analysis, Kaiser-Meyer Olkin (KMO) and Barlett tests were performed to test the compatibility of the data and the sample for basic components analysis (Tabachnick & Fidell, 2007).

KMO coefficient was calculated 0,902 and since it was greater than 0,6 it was concluded that the sample size is sufficient for factor analysis (Hutcheson & Sofroniou, 1999). On the other hand, Barlett test was significant at (p<0,000). That Barlett test was significant at (p<0,001) indicates that the data set is significant for factor analysis and can be factorized (Field, 2005).

Following KMO and Barlett tests basic component analysis was carried out. Factor load values of the items were calculated. While factors were being formed, it was tried to be arranged so as to each item has only one great factor load value in a single factor. If it has two great factor load value in two different factors the difference between them was arranged greater than 0,1 and the factor load of the factor which they belong was arranged greater than 0,4 (Bandalos& Finney, 2010; Field, 2005).

When the factor load values formed by the conducted factor analysis were examined, it was found out that 14 items were coincident and they were excluded from the scale. As a result of the repeated basic component analysis with the remaining 45 items, it was observed that 7 more items had more than one great factor load values and they were excluded, too. In order to define the factor structure of the scale basic component analysis re-applied with

Varimax Rotation Technique on these 32 items. KMO value was found as 0,917 and BTS revealed an Approx. Chi-Square value of 4938,380 with a significance value of 0,000. These values granted that the sample is proper and data set can be factorized. As seen in Table 3, the factor loadings of 32 items on the scale ranged from 0,42 to 0,69. The communalities were acceptable. Components factor loading and varimax factor loading values are presented in Table 2.

The analysis of these 32 items yielded 5 factors with factor load values greater than 1.0 and explaining the 52,26% of the total variance. The variance ratio that is explained by 5 factors (52,26%) is acceptable for social sciences (Scherer, Wiebe, Luther & Adams, 1988). Eigenvalue and variance explained by each factor were; 1st factor 9,913 and %30,98 ; 2nd factor 2,678 and %39,347; 3rd factor 1,482 and %4,63; 4th factor1,459 and %4,559; 5th factor 1,193 and %3,729 respectively. When the values obtained by Varimax Rotation was reviewed, it was observed that rotated factor load values of the items varied 0,43 – 0,73. The related values were within the acceptable limits (Field, 2005). Table 3 presents variance ratios explained by each factor and eigenvalues before and after the Varimax Rotation:

Table 2. Factors and Factor Loading Values of the Items

Factors & Items	**Factor Loadings**	**Varimax Factor Loadings**	**Factors & Items**	**Factor Loadings**	**Varimax Factor Loadings**
Factor 1 Affective Critical Thinking Skills			Factor 2 General Thinking Skills		
Item28	,601	,725	Item15	,593	,737
Item23	,590	,717	Item17	,548	,714
Item22	,653	,685	Item24	,553	,689
Item29	,576	,678	Item16	,537	,612
Item31	,557	,673	Item18	,631	,561
Item32	,526	,672	Item19	,624	,430
Item30	,613	,657	Factor 3 Basic Critical Thinking Skills		
Item24	,553	,641	Item3	,461	,696
Item25	,667	,632	Item4	,536	,635
Item20	,640	,627	Item8	,485	,596
Item27	,553	,614	Item2	,392	,533
Item26	,616	,556	Item5	,540	,519
Item21	,672	,551	Item1	,506	,512
Factor 4 Associating with Daily Life			Factor 5 Reasoning Skills		
Item7	,487	,722	Item12	,559	,681
Item6	,506	,669	Item13	,518	,578
Item9	,599	,668	Item11	,529	,484
Item10	,525	,602			

Table 3. Principal Component Factoring Analysis: Total Variance Explained

	Extraction Sums of Squared Loadings		Rotation Sums of Squared Loadings	
Factor	Eigen Value	Percentage of Variance (%)	Eigen Value	Percentage of Variance (%)
1	9,913	30,980	5,957	18,614
2	2,678	8,368	3,747	11,711
3	1,482	4,630	2,243	7,009
4	1,459	4,559	2,203	6,883
5	1,193	3,729	1,824	5,701
Cumulative %		55,579		55,579

After the explanatory factor analysis 27 items were excluded from 59-item pilot scale and a 32-item, 5-factor scale was obtained (*see CTLOBICS available as supplementary material accompanying the online article*). The first factor (*Affective Critical Thinking Skills*) had 13 items, the second (*General Thinking Skills*) and the third (*Basic Critical Thinking Skills*) had 6 items, the forth (*Associating with Real Life*) and the last (*Reasoning Skills*) had 3 items. The factor names were given concerning the factor content and the related literature.

