International Journal of Assessment Tools in Education
2022, Vol. 9, No. 4, 1030–1061
https://doi.org/10.21449/ijate.1058145
Published at https://ijate.net/ https://dergipark.org.tr/en/pub/ijate *Research Article*

A Comperative Adaptation of the Crick Learning for Resilient Agency (CLARA) with Classical Test Theory and Item Response Theory

Hasan Fehmi Ozdemir[1*], **Omer Kutlu**[2], **Shaofu Huang**[3], **Ruth Crick**[4]

[1]Şırnak University, School of Physical Education and Sports, Department of Physical Education and Sports, Türkiye
[2]Ankara University, Faculty of Education, Department of Educational Sciences, Türkiye
[3]WILD Learning Sciences CIC, United Kingdom
[4]University of Technology Sydney, Connected Intelligence Centre, Australia

ARTICLE HISTORY
Received: Jan. 15, 2022
Revised: Aug. 01, 2022
Accepted: Sep. 18, 2022

Keywords:
Learning power, Resilient agency, Classical Test Theory, Item Response Theory, Intercultural test adaptation.

Abstract: The aim of this study is to adapt the Crick Learning for Resilient Agency (CLARA) to Turkish culture, and to examine the psychometric features of the Inventory according to both Classical Test Theory (CTT) and Item Response Theory (IRT). In this respect, it is a descriptive level survey design research. Two different study groups were formed in accordance with the purpose of the study. Lingual equivalence applications were performed on two separate groups, one of which consisted of English Language and Literature Department students and the other consisted of English Language instructors. 1054 students participated in the validity and reliability studies from 101 different undergraduate programs at Ankara University. Before testing the research questions, it was examined whether the assumptions of CTT and IRT were met. With the application data; the predicted item discrimination indices, ability levels, students' scores forming their learning power profiles, and reliability coefficient values were found to be similar in both theories. It can be said that with CLARA-Tr, obtained by adapting CLARA, a valid and reliable tool has been provided to the Turkish literature to be used in future studies.

1. INTRODUCTION

Psychological tests are the subject of determining the cognitive, affective and dynamic characteristics of people and are used in scientific fields such as medicine, psychology and education. In general terms, tests provide information about the psychological characteristics of individuals and help to make decisions about individuals based on the results obtained from their application (Cohen & Swerdlik, 2010; Cronbach, 1960).

Wherever there are psychological activities, emphasis is placed on studies related to psychological tests. Studies on test or scale development and adaptation have an important place in Turkish literature. As different aspects and characteristics of human behavior are discovered, the need for different assessment tools to measure these characteristics is increasing. Instruments, measuring different psychological structures for different age groups are needed.

*CONTACT: Hasan Fehmi Ozdemir ✉ hfozdemir@sirnak.edu.tr Şırnak University, School of Physical Education and Sports, Department of Physical Education and Sports, Türkiye

e-ISSN: 2148-7456 /© IJATE 2021

This requirement can be met by the development of new measurement tools or by adapting suitable measurement tools, developed in different cultures, to Turkish culture. Both ways have either superior, and advantageous or inferior, and disadvantageous aspects. However, scale adaptation studies have benefits such as; the widespread use of technical knowledge, the establishment of international joint research relationships and the increase of information exchange, the localization of psychology, the initiation of cross-cultural comparative studies, the increase in the potential of collecting objective data on various subjects in the country, and contributing to the production of knowledge through its use in other research studies (Hambleton & Patsula, 1999; Hambleton et al., 2005; International Test Commission, 2018; Savaşır, 1994).

Undoubtedly, one of the most important steps in scale development or adaptation studies is to demonstrate the experimental reliability and validity proofs of the instrument being developed or adapted. Because the value and usability of the findings or results, obtained from psychological measurement tools, to make decisions about individuals is directly related to the psychometric properties of these tools at scale and item levels. One step further, no matter how strong the theoretical background of a scientific research is, if the tools used in the data collection process do not have the necessary psychometric qualities, there will be a trust problem in the interpretation of the findings of a research study, and it will be inevitable to make wrong decisions with the results obtained from this tool (Özdemir et al., 2019). Another important point is to use different theories and various methods and techniques developed based on these theories to determine the psychometric properties of measurement tools.

In the Turkish literature, there are studies in which Classical Test Theory (CTT) and Item Response Theory (IRT) are used in measurement tool development processes or in the prediction of item and test parameters of previously developed tools (Karakılıç, 2009; Kelecioğlu, 2001; Nartgün, 2002; Uysal, 2015). However, it is observed that there are many measurement tools adapted to Turkish culture in order to measure psychological characteristics, and almost all of these instruments' adaptation processes are based on CTT, due to the ease of implementation. However, when both theories are compared, it is known that the CTT has some limitations compared to the IRT (Hambleton, Swaminathan, & Rogers, 1991). In the intercultural adaptation studies of measurement tools, it is important to determine the psychometric properties not only according to CTT, but also to IRT, which has stronger assumptions. As a result of examining the harmony of the qualities determined by the methods and techniques based on different test theories, the usability of the scores to be obtained by the application of the said measurement tools will also increase.

