

# Reflective Thinking Scale: A Validity and Reliability Study\*

GÜLŞAH BAŞOL<sup>a</sup>

Gaziosmanpaşa University

İLKE EVİN GENÇEL<sup>b</sup>

Çanakkale Onsekiz Mart University

## Abstract

The purpose of this study was to adapt Reflective Thinking Scale to Turkish and investigate its validity and reliability over a Turkish university students' sample. Reflective Thinking Scale (RTS) is a 5 point Likert scale (ranging from 1 corresponding Agree Completely, 3 to Neutral, and 5 to Not Agree Completely), purposed to measure reflective thinking in habitual actions, understanding, reflecting and critical reflecting subscale levels. With the help of education area and language experts, the scale adaptation thoroughly completed and the correlation between the original scale and its Turkish form indicated satisfactory language equality ( $r=.80$ ). During the 2011-2012 school years, 1413 undergraduate students, from varying departments of Faculty of Education, Çanakkale Onsekiz Mart and Gaziosmanpaşa University, were participated in this study. In order to maximize student participation, convenient sampling method was followed. After the item analysis, internal consistency of the scale and subscale items were checked through Cronbach Alpha internal consistency coefficient and Spearman Brown Split-Half correlation coefficient and both indicated good internal consistency on the scale and the subscale levels. The scale also demonstrated sufficient test-retest reliability over the application of the scale in two weeks and two months intervals. The construct validity is established through a confirmatory and an exploratory factor analyses and the four factor structure, suggested in the original study, was confirmed by both analyses. The concurrent validity was also confirmed through the scale's satisfactory Pearson correlation coefficient with Critical Thinking Scale.

## Key Words

Reflective Thinking, Critical Thinking, Reflective Thinking Scale, Validity, Reliability

Producing reflective thinkers is one of the main goals of a constructivist approach to education and also one of the primary objectives of Dewey's pragmatism. Dewey (1933) described reflective thinking as "active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusion to which it tends". He also stated that reflective thinking adds meaning to the experience through its

reorganization and reconstruction and leads the way for further goals requiring more comprehensive tasks. Reflective thinking requires the learner to achieve the learning goals and turn into the behavior. Later, learner evaluates his/her own performance and has a perception of the progress he/she made, makes a discussion about the product, this way it will be possible to benefit from that experience for the future occurrences. A deed can be counted as successful

\* This study was presented in the Third National Educational Measurement and Evaluation Congress between September 19 and 21 in Bolu Abant İzzet Baysal University.

**a** Gülşah BAŞOL, Ph.D., is an associate professor in educational measurement and evaluation area. Her study interests are focused on meta analysis, methodological evaluation, psychometric evaluation of measurement instruments focused on instrument adaptation and development studies. She has a special interest in attitudes toward women, teacher attitudes toward multiculturalism, burnout levels in school principals and teachers and its relation with social support. Correspondence: Gaziosmanpaşa University, Faculty of Education, Department of Educational Sciences, 60100, Tokat, Turkey. E-mail: gulsah.basol@gop.edu.tr, basol\_g@hotmail.com Phone: +90 356 252 1616/3412.

**b** İlke EVİN GENÇEL, Ph.D., is an assistant professor in curriculum and instruction area. Contact: Çanakkale Onsekiz Mart University, Faculty of Education, Department of Educational Sciences, Anafartalar Yerleşkesi, 17100 Çanakkale, Turkey. E-mail: ilkeevincel@comu.edu.tr, ilke.evin@gmail.com

to the extent that its doer reflects his/her education towards life experiences. For this purpose, a teacher needs to promote students through helping them with homework assignments and demonstrations. Pollard (2002) emphasized that teacher reflection has become increasingly important for classroom practitioners, and both Tripp (2003) and Pollard stated that reflective thinking is a cyclic process. According to Ennis (1987), reflective thinking is a component of the critical thinking process. Tan and Goh (2008) stated that reflection links the past, the present, and/or the future.

Kember and his colleagues (Kember et al., 2000; Kember, McKay, Sinclair, & Frances, 2008) explained reflective thinking in four dimensions as Habitual Action, Understanding, Reflection, and Critical Reflection. They also concluded that this classification can be used to evaluate the appropriateness of students' writings, compositions, dairies, and their answers to open ended questions to reflective thinking.

**Habitual Actions:** These are the actions, learned before and mastered through frequent use and started to be performed automatically. Using a keyboard, riding a bike, driving a car can be given as examples of this level.