Step Four: Confirmatory Factor Analysis

Following to the explanatory factor analysis, confirmatory factor analysis (CFA) was applied to test the suitability of formed structure (model) to the data and structural validity of the factors (Maruyama, 1998; Kline, 2005). "*The purpose of CFA is to identify latent factors that account for the variation and co-variation among a set of indicators.*" (Brown. p.40). In this study, Chi-Square (χ^2), χ^2/degree of freedom, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Non-Normed Fit Index (NNFI) and Comparative Fit Index (CFI) were referred as fit indexes. After CFA, the mentioned fit index values were formed as; χ^2= 995,80; χ^2/df=2,37; RMSEA=0,060; SRMR=0,06; GFI=0,86; AGFI=0,83; NNFI=0,96; CFI=0,96. Since χ^2 value is not significant, χ^2/df ratio is below 3, RMSEA value is smaller than 0.08, SRMR index value smaller than 0.08, GFI index value greater than 0,90, AGFI index value greater than 0,80 and NNFI and CFI index values are greater than 0,90, the produced model can be said highly fit (Brown, 2006; Klem, 2000; Kline, 2005; McDonald & Ho, 2006). When fit index values after CFA were evaluated, it was concluded that the 32-item scale is fit and applicable without any modification.

Step Five: Reliability Analysis

Reliability concept; explains the consistency of the measurement process (Fraenkel & Wallen, 2008). Internal Consistency Reliability was applied to test the consistency among the individual items in the scale as a measure of reliability (Scott & Morrison, 2006). The reliability of the 32-item scale was expressed with Cronbach's Alpha (α) correlation coefficient as 0,92.

The individual reliability coefficients of the factors were determined as; *1st Affective Critical Thinking Skills*: 0,89, *2nd General Thinking Skills*: 0,77, *3rd* Basic Critical Thinking Skills: 0,74, *4th Associating with Real Life*: 0,72, *5th Reasoning Skills*: 0,69. It was obvious that overall scale and individual factor reliability coefficients were very close to reliability range 0,70 and above the ideal reliability limit (Creswell, 2005). On the other hand item total correlation values were calculated to determine discrimination power of the items in the scale. They varied between 0,31 and 0,63. These values imply that discrimination powers of the items in the scale were in acceptable limits (Buyukozturk *et.al.*, 2010). This step finalized the 32-item, 5-factor scale.

The final form of the survey was translated from Turkish into English by two different experts and the translations were compared. In the end, the English survey form translated back into Turkish and it was compared with the original Turkish text and the survey form was reached. The parts in *italic* were items excluded from the survey after consecutive factor analyses (*see CTLOBICS available as supplementary material accompanying the online article*).

In order to obtain feedback about the content validity of the survey and a kind of informal consent from the founder of the initial theoretical framework of the study, the survey was sent to a field expert (Dr. Enoch Hale, a fellow of the CTC and co-worker of Dr. Richard Paul, an author of Paul *et al.* (1990), personal communication *via* email, September 22, 2011). He offered some alterations like specifying words towards comprehensibility and critical thinking terminology for items 1, 3, 4, 5, 6, 9, 10, 12, 13, 20, 22, 25, 29, 32, 33, 37, 40, 41, and 54. Concerning statistical requirements, possible changes were reflected on the survey form. One comment by him needed splitting an item so it could not be reflected. He thought that it may be more useful if item 1 was broken into two as: *"#1 Grounds content concepts, principles and processes with examples. #2 Challenges us to extrapolate general rules from the examples presented in class"*

Discussion & Conclusion

Setting sail from the point of supporting HOTS and assessing them with proper tools is vital part of the contemporary teaching and learning theories, we end up with a survey having 32 items and 5 factors with the inner consistency coefficient α=0,92. Although there are variety of scales to detect different HOT components, CTLOBICS seems to be unique with measuring the state of CriT holistically by referencing the student opinions about the frequency of CriT inducing events in classroom environment. This environment is shaped not only by teachers but also by the interaction between teachers and students. CTLOBICS focuses on not individualized talent, but the total social effect. This perspective, we believe, will yield another layer of quantitative data and pave way for further researches into this field.