Another important issue that has been frequently criticized in recent years is the proliferation of the test-oriented teaching and learning practices. The widespread use of large-scale tests and evaluations based on their results, force many tutors around the world to teach learners only multiple-choice test taking tips and the strategies to deal with them. This causes many learners to fail, by preventing them from gaining knowledge about participation in learning processes and self-learning (Deakin Crick et al., 2004). The way that will lead individuals to a solution is to encourage them to learn in a willing and relevant way in the face of new needs and opportunities. For this reason, in order to raise individuals with the mentioned qualities, education and measurement policies should be structured differently, and educational institutions at all levels should be structured to serve this.

If the capacity and willingness to learn and continue learning throughout life is accepted as the central point in the concept of "learning", it is of great importance to use tools that measure the capacities and desires of individuals and their constantly evolving and changing qualities. "What makes the individual participate in the learning process, continue his/her learning, and want to learn effectively and efficiently while doing this?" The answer to this question has been

an important starting point for the development of different measurement tools. In this context, one of the measurement tools that we come across in the literature is the Crick Learning for Resilient Agency (CLARA), which defines and measures the "learning power" of an individual (Deakin Crick et al., 2015). In the Turkish literature reviews, no measurement tool was found to measure learning power. Therefore, it is thought that adapting the CLARA, which is widely used in the international literature and has appropriate psychometric properties, to Turkish culture will contribute to the Turkish society and the field of measurement and evaluation.

1.1. Aim of the Study and Research Questions

The main aim of this study is to adapt Crick Learning for Resilient Agency (CLARA) Inventory to Turkish culture as CLARA-Tr, and to analyze and compare the psychometric properties of the Inventory in the adaptation process according to the methods and techniques of both Classical Test Theory (CTT) and Item Response Theory (IRT). In line with this main aim, answers to the following research questions were sought:

1. Is there a relationship between the scores obtained from the English and Turkish forms of CLARA?
2. Is CLARA's original factor structure confirmed in Turkish culture?
3. What is the relationship between the values of the psychometric properties of Turkish form of CLARA estimated according to the CTT and IRT?
   - 3.1. Is there a relationship between the item discrimination indices (item score - corrected total score correlation and a_i parameter) of CLARA-Tr's items according to the CTT and IRT?
   - 3.2. Is there a relationship between the levels of the features/traits measured by CLARA-Tr items (arithmetic mean and b parameter) estimated according to the CTT and IRT?
   - 3.3. Is there a relationship between learning power levels estimated from CLARA-Tr according to CTT and IRT?
   - 3.4. What is the reliability of CLARA-Tr according to the CTT and IRT?

1.2. Significance of the Study

One of the priorities included in the Lifelong Learning Strategy Document and Action Plan for the period 2014-2018 (Ministry of National Education [MoNE], 2014; 2018), which was prepared to increase the effectiveness and efficiency of the lifelong learning system in Türkiye, is "constructing the culture and raising the awareness of lifelong learning in society". In this context, it is planned to expand the studies for the adult population to acquire basic skills (such as communication in mother tongue and foreign language, digital competencies, learning to learn etc.). In order to achieve this, the individual must first recognize himself/herself, recognize his/her weaknesses and strengths as a learner, and see learning as a necessity. Within the scope of "development of lifelong learning monitoring and evaluation system", which is another priority in the said Strategy Document, creating statistics and researches is expected to be done from responsible institutions and organizations (Ministry of National Education and Higher Education Council, Universities in this context) that will help develop policies and strategies.

One of the important reasons for conducting such a study is that there is no measurement tool that measures learning power in the Turkish literature reviews and the need to do more research on metacognitive skills such as self-awareness, curiosity, creativity, readiness to learn, and resilience, which are among the basic life skills. Due to the requisite and important need for resilient agency at every stage of individuals' learning journeys from purpose to performance, it has been acted with the thought that it will make a significant contribution to the priorities and achievement of these priorities in Türkiye Lifelong Learning Strategy Document and

Action Plan 2014-2018. In this context, it was decided to adapt and introduce Crick Learning for Resilient Agency (CLARA) self-assessment tool into Turkish culture.