**Understanding:** According to Mezirow (1991), thoughtful action makes use of the knowledge, without attempting to appraise that knowledge (cited in Kember et al., 2000). Our learning that has taken place in schools; Mezirow's "book learning" is classified under this category. Although first Kember et al. (2000) attempted including all categories of Blooms taxonomy in the Understanding sub scale of the questionnaire, later they realized that the scale had poor psychometric properties; therefore they have decided to focus on Understanding or Comprehension in order to increase reliability and validity through concentrating on a narrower construct.

**Reflection:** There are many definitions of reflective learning and how it should be. Kember et al. (2000) based their article on Mezirow's definition. They were also influenced by Boyd and Fales (1983) and Boud, Keogh, and Walker (1985) who emphasize the importance of experiences in the reflection process. As cited by Kember et al. (2000), Mezirow describes reflection as validity testing.

Reflection involves the critique of assumptions about the content or process of problem solving. The critique of premise or presupposition pertains to problem *posing* as distinct from problem *solving*. Problem posing involves making a taken-for-granted situation problematic, raising questions regarding its validity (Mezirow, 1991, p. 105).

According to Boyd and Fales, reflective learning is the process of internally examining and exploring an issue of concern, triggered by an experience, which creates and clarifies meaning in terms of self, in a changed conceptual perspective (cited in Kember et al., 2000).

**Critical Reflection:** This is a higher level reflective thinking. Mezirow calls it "premise reflection". In this level, learners become aware of why he/she perceive, think, feel, or act (cited in Kember et al., 2000). Kember et al. (2000) stated that it would not be easy to transform perspectives. Dewey (1933) also underlines the importance of critical thinking by suggesting not reaching a hasty conclusion without examining all the possibilities. Reflective thinking is a much discussed construct and there are not many scales measuring. However, critical thinking is regarded as a close call to reflective thinking. Reflective Thinking Scale is developed by Kember and his colleagues in 2000. There is also a scale developed by Semerci (2007) purposed to measure teacher candidates' reflective thinking inclination.

Although it was developed not long before, Reflective Thinking Scale has been used cross culturally. Leung and Kember (2003) used the scale on health class students in Hong Kong University and found that students who tend to learn on surface, scored higher in Habitual Action level, while deep learners are more likely to be in Understanding, Reflection or Critical Reflection Levels. According to Leung and Kember this finding can be taken as a sign of convergent validity of the scale. Pahn (2007) also confirmed these findings as a result of his study, searching reflective thinking in comparison with self-efficacy beliefs of students. Lie (2007) also found that students who are in the first year of their education based on problem solving methods had lower critical thinking levels compared to other students in upper grades. There are also findings of a quasi-experimental study indicating higher scores in the experimental group for Understanding and Critical Thinking Sub-scales and Higher scores for the control group in Habitual Action Sub-scale (Mahardale, Neville, Jais, & Chan, 2008). In addition, there are some researches about reflective thinking and peer education (Loke & Chow, 2003), and learning approaches (Leung & Kember). Besides Lucas and Tan (2006) emphasized that Reflective Thinking Scale is valid on social sciences field students. Dolapçioğlu (2007) and Karadağ (2010) stated that reflective thinking levels of the Turkish teachers were found high. Erginel-Şanal (2006) and Köksal and Demirel (2008) found that reflective thinking education has positive effects on the reflective thinking level.

Literature review indicated that there was not any attempt to adapt this scale into Turkish. Therefore, it would be a good addition to the literature, providing researchers an instrument to measure reflective thinking in order to increase students' achievement.

## Method

### Model

This study was based on a descriptive survey method. The reliability and validity of Reflective Thinking Scale were studied from many aspects.

### Study Sample and Application

Since the main purpose of the current study was to search construct validity of Reflective Thinking Scale over students, taking different classes by different instructors, it was decided to carry on a big scale study. The Sample of the study consisted of 1413 university students from the Faculty of Education of Çanakkale Onsekiz Mart University (n=968) and Gaziosmanpaşa University (n=445). The sample drawn by convenient sampling consists of students, whose ages were between 18 and 37, coming from different class level. 70% of students were girls and 30% were boys.

In order to prove scale's language equity, original form and its Turkish form applied to a group of 151 3rd grade students in an English Teacher Education Program, with a one week interval. Test-retest reliability was also tried on a 147 Teacher Candidate sample in the Faculty of Education of Gaziosmanpaşa University.