As well as teachers and school managers may use CTLOBICS to evaluate classroom environments, it can also be used in variety of study areas particularly scientific studies conducted in contemporary teaching/learning approaches based classroom environments prioritizing HOT. CTLOBICS can gather data to constitute quantitative legs of such groups of existing studies (Cimer & Timucin, 2009):

- Determining CriT skill level of various learner and teacher groups,
- Developing or improving educational components like learning environments, instructional materials *etc*. by concerning CriT and/or other HOTS
- Investigating effects of educational innovations on CriT features of learners/teachers,
- Evaluating or developing curricula with respect to CriT or HOT elements.

CTLOBICS can be scored in a way to have 1 for the least frequent, 2 for the next and so on, and 5 for the most frequent choices for each item. Therefore all-32 items will yield maximum total score of 160. Classroom averages can be used to have an idea about the CriT state of the Biology course for that classroom. In further studies, all over survey scores of CTLOBICS may be compared to sub-scale scores and their relations can give further ideas about the CriT state of the classroom. As there are more studies to be referenced, the factor scores and the overall score of the survey can be classified as high, medium, and low. If there are no results to compare, the median value of the applications can be used to manage such classification.

To a large extent, factorization of the survey was in an expected way. Based on CriT literature and definitions, the factors were named as; *Affective Critical Thinking Skills, General Thinking Skills, Basic Critical Thinking Skills, Associating with Real Life, Reasoning Skills*. It was realized that the clarity and comprehensibility of the items had an important part in healthy factor formation and eventually the existence of the items in the final form. Therefore, it must be focused on preparing items with single statement and with clear, single meaning when similar scales are being developed or the scale is being translated into other languages. Still, the researchers or teachers to apply this survey should reconsider the comprehensibility of the survey concerning the level of their students.

For availability reasons this study was conducted with 4 teachers teaching 9th, 10th and 11th grades, the final form of CTLOBICS can be applied or re-piloted in different school types or with more homogenous samples. Moreover, the survey can be applied cross-culturally in different languages and the results may be compared. The effect of gender was not monitored in this study; further studies may inquire about the probable impact of the gender factor.

CTLOBICS has no direct reference to 'Biology Course' in the body text so pilot studies for different courses may be designed. Backed with literature and the observations conducted within the other steps of the present study (Cimer & Timucin 2010), it can be argued that adding items or even factors concerning student- student interaction and humour domains may positively affect content validity of further studies. For future improvements of similar scales it can be recommended that; instead of 'slicing out' CriT Skills, holistic scales sensitive to all types of HOTS in the classroom environment may be prepared containing factors or sub-factors as; Critical Thinking, Creative Thinking, Reflective Thinking and Problem Solving. Last but not least, customized assessment of thinking skills is generally more valid (Stein, 2003). So, quantitative instruments should always be combined with qualitative complements, preferably rubrics prepared for the unique cases of teachers or researchers.

REFERENCES

Alev, N. (2003). Integrating Information and Communications Technology (ICT) into Pre-service Science Teacher Education: The Challenges of Change in a Turkish Faculty of Education. Unpublished Doctoral dissertation, University of Leicester, School of Education, UK.

Allen J., (2003). Fostering Ethical Competence in Nursing Education. Clinical Research and Regulatory Affairs, 20(4), 373-377.

Azar, A. (2010). The effect of critical thinking dispositions on students achievement in selection and placement exam for university in Turkey. Journal of Turkish Science Education, 7(1), 61-73.

Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), The reviewer's guide to quantitative methods in the social sciences (pp. 93-114). New York: Routledge

Blake, B., & Pope, T. (2008). Developmental psychology: Incorporating Piaget's and Vygotsky's theories in classrooms. Journal of Cross-Disciplinary Perspectives in Education, 1(1), 59-67.

Bloom, B., Englehart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive Domain. New York: Longmans Green.

Brooks, J.G., & Brooks, M.G. (1993). In search of understanding: The case for constructivist classrooms. Alexandria, VA: Association of Supervision and Curriculum Development.

Buyukozturk, S. (2009). Sosyal Bilimler icin Veri Analizi El Kitabı (10th Ed.). Ankara: PegemA Yayıncılık.

Creswell, J. W. (2005). Educational research: Planning, conducting and evaluating quantitative and qualitative research. New Jersey: Pearson Prentice Hall.

Cepni, S. (2001). Arastırma ve Proje Çalışmalarına Giriş. Trabzon: Erol Ofset.

Cimer, S. (2004). An Investigation into Biology Teachers' Perceptions of Classroom Assessment in Secondary Schools in Turkiye, Unpublished Doctoral dissertation, The University of Nottingham, UK.