Another important reason for the selection of this Inventory is that CLARA has not only the ability to satisfy the requirements of the researches in which it is used as a data collection tool, but also to provide instant feedback and forward notifications based on monitoring (formative) to individuals as learners. Because, some studies conducted in Türkiye reveal the low lifelong learning disposition levels of undergraduate students (Diker Coşkun & Demirel, 2012; Tunca et al., 2015). Changing this negative perception and letting the university students to see the strengths and weaknesses of their own learning power as independent learners at the undergraduate level, is seen as an important investment in their learning journeys that continue from cradle to grave, and therefore to themselves and the society they live in.

As a result, both a new measurement tool has been added to the Turkish literature for the researchers who want to have knowledge about learning power and for their future studies, and also an example was provided for the comparison process of test theories in the intercultural test/scale adaptation process.

1.3. The Crick Learning for Resilient Agency (CLARA) Profile

The research programme which has led to the publication and various applications of CLARA began in 2000 at the University of Bristol, UK. Originally funded by the LifeLong Learning Foundation, and building on the work of Carr and Claxton (2002) it addressed the challenge of identifying personal qualities and characteristics which define a 'good learner'- someone who is able to engage effectively and profitably with new learning opportunities across the lifespan. As well as identifying these qualities, the purpose of the research was to devise a learning analytics tool that could be used to assess where an individual was located on those qualities at any given time and in any given context and thus provide them with data that could be used formatively to enable them to develop their capacity to learn how to learn. Then the Assessment Reform Group (2010, December) in the UK had developed a significant programme of work, which aimed make 'assessment for learning' a focus for policy and practice Broadfoot (1998). There was, even then, substantial evidence of the negative impact of high stakes testing and summative assessment on students' motivation for learning Harlen and Deakin Crick (2003a and 2003b) and this programme of research set out to develop alternative forms of assessment for learning that could be both formative for teachers and ipsative for learners in that it could provide a foundation for teacher supported but student-led, self-directed change in learning how to learn.

The original research (Deakin Crick et al., 2004; Deakin Crick and Wilson, 2005; Deakin Crick, 2005; Deakin Crick, 2007) was a factor analytic study which drew together items created to reflect what was known at the time about lifelong learning and 'learning power' a popular term coined first by (Claxton, 1999) to refer to a person's capacity for learning how to learn. It drew on a susbtantive literature review and included items from socio-cultural learning theory and pedagogical studies. The factor analysis produced seven latent variables, which have remained constant over time through successive quantitative studies (Arthur et al., 2006; Deakin Crick and Yu, 2008; Deakin Crick et al., 2013; Deakin Crick et al., 2015). The original tool was called ELLI (The Effective Lifelong Learning Inventory).

From 2001 the data was used in practice as well as research, and returned to teachers in classrooms as a digital learning analytic in the form of a spider diagram. In keeping with the theoretical foundations of the study, this was designed so that teachers and learners were encouraged to explore patterns and interpretations, rather than a numerical score, or set of scores, which would inevitably lead to a more summative self-judgement (Deakin Crick, 2005; Deakin Crick, 2006; Deakin Crick and McCombs, 2006; Deakin Crick, 2009a and 2009b).

A programme, which attended to both practice, research and policy, was a challenge in a traditional University. After three failed attempts to commercialise the work, the University enabled a re-analysis of accumulated data and the publication of a revised version known as the Crick Learning for Resilient Agency profile (Deakin Crick et al., 2015). As a creative commons publication this opened up new opportunities for ongoing research and development. The study reported in this paper builds on this work.

1.4. Learning Power

The term learning power has come into popular usage to describe the capacity a person has to learn and to engage profitably with risk, uncertainty and challenge. In other words, they know how to go about finding out what to do when the solution to a challenge is not known in advance. The ELLI tool and subsequently the CLARA tool built on theoretical foundations which took seriously a holistic approach to learning. This included the role of (i) dispositions, awarenesses and skills (ii) identities – the beliefs, values and attitudes about self, learning and knowledge held by the learner, (iii) narratives – the socio-cultural formation of learners over time and (iv) the quality and substance of learning relationships (Deakin Crick, Broadfoot and Claxton, 2004). This led to a set of Scales, known as dimensions of learning power, which measured eight variables. Each of these included cognition, affect and volition and were presented to learners in real time as a reflection of their 'learning power' in a particular context at a particular point in time. On the basis of the underlying theory of agency and choice, the feedback was designed to stimulate learner ownership, awareness and responsibility for self-directed change. For this reason, the visual imager was important in assessment terms because it stimulates reflection on one's self-identity and story and offers opportunities for reflexive self-awareness and change in purposeful agency.

The Scales of CLARA are presented in greater depths elsewhere (Deakin Crick et al., 2015). A summary is presented in Figure 1 below.