### Measurement Instruments

**Reflective Thinking Scale:** It is a 16 item 5-point Likert scale ranging from 5 Definitely Agree, 4 Agree with reservation, 3 only to be used if a Definite answer is not possible, 2 Disagree with reservation, to 1 Definitely disagree. The scale has four subscales which are Habitual Action, Understanding, Reflection, and Critical Reflection. The lowest possible score that can be earned from each sub scale is 4 and the highest is 20. The scale is very easy to understand and it does not take more than 10 minutes to complete. The original validity study was carried out in Hong Kong University in 2000 on 303 university student sample. The Cronbach Alpha coefficients for the subscales ranged between .62-.76. A confirmatory factor analysis affirmed the four factor model for reflective thinking.

The literature review indicated that there is a need

for adapting the Reflective Thinking Scale into Turkish since it is measuring an important feature of constructivist approach. The current study was intended to fill this gap in the literature.

### The Scale Adaptation into Turkish

In the process of adaptation, the scale first translated into Turkish by the authors of the study. The language equity of the scale was confirmed by three area experts who are fluent in both languages. The Turkish form of the scale was improved accordingly to these feedbacks. The final form was back translated by two instructors from the English Department. For delivering handouts before exams by instructors is not a common application by Turkish instructors, it was decided to re-write Item 9 as "As long as I study for the exams and remember what we learned in class, there is no need for me to think too much". Also, Item 14 is also rewritten to imply understanding to be in line with the other three items in the Understanding sub scale. Later, the form was read by a group of graduate students in order to evaluate its appropriateness for undergraduate students. The final form was re-evaluated by two area experts and two philosophers to decide its equality to its original form after the final improvements. After a careful review by a Turkish language expert, the scale adaptation process was completed. A pilot study of the scale on a group of 15 students indicated that all students completed the scale in 10 minutes most. The only problem with the scale, during the application process was that a few students asked what was meant by "this course" in the items since they were taking several courses. The administrator explained that it meant the course they were having a class meeting right at that moment. It was decided that during the application process students were better to be reminded what was meant by "this course" in the items. This will not be a problem when the scale is applied to measure students' reflection levels in a certain course.

### The California Critical Thinking Disposition Inventory

The California Critical Thinking Disposition Inventory (CCTDI) was developed as a part of American Philosophy Association's Delphi Project in order to measure critical thinking intentions by Facione and Facione (1992). The scale adapted to Turkish by Kökdemir in 2003 has 51 items and six subscales as analyticity, self confidence, inquisitiveness, maturity, open-mindedness, being systematical, and truth seek-

ing. Cronbach Alpha internal consistency for the scale is .88 and for the sub-scales it ranges from .61 to .78.

### Analysis

After a routine check for unusual data points and normality check for each item, the factorial structure of the scale was checked by an exploratory factor analysis through SPSS and a confirmatory factor analysis was performed by Lisrel. Item discrimination levels were searched through item-total correlation coefficients and upper and lower group comparisons through independent group *t* tests. Reliability checks were accomplished by Cronbach Alpha internal consistency coefficient, Spearman Brown Split Half reliability, and test-retest correlation coefficients.

### Results

Turkish and English versions of the scale were applied in two week intervals to 151 English Department 3rd grade students. It was seen that the correlations ranged from .74 to .92. Therefore, it was decided that the scale had adequate language equality.

### Item Analysis

Factor loadings are all positive and over .50. A negative factor loading means that the item has a negative relationship with the rest of the items. There were no reverse coded items in the scale and no negative loadings indicating all items have positive relationships with each other. Communalities ranged between .36 and .59. High communalities indicate that the item has power in explaining the variation in data. Corrected item-total correlations are found between .21 and .50 which is above the cut point of .20. Also, upper and lower group comparison indicated that all items had discriminating power.

### Reliability

Reliability of the scale was attained through the test-retest reliability, Cronbach Alpha coefficient and split half correlations. Overall total scale scores, Understanding, Reflective Thinking and Critical Reflection Subscales had sufficient reliability. Results indicated low reliability with Habitual Learning Subscale. Considering that the scale has four items under each sub scale level, this can be understandable.

Correlations between the subscales and the total score were higher than the correlations among the subscales. The low correlations among the subscales indicated that the scale has more than one dimensions (subscales). Considerably high correlations between the total score and subscales can be tak-

en as a sign that the subscales indeed measure the same domain altogether.

Significant mean differences between the subscales can also indicate that subscales measure different domains. All of the pair-wise comparisons between the subscales were significant; therefore the subscales measure different parts of reflective thinking.