Cimer, S., & Timucin, M. (2008). Formative assessment perceptions and habits of primary school english teachers: Trabzon Case, VII. International Educational Technology Conference (IETC-2008), Eskisehir, Turkey.

Cimer, A., & Timucin, M. (2009). Critical Thinking Theses in Turkey. Paper presented in 3rd International Computer and Instructional Technologies Symposium (ICITS 2009), Trabzon, Turkey, p.173.

Cimer, A., & Timucin, M., Content of an in-service training to develop and assess activities minding critical thinking. World Conference on Teaching Learning and Administration, The American University Conference and Visitor Center, Cairo, Egypt (29-31 October 2010).

Cimer, A., & Timucin, M. (2010). Content of an in-service training to develop and assess activities minding critical thinking. Procedia Social and Behavioral Sciences, 9, 958–962.

De Bono, E. (1995). Mind Power. New York:Dorling Kindersley.

Dewey,J., (1982). How we think. Lexington, MA: Heath, (Originally published in 1910).

Dolapcı. O.C., (2009). The evaluation of the workshop about informing the teachers on the subject of critical thinking. (Unpublished master's thesis) Institute of Social Sciences. Yeditepe University. Istanbul, Turkey.

Duron, R., Limbach, B., & Waugh, W. (2006). Critical thinking framework for any discipline. International Journal of Teaching and Learning in Higher Education, 17(2), 160-166.

Ennis, R. H. (1993). Critical Thinking Assessment. College of Education, 32 (3), 179-184.

Facione, P.A., & Facione, N.C. (1994). The California Critical Thinking Skills and Test Manual. Millbrae, CA: California Academic Press.

Field, A. (2005). Discovering statistics using SPSS. London: Sage.

Greenockle, K.M., & Purvis G.J. (1995). Redesigning a secondary school wellness unit using the critical thinking model. The Journal of Physical Education, Recreation & Dance, 66(6), 49-52.

Harrigan, A., & Vincenti. V. (2004). Developing higher-order thinking through an intercultural assignment: A scholarship of teaching inquiry project. College Teaching 52(2), 113-120.

Ho, R. (2006). Handbook of univariate and multivariate data analysis and interpretation with SPSS. Florida: Chapman & Hall/CRC.

Huitt, W. (1998). Critical thinking: An overview. Educational Psychology Interactive. Valdosta, GA: Valdosta State University. Retrieved 19 April, 2011, from, http://www.edpsycinteractive.org/topics/cogsys/critthnk.html

Icbay, M. A. (2005). A SWOT Analysis on the university entrance examination in Turkey: A case study. Mersin Universitesi Eğitim Fakultesi Dergisi. 1(1), 126–140.

Insight assessment. (2011). A Division of California Academic Press. Retrieved 10 December, 2011, from http://www.insightassessment.com/Products

Jensen, E.P. (2008). A fresh look at brain-based education. (cover story). Phi Delta Kappan, 89(6), 408-417.

Keles, E. & Cepni, S. (2006). Brain and Learning. Journal of Turkish Science Education, 3(6), 66-82.

King, F.J., Goodson, L., & Rohani, F. (n.d.) Higher Order Thinking Skills: Definition, Teaching Strategies, Assessment. Report by Center for Advancement of Learning and Assessment. Retrieved September 15, 2011, from www.cala.fsu.edu/files/higher_order_thinking_skills.pdf

Kirsch, I. (2001). The International adult literacy survey (IALS): Understanding what was measured, educational testing service research report RR-01-25, Princeton, New Jersey.

Klem, L. (2000). Structural equation modeling. In L. Grimm & P. Yarnold (Eds.), Reading and understanding multivariate statistics, Vol. II. Washington, DC: American Psychological Association.

Kline, R.B. (2005). Principles and Practice of Structural Equation Modeling (2nd ed.). New York: The Guilford Press.

Kuhn D. (1999). A developmental model of critical thinking. Educational Researcher, 28(2), 16- 46.

Krathwohl, D. R, & Anderson, L.W. (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Allyn & Bacon. Boston, MA

Lewis, A., & Smith, D. (1993). Defining higher order thinking. Theory into Practice, 32(3), 131−137.

McDonald, R. P., & Ho, M-H. R. (2002). Principles and practice in reporting structural equation analyses. Psychological Methods, 7, 64 – 82.

MoNET (2007). Directorate of Student Discipline Board, Secondary Education 9th Grade Biology Course Curriculum, Ankara.