Figure 1. *The Scales of CLARA.*

1.4.1. *Mindful agency scale (9 items)*

Mindful Agency is taking responsibility for your own learning. It's about how you manage your feelings, your time, your energy, your actions and the things you need to achieve your goals. It's knowing your purpose - then knowing how to go about achieving it; stepping out on the path towards your goals.

1.4.2. *Hope and optimism scale (3 items)*

Hope and Optimism is being confident that you can change and learn and get better over time. It is helped by having a positive learning story to reflect upon, that gives you a feeling of having 'come a long way' and of being able to 'go places' with your learning.

1.4.3. *Sense making scale (7 items)*

Sense Making is making connections between ideas, memories, facts - everything you know - linking them and seeing patterns and meaning. It's about how 'learning matters' to you, connecting with your own story and things that really matter.

1.4.4. *Creativity scale (8 items)*

Creativity is using your imagination and intuition, being playful and 'dreaming' new ideas, having hunches, letting answers come to you, rather than just 'racking your brains' or looking things up. It's about going 'off the beaten track' and exploring ideas.

1.4.5. *Curiosity scale (6 items)*

Curiosity is your desire to get beneath the surface, find things out and ask questions, especially 'Why?' If you are a curious learner, you won't simply accept what you are told without wanting to know for yourself whether and why it's true.

1.4.6. *Collaboration scale (3 items*)

Collaboration is how you learn through your relationships with others. It is about knowing who to turn to for advice and how to offer it too. It's about solving problems by talking them through, generating new ideas through listening carefully, making suggestions and responding positively to feedback.

1.4.7. *Belonging scale (3 items)*

Belonging reflects how much you feel you belong as part of a 'learning community' – at work or at home, or in your wider social network. It's about the confidence you gain from knowing there are people you learn well together with and to whom you can turn when you need guidance, support and encouragement.

1.4.8. *Orientation to learning scale (10 items)*

Orientation to Learning is about the degree to which a person is open to new ideas and to challenge and having the 'inner strength' to move towards learning and change, rather than either giving up and withdrawing or 'toughing it out' and getting mad with the world. Becoming more open to learning is like a pathway to all the other Scales of learning power, and just as the other Scales it also help you become more open to learning. This Scale is sometimes referred to simply as 'Openness to Learning'.

1.5. Resilient Agency

The term resilience is much used in various contexts and domains. In the psychological literature resilience refers to those qualities that an individual has that enables them to succeed despite adverse conditions or circumstances (Rutter, 1985; Rutter, 2012; Masten, 2007). In the 2015 revision of the learning power assessment tool, the term was chosen to describe the overall purpose of the whole assessment event, in response to all of the now eight Scales of learning power, which is to empower the individual to understand themselves as a learner and to use that

understanding to explore strategies for change. In the early version, ELLI, the Scale now called orientation to learning was described as Resilience in keeping with commercial applications of learning power (Gornall et al., 2005). However, the data demonstrated that simply persisting in particular behaviours did not necessarily enable one to succeed despite adverse conditions or circumstances. Indeed, in some contexts, it led to more negative outcomes (Deakin Crick and Salway, 2006). Resilience in learning is complex and includes the capacity to persist, but also must include the capacity to explore identity and purpose, to generate questions, utilize one's imagination and develop positive relationships. In the context of developing learning power, Resilient Agency was identified as a descriptor for the purpose of the whole assessment event, which is to stimulate self-leadership and self-directed change strategies which lead towards a more profitable future.

2. METHOD

2.1. Research Model

In this study, it was examined whether the psychometric properties of Crick Learning for Resilient Agency (CLARA) determined by different methods and techniques of Classical Test Theory (CTT) and Item Response Theory (IRT) differed or not in the process of adapting the Inventory to Turkish culture. The study aims to reveal the psychometric properties of the said Inventory as they exist on the basis of two different test theories. In this respect, this study is a descriptive survey research (Büyüköztürk et al., 2014; Erkuş, 2013; Scott & Usher, 2011).

2.2. Study Groups

In scale adaptation studies, due to the limitations in terms of time, money and labor, the sample is chosen from easily accessible and practicable units. For this reason, instead of working with the population and sample, it is preferable to conduct the research with a "study group", which is reached through convenient sampling from individuals similar to the target group. In this study, the target group was determined as undergraduate students, and in line with the purpose of the study, two different study groups were formed from students studying at different departments of Ankara University, and also a group of English lecturers working at Ankara University have participated in the lingual equivalence applications.

2.2.1. *Linguistic equivalence application groups*

Linguistic equivalence applications were carried out on two separate groups that were deemed to be sufficient in both languages. In the first group, there were a total of 31 students from the 2nd and 4th grade students who are continuing their education in Ankara University, Department of English Language and Literature. In the second group there were 35 English lecturers working at Ankara University Turkish and Foreign Language Application and Research Center.