### Validity

Construct validity of the scale is searched through an explanatory and confirmatory factor analysis. In addition, criterion validity of the scale is established through its correlation with California Critical Thinking Scale.

**Exploratory Factor Analysis:** Kaiser Mayer Olkin Measure found .83 indicating that sample was adequate to run a factor analysis. Bartlett's Chi Square test (2352.11,  $p = .001$ ) was significant, indicating that data was appropriate to draw factors. Principal component analysis was run with a Promax rotation. According to Tabachnick and Fidel (1996) in a principal component analysis when the sub-scales have high correlations, one of the orthogonal rotations should be used. According to Kaiser Rule (Kaiser, 1960) and Cattell's Scree Plot (Cattell, 1966) there were four factors. According to Kline (1994) 40% of the total variation explained by factors, can be taken as the minimum acceptable explained variance level. Sixteen items together explained 53% of variation in data. The same items appeared under the same headings as in the original study. In order to accept an item under a dimension, factor loading value .40 and above were taken as a cut off value (Stevens, 1996). Item loadings were all positive and ranged between .76 at highest and .54 at lowest. Communalities were between .41 and .63. Reflection had the highest eigenvalue with 3.78 and explained 24% of the variance. Reflection followed by Critical Reflection with an eigenvalue of 2.11 and 13% of the variation. There was one item (Item 1) with a .21 item-total correlation level. According to Şencan (2005), with a sample size over 400 cases, .20 and above item total correlation values can be regarded enough. Ebel (1965) also suggests that an item with an item-total correlation value between .20 and .30 can be used if it does not have an alternative and when it is necessary (cited in Turgut, 1990, p. 270).

**Confirmatory Factor Analysis:** Through Lisrel 8.54 (Jöreskog & Sörbom, 2003), it was searched whether the data support the advocated four factor model for the scale, suggested by its developers. Proposed four factor model compared to one factor

model. Since model-data suitability was affected by sample size,  $X^2/df$  division is used. Any value below five indicates a good fit. Goodness of fit index (GFI), Normalized Standard Fit Index (NCFI), Comparative Fit Index (CFI), Normative Comparative Fit Index (NCFI), Adjusted Goodness of Fit Index (AGFI), and Root Mean Square Error Adjusted (RMSEA) were used to decide which model is a better fit to the data. According to Anderson and Gerbing (1984) when NNFI and CFI values are close to 1, this indicates a good fit, this is true also for GFI and AGFI indexes. RMSEA value is expected to be less than .08 (Byrne, 2001; Hu & Bentler, 1998), NNFI, CFI, GFI and AGFI values are also expected to be close to 1 for a good fit.

According to table 1, it was seen that single factor model has the poorest properties as it was in the original study. Four-factor model indicated a good fit with the  $X^2/df$  value less than 5 and all the fit indexes close to 1.

**Table 1.**  
Fit Indexes for Different Models

Model	$X^2/df$	GFI	AGFI	NNFI	CFI	RMSEA
1. Single Factor Null Model	10.49(104)	.80	.74	.77	.80	.13
2. Four Factor Model	4.48 (98)	.93	.90	.92	.93	.07*

*p*<.001\*\*

### Concurrent Validity of the Scale

Concurrent validity of the scale was searched through its correlation with California Critical Thinking Scale. Results indicated that total scores for two scales had a satisfactory correlation.

### Results and Discussion

Reflective Thinking Scale's Turkish Form confirmed the four factor structure of the original scale. EFA and CFA results both confirmed the construct validity of the scale. The low correlations among the subscales indicate that subscales do not overlap. The item analysis indicated that all items were psychometrically sound. Cronbach Alpha and Spearman Brown Split-Half correlation coefficients indicated that the scale had good internal consistencies both in the scale and subscale level, all of them near or above the critical level of .70, set by Nunnally and Bernstein (1994). The scale also demonstrated sufficient test-retest reliability over the application of the scale in two weeks and two month intervals. The scales' concurrent validity was also confirmed through its satisfactory correlation with Critical Thinking Scale (Kökdemir, 2003).