MoNET (2008a). Directorate of Student Discipline Board, Secondary Education 10th Grade Biology Course Curriculum, Ankara.

MoNET (2008b). Directorate of Student Discipline Board, Secondary Education 11th Grade Biology Course Curriculum, Ankara.

MoNET (2009). Directorate of Student Discipline Board, Secondary Education 12th Grade Biology Course Curriculum, Ankara.

MoNET (2010). Oğretim Programlarının Yenilenmesini Zorunlu Kılan Nedenler. MEB (The Reasons Obliging the Renewal of the Curricula). Retrieved June 21, 2010, from http://ttkb.meb.gov.tr/programlar/prog_giris/prog_giris_6.html

Mulnix, J. W. (2012). Thinking Critically about Critical Thinking. Educational Philosophy and Theory, 44(5), 464-479. doi: 10.1111/j.1469-5812.2010.00673.x

OECD (Organisation for Economic Co-operation and Development), Directorate for Education (2005). Definition and Selection of Key Competencies: Executive Summary. Retrieved from: http://www.oecd.org/dataoecd/47/61/35070367.pdf (July 5, 2011)

Ozden, M. (2007) Problems with science and technology education in Turkey. Eurasia Journal of Mathematics, Science & Technology Education, 3(2), 157-161.

Paul, W.R., Binker., A., Jensen, K., & Kreklau, H. (1990). Critical Thinking Handbook: A guide for remodeling lesson plans in language arts, social studies and science. Rohnert Park, CA: Foundation for Critical Thinking.

Paul, R., & Elder, L. (2002). Critical Thinking: Tools for Taking Charge of Your Professional and Personal Life. New Jersey: Financial Times Prentice Hall

Paul, W.R. (2005). The state of critical thinking today. New Directions for Community Colleges, 130, 27-38.

Radin, J.L. (2009). Brain-compatible teaching and learning: Implications for teacher education. Educational Horizons, 88(1), 40-50.

Ross. J., & Ross, C. (1976). Ross test of higher cognitive processes manual. Los Angeles, CA: Western Psychological Services.

Scott, D., & Morrison, M. (2006). Key ideas in educational research. London: Continuum International Publishing Group.

Stein, B.S., Haynes, A.F., & Unterstein, J., (2003). Assessing Critical Thinking Skills. Contribution to SACS/COC Annual Meeting / Nashville, Tennessee / December 6 - 9, 2003.

Tabachnick, B.G., & Fidell, L.S. (2007). Using Multivariate Statistics (5th ed.). Pearson Education, Inc. / Allyn and Bacon.

Taylor, P.C., & Fraser, B.J. (1991, April). Development of an instrument for assessing constructivist learning environments. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, L.A.

The Critical Thinking Community (2011). The Critical Thinking Community Foundation for Critical Thinking. Retrieved December 10, 2011, from http://www.criticalthinking.org/pages/dr-richard-paul/818

The Critical Thinking Community-Strategy List. (2011). The Critical Thinking Community Foundation for Critical Thinking. Retrieved December 10, 2011, from http://www.criticalthinking.org/pages/strategy-list-35-dimensions-of-critical-hought/466

Torrance, E.P. (1966). Torrance Tests of Creative Thinking: Norms technical manual. Princeton NJ: Personnel Press.

Tileston, D. (2005). 10 best teaching practices (2nd ed.). Thousand Oaks, CA: Corwin Press. Retrieved September 20, 2011, from Google Books, http://books.google.gm/books/about/10_best_teaching_practices.html

Timucin, M. (2008). Turkish Primary School Teachers' Perceptions About Formative Assessment, XIII World Congress Of Comparative Education Societies Living Together: Education And Intercultural Dialogue Program, Sarajevo, Bosnia-Herzegovina

Timucin, M. (2009). History of Critical Thinking. Paper presented in 3rd International Computer and Instructional Technologies Symposium (ICITS 2009), Trabzon, Turkey.

Walsh, D., & Paul, R. (1988). The goal of critical thinking: From educational ideal to educational reality. Washington, D.C.: American Federation of Teachers.

Watson, G., & Glaser, E. M. (1980). Watson-Glaser Critical Thinking Appraisal. San Antonio, TX: Psychological Corp.

Willingham, D.T. (2007). Critical thinking: Why is it so hard to teach? American Educator, 31(2), 8-19.

Zohar, A., & Dori, Y.J. (2003). Higher order thinking skills and low achieving students: Are they mutually exclusive? The Journal of the Learning Sciences, 12, 145-182.