2.2.2. *Validity & reliability studies application group*

It has been taken into consideration that the analyzes to be made in order to determine the psychometric properties of the adapted instrument will be made according to both CTT and IRT. For this reason, taking into account the lower limits of the number of participants suggested by researchers such as Crocker and Algina (1986), Reise and Yu (1990) and De Ayala (2009), which is sufficient for statistical methods to be used and necessary to provide assumptions and to ensure variability, this application was conducted on a group of 1054 students who are continuing their education at 101 different undergraduate programs of Ankara University. 33.11% (n = 349) of the students in this group are male and 66.89% (n = 705) are female. Considering the grade levels, 2.56% (n = 27) of the group was preparatory class, 7.97% (n = 84) were 1st grade, 16.41% (n = 173) 2nd grade, 24.67% (n = 260) 3rd grade, 42.41% (n = 447) 4th grade, 4.74% (n = 50) 5th grade and 1.23% (n = 13) 6th grade students.

2.3. The Adaptation Process of the Crick Learning for Resilient Agency (CLARA)

The following steps have been followed in the process of adapting CLARA, which is planned to be introduced into Turkish psychometry field:

1. Participation in the workshop organized in Bristol / England, in order to receive the necessary training on CLARA's application, scoring and interpretation of the scores.

2. CLARA, was translated from its original language English to Turkish by a group of expert translators who have mastered the language and culture, and then back translated into Turkish by a different group of translators. The back-translations of the Inventory, and the Scale names, and also the items, and the response categories were shared with the developers, and their opinions and approvals were received. The original form, the form translated into Turkish and the back translation form were presented to the evaluation of a group of instructors who know both languages well and who are knowledgeable about measurement and learning. While considering the back-translations, the evaluators were asked to compare the Turkish translation form with the original form, in terms of language and meaning.

3. The necessary corrections were made in line with the suggestions and evaluations of the expert group, and the final version of the Turkish form was presented to the opinion of the Turkish language experts and final checks were carried out.

4. Bilingual group design was used to ensure linguistic equivalence. In this direction, it is necessary to apply the instrument's original and translated forms on a group that is deemed to be sufficient in both languages. For this reason, applications were made in two separate groups in order to test whether linguistic equivalence was achieved. In both groups, the original form and the translation form of the tool were applied every three weeks. After the applications, the relationship between the scores obtained from the original and target language forms of the scale was examined.

In this study, the procedure steps suggested by Hambleton and Swaminathan (1985) for the estimation of psychometric properties of Likert type measuring instruments based on IRT were followed. In the estimation of the psychometric properties of CLARA based on IRT, the inventory was first applied to a group with a high number of participants. It was tested whether the data meet the IRT assumptions; unidimensionality and local independence, and whether the data fit the selected model. Ability levels (θ) and item parameters were estimated with MULTILOG 7.03 program. Also, IBM SPSS 22 and LISREL 8.8 were used for statistical analysis of the data within the scope of the study. Before starting the testing phase of the research questions, it was examined whether the data met the CTT and IRT assumptions required for analysis. Kolmogorov-Smirnov and Shapiro Wilk tests were used together with descriptive statistics, and the histogram graphs in the analysis of whether the data provided the assumption of normality. In testing the assumptions of unidimensionality and local independence, the results of two confirmatory factor analysis were used. In terms of Item Response Theory, data model fit was analyzed using the "-2 lnL" statistic, and also the level of data-model fit was examined by the difference between the observed and expected proportions of responses to the item response categories.

An example of the MULTILOG program output (Belonging Scale) showing the a and b parameters estimated according to the IRT of the CLARA-Tr items used in this study, as well as the model-data fit and marginal reliability coefficient values are given in Appendix 1.

3. RESULT

The results / findings obtained regarding the research questions are given and discussed below respectively.

3.1. Findings Regarding the First Research Question – The Relationship Between the Scores Obtained from the Application of English and Turkish Forms of CLARA

In order to search for an answer to the question "Is there a relationship between the scores obtained from the application of English and Turkish forms of CLARA?" and to test whether linguistic equivalence was achieved between the original and Turkish forms of the Inventory, linguistic equivalence applications were carried out in two separate groups (n1 = 31 and n2 = 35). In both groups, the original and the translation forms of the tool were applied three weeks apart, and the relationship between the scores obtained from these applications was examined with the Pearson Product-Moments Correlation coefficient. The correlation values are presented in Table 1.