### References/Kaynakça

- Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49, 155-173.
- Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*. London: Kogan Page.
- Boyd, E. M., & Fales, A. W. (1983). Reflective learning: Key to learning experience. *Journal of Humanistic Psychology*, 23 (2), 99-117.
- Byrne, B. M. (2001). *Structural equation modeling with Amos: Basic concepts, applications and programming*. Hillsdale, NJ: Erlbaum.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 629-637.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: D. D. Heath.
- Dolaçoğlu, S. D. (2007). *Sınıf öğretmenlerinin yansıtıcı düşünme düzeylerinin değerlendirilmesi*. Yayınlanmamış yüksek lisans tezi, Mustafa Kemal Üniversitesi, Sosyal Bilimler Enstitüsü, Hatay.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Steinberg (Eds.) *Teaching thinking skills: Theory and practice* (pp. 9-26). New York: Freeman.
- Erginel-Şanal, S. (2006). Developing reflective teachers: A study on perception and improvement of reflection in pre-service teacher education. Yayınlanmamış doktora tezi, Orta Doğu Teknik Üniversitesi, Ankara.
- Facione, P. A., & N. Facione. (1992). *The California Critical Thinking Disposition Inventory (CCTDI)*. Millbrae, CA, California Academic Press.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to under parameterized model misspecification. *Psychological Methods*, 3, 424-453.
- Jöreskog, K., & Sörbom, D. (2003). *LISREL VIII: User's reference guide*. Chicago, IL: SSI.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Karadağ, M. (2010). *Sosyal bilgiler öğretmenlerinin yansıtıcı düşünme düzeylerinin incelenmesi (Şanlıurfa ili örneği)*. Yayınlanmamış yüksek lisans tezi, Çukurova Üniversitesi, Sosyal Bilimler Enstitüsü, Adana.
- Kember, D., Leung, D. Y. P., Jones, A., Loke, A. Y., McKay, J., Sinclair, K. et al. (2000). Development of a questionnaire to measure the level of reflective thinking. *Assessment & Evaluation in Higher Education*, 25 (4), 381-395.
- Kember, D., McKay, J., Sinclair, K., & Frances, Y. W. (2008). A four-category scheme for coding and assessing the level of reflection in written work. *Assessment Evaluation in Higher Education*, 33 (4), 369-379.
- Kline, P. (1994). *An easy guide for factor analysis*. U.K. Routledge.
- Kökdemir, D. (2003). *Belirsizlik durumlarında karar verme ve problem çözme*. Yayınlanmamış doktora tezi, Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Sosyal Psikoloji Anabilim Dalı, Ankara.
- Köksal, N. ve Demirel, Ö. (2008). Yansıtıcı düşünmenin öğretmen adaylarının öğretmenlik uygulamalarına katkıları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34, 189-203.

Leung, D. Y. P., & Kember, D. (2003). The relationship between approaches to learning and reflection upon practice. *Educational Psychology*, 23, 61-71.

Lie, A. L. Y. (2007, March). *Students' reflective development in a problem based learning environment*. Paper presented at the International Problem Based Learning Symposium, Singapore.

Loke, A. Y., & Chow, F. L. W. (2003). *Effect of peer tutoring in enhancing reflective learning and thinking* (Tackling The Challenges to Professional Education Project). Hong Kong Polytechnic University.

Lucas, U., & Tan, P. L. (2006, May). *Assessing levels of reflective thinking: The evaluation of an instrument for use accounting and business education*. Paper presented at the 1st Pedagogic in Higher Education Conference, Liverpool Hope University, Liverpool.

Mahardale, J., Neville, R., Jais, N. & Chan, C. (2008). *Reflective thinking in a problem based English programme: A study on the development of thinking in elementary students*. Retrieved March 1, 2008 from <http://www.pbl2008.com/PDF/0048.pdf>

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

Phan, H. P. (2007). Examination of student learning approaches, reflective thinking, and self efficacy beliefs at the university the South Pacific: A path analysis. *Educational Psychology*, 27 (6), 789-806.

Pollard, A. (Ed.). (2002). *Readings for reflective teaching*. London: Continuum.

Semerci, Ç. (2007). Öğretmen ve öğretmen adayları için Yansıtıcı Düşünme Eğilimi (YANDE) Ölçeğinin geliştirilmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 7, 733-740.

Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenirlik ve geçerlilik*. Ankara: Seçkin Yayınları.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. New Jersey: Lawrence Erlbaum Associates Publishers.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.

Tan, K. S., & Goh, N. K. (2008). Assessing students reflective responses to chemistry-related learning tasks. *USChina Education Review*, 5 (11), 28-36.

Tripp, D. (2003). Action inquiry. *Action Research e-Reports*, 017. Retrieved July 3, 2008 from [www.fhs.usyd.edu.au/arow/arer/017.htm](http://www.fhs.usyd.edu.au/arow/arer/017.htm)

Turgut, F. (1990). *Eğitimde ölçme ve değerlendirme metodları*. Ankara: Saydam Matbaacılık.