Table 1. *Relationship Between Scores Obtained from English and Turkish Forms of CLARA.*

Scales	n_1=31		n_2=35	
(English / Turkish)	*r*	*p*	*r*	*p*
Belonging	0.75	0.000	0.78	0.000
Collaboration	0.72	0.000	0.71	0.000
Creativity	0.76	0.000	0.82	0.000
Curiosity	0.81	0.000	0.87	0.000
Hope & Optimism	0.70	0.000	0.73	0.000
Mindful Agency	0.78	0.000	0.79	0.000
Orientation to Learning	0.71	0.000	0.81	0.000
Sense Making	0.79	0.000	0.80	0.000

When Table 1 is examined, it is determined that there is a positive, high and significant (r = 0.70-0.87, p <0.01) relationship between the scores obtained from the English and Turkish forms of CLARA's both linguistic equivalence applications. Accordingly, it can be accepted that linguistic equivalence is provided between the original and Turkish forms of CLARA (Büyüköztürk et al., 2014). The item examples included in CLARA and CLARA-Tr that emerged as a result of this process are presented in Appendix 2.

3.2. Findings Regarding the Second Research Question – The Structure of CLARA Verified in Turkish Culture

Randomly chosen, with sufficient sample sizes two separate (n_1 = 550 and n_2 = 504) confirmatory factor analyzes were conducted on the data obtained from the validity & reliability studies application to find an answer to the question "Is the original structure of CLARA verified in Turkish culture?" and to determine whether the eight-scale original structure of the Inventory was also confirmed by Turkish undergraduates or not. The analyzes were carried out using LISREL 8.8 program. Covariances were used as the moment matrix, and maximum likelihood (ML) estimation method was used in CFA. Fit indices obtained as a result of the analyzes are given in Table 2.

Table 2. *Confirmatory Factor Analysis of CLARA Turkish Form Fit Indices.*

Fit Indices	CFA 1 (n=550)	CFA 2 (n=504)
	Values	
Chi - Square (X^2)	3956.82	2983.64
Degrees of Freedom (df)	1398	1152
X^2/sd	2.83	2.59
Non-Normed Fit Index (NNFI)	0.95	0.95
Comparative Fit Index (CFI)	0.96	0.95
Root Mean Square Error of Approximation (RMSEA)	0.069	0.066
Root Mean Square Residual (RMSR)	0.016	0.015
Standardized RMR	0.08	0.08

Fit indices of the models obtained from CFA's were examined and Chi-square values (χ^2 = 3956.82, N = 550, *df* = 1398, *p* = 0.00; χ^2 / *df* = 2.83 and χ^2 = 2983.64, N = 504, *df* = 1152, *p* = 0.00; χ^2 / *df* = 2.59) were found to be significant. Fit index values were obtained as RMSEA = .069 and .066, NNFI = .95 and .96, CFI = .95 and .95, RMR = .016 and .015, Standardized RMR = 0.08 and 0.08 respectively. 90% confidence interval of RMSEA are between 0.057-0.071 and 0.054-0.069. According to Jöreskog and Sörbom (1993), Hu & Bentler (1999), Kline (2005), Özdamar (2013), Sümer (2000), Şimşek (2007), Vieira (2011) the values in Table 2 indicate acceptable fit. According to these data, it was decided that the original structure of CLARA was also verified by Turkish undergraduate students, and that data on learning power could be collected from university students in a valid and reliable manner by its application.

3.3. Findings Regarding the Third Research Question – Relationship Between the Values of the Psychometric Properties of the Turkish Form of CLARA

The third research question of the study is "What is the relationship between the values of the psychometric properties of the Turkish form of CLARA, which are estimated based on CTT and IRT?" Findings and comments regarding the sub-questions to be answered within the scope of this question are presented below.

3.3.1. *Research question 3.1. findings – relationship between the item discrimination index values of CLARA-Tr*

"Is there a relationship between the item discrimination index values of CLARA-Tr, which are estimated based on CTT and IRT?" For this question, the relationship between the item discrimination indices of each item estimated according to two theories was tested with the Spearman Rank Difference Correlation Coefficient.

In the estimation of item discrimination index according to CTT, correlation based item analysis technique was used. For this purpose, the relationship between the responses of the participants to the items and their corrected total scores from the scale in which that item is included was calculated with the Pearson Product-Moment correlation coefficient. The corrected total score was calculated by subtracting each participant's relevant item score from his/her raw score obtained from that scale. In IRT, on the other hand, a parameter was estimated for each item according to the Graded Response Model of Samejima (Samejima, 1969) and the relationship between the values obtained according to both theories was examined. The Graded Response Model is an extension of the two-parameter logistic model (2PL). This model is appropriate when the responses of an individual to an item can be classified into more than two ordered categories, such as to represent different levels of agreement or frequency to a certain statement. In Table 3, the discrimination indices of the items in the Inventory, which is estimated based on CTT and IRT, are presented.

Table 3. *Discrimination Index Values of CLARA-Tr Items Estimated According to CTT and IRT.*

Scale	Item No	CTT (Item Score-Corrected Total Score Correlation)	IRT (a parameter)
Belonging	7	0.970	4.12
	17	0.972	4.46
	45	0.964	2.29
Collaboration	6	0.886	2.39
	35	0.930	0.98
	48	0.885	0.92
Creativity	9	0.982	1.80
	29	0.982	1.22
	1	0.983	1.86
	41	0.985	1.12
	16	0.984	0.94
	31	0.986	2.01
	39	0.982	1.58
	11	0.990	1.09
Curiosity	2	0.972	1.10
	47	0.978	1.34
	33	0.977	3.04
	22	0.978	2.90
	5	0.988	1.13
	38	0.986	1.23
Hope & Optimism	13	0.962	3.22
	24	0.967	1.91
	49	0.971	6.00
Mindful Agency	3	0.986	1.09
	10	0.992	1.41
	15	0.990	1.57
	23	0.989	1.32
	26	0.986	1.26
	34	0.985	1.34
	36	0.990	1.69
	43	0.986	1.36
	46	0.993	2.07
Orientation to Learning	14	0.979	1.03
	18	0.981	1.64
	20	0.985	1.98
	21	0.989	1.60
	25	0.986	1.46
	28	0.983	1.98
	30	0.989	1.15
	32	0.975	1.58
	37	0.992	1.61
	42	0.983	1.87
Sense Making	4	0.979	1.11
	8	0.963	1.13
	12	0.948	1.82
	19	0.910	1.61
	27	0.947	1.10
	40	0.973	1.32
	44	0.933	1.61

In Table 4, descriptive statistics of the discrimination index values of the items of the Inventory, which are estimated based on the CTT and IRT, are presented.

Table 4. *Descriptive Statistics of the Discrimination Index Values of CLARA-Tr Items Estimated According to CTT and IRT.*

Descriptive Statistics				
	CTT	Stand. Error	IRT	Stand. Error
Minimum	0.890		0.92	
Maximum	0.990		6.00	
$\bar{X}$	0.973	0.0035	1.78	0.139
Median	0.982		1.58	
SD	0.249		0.97	
Kurtosis	0.668		7.68	0.67
Skewness	0.340		2.55	0.34
Range	0.100		5.08	
Number of Items (k)	49		49	
Number of Students	1054		1054	

When Table 3 and 4 are examined together, it is seen that the discrimination indices of the items of the eight scales that make up the Inventory vary between 0.885 (Collaboration Scale, item 48) and 0.993 (Mindful Agency Scale, item 46) and the median is 0.982. In the analysis of correlation-based item discrimination, it is concluded that as the values approach 1.00, the item measures the feature/trait that is measured with the whole scale to which it belongs, and it can better discriminate the individuals who have this feature/trait and those who do not. Based on this, it was observed that all 49 items in 8 Scales of the Inventory, which was adapted to Turkish culture, had a high level of discrimination.

It is seen that the values of a parameter estimated according to the IRT vary between 0.92 (Collaboration Scale, item 25) and 6.00 (Hope and Optimism Scale, item 49) and the median is 1.58. In the IRT, it is accepted that the items with a discriminative power of 1.00 and above are sufficiently discriminating (Hambleton & Swaminathan, 1985). This can be interpreted as that 49 items of the inventory can discriminate the individuals who have the desired feature/trait to be measured with the scales they belong to, and those who do not.

Despite the good discrimination index values obtained according to the two test theories, only the items of the "Belonging" and "Mindful Agency" scales discrimination values determined according to CTT and IRT showed significant relationship when examined with the Spearman Rank Correlation Coefficient ($p < 0.05$), no significant relationship was found for the items of the other six scales. According to this result, it can be interpreted that the item discrimination indices of "Belonging" and "Mindful Agency" scales estimated according to the two theories are similar to each other and these values are comparable.

3.3.2. *Research question 3.2. findings – relationship between the levels of the features/traits measured by CLARA-Tr*

Another sub-question to be answered within the scope of the third research question of the study is "Is there a relationship between the levels of the features/traits measured by CLARA-Tr items (arithmetic mean and b parameter) estimated according to the CTT and IRT?" For this question, the relationship between the levels of the features/traits measured by each CLARA-Tr item based on two theories, was tested with Spearman Rank Differences Correlation.

According to the CTT, the levels of the features/traits measured by the items were calculated by the arithmetic mean of the responses given to the relevant item by the students in the study group. According to the IRT, the levels of the features/traits measured by each item were determined by taking the arithmetic mean of the b parameter values estimated according to

Samejima's Graded Response Model (Samejima, 1996). The values of the levels of the features/traits measured by the items based on both theories are given in Table 5.

Table 5. *Levels of the features/traits measured by CLARA-Tr Items Estimated According to CTT and IRT.*

Scale	Item No	CTT	IRT					
		Art. Mean	b_1	b_2	b_3	b_4	b_5	b_{AM}
Belonging	7	3.12	-1.38	-0.68	-0.14	0.18	0.72	-0.26
	17	4.43	-1.35	-0.76	-0.16	0.22	0.70	-0.27
	45	4.87	-1.51	-0.49	0.21	0.66	1.27	0.03
Collaboration	48	4.05	-2.94	-1.31	-0.04	0.95	2.29	-0.21
	6	4.05	-2.82	-1.96	-1.30	-0.80	0.00	-1.38
	35	4.31	-3.61	-2.16	-1.06	-0.26	0.80	-1.26
Creativity	9	4.63	-2.43	-1.26	-0.45	0.21	0.99	-0.59
	29	4.70	-2.61	-1.00	0.10	0.90	1.95	-0.13
	1	3.55	-5.60	-3.17	-1.19	-0.24	1.23	-1.79
	41	4.03	-4.36	-2.56	-1.11	-0.17	1.02	-1.44
	16	4.25	-4.66	-2.84	-1.48	-0.43	0.97	-1.69
	31	5.03	-2.74	-1.54	-0.68	0.02	0.85	-0.82
	39	3.38	-2.34	-0.92	-0.12	0.53	1.48	-0.27
	11	5.00	-6.28	-3.82	-2.29	-1.25	0.06	-2.72
Curiosity	2	3.23	-2.52	-0.66	0.59	1.40	2.68	0.30
	47	4.10	-3.77	-2.25	-1.23	-0.48	0.44	-1.46
	33	4.45	-2.01	-1.03	-0.28	0.21	0.83	-0.46
	22	4.17	-1.89	-1.04	-0.39	0.16	0.95	-0.44
	5	5.16	-5.25	-4.18	-2.26	-1.25	-0.01	-2.59
	38	5.01	-4.93	-3.47	-1.98	-0.92	0.26	-2.21
Hope & Optimism	49	5.64	-2.13	-1.16	-0.46	0.08	0.79	-0.58
	13	4.56	-2.02	-1.20	-0.47	0.13	0.93	-0.53
	24	5.16	-3.51	-2.51	-1.64	-0.95	0.00	-1.72
Mindful Agency	3	3.86	-3.60	-2.03	-0.81	0.25	1.87	-0.86
	15	4.71	-3.25	-2.44	-1.29	-0.38	0.82	-1.31
	43	4.24	-2.36	-0.73	0.39	1.15	2.26	0.14
	36	3.29	-3.76	-2.07	-1.01	-0.19	0.89	-1.23
	46	4.23	-3.46	-2.51	-1.48	-0.64	0.39	-1.54
	23	4.92	-4.15	-2.65	-1.67	-0.72	0.35	-1.77
	34	3.70	-2.62	-1.46	-0.40	0.48	1.70	-0.46
	26	3.86	-3.46	-2.13	-1.07	-0.25	0.90	-1.20
	10	4.99	-4.64	-3.49	-1.89	-0.90	0.34	-2.12
Orientation to Learning	20	4.56	-1.59	-0.37	0.47	1.08	1.80	0.28
	30	3.14	-0.63	1.22	2.13	2.81	3.82	1.87
	25	2.08	-8.20	-5.87	-3.16	-1.04	1.60	-3.33
	28	4.46	-1.40	-0.34	0.48	1.00	1.80	0.31
	14	5.41	-0.72	0.58	1.33	1.85	2.64	1.14
	42	3.89	-2.49	-1.36	-0.46	0.05	0.74	-0.70
	21	2.58	-1.16	0.18	1.15	1.89	2.70	0.95
	18	4.04	-5.08	-2.61	-0.81	0.50	2.31	-1.14
	32	3.73	-3.58	-1.70	-0.27	0.77	2.62	-0.43
	37	5.20	-8.78	-6.12	-4.18	-2.25	-0.06	-4.28
Sense Making	19	4.05	-3.32	-2.15	-1.23	-0.32	0.89	-1.23
	40	3.61	-8.51	-4.09	-1.58	0.60	3.74	-1.97
	4	2.51	-5.36	-4.54	-3.22	-1.96	-0.51	-3.12
	27	4.51	-5.07	-4.41	-3.83	-2.81	-1.23	-3.47
	8	4.46	-4.53	-2.44	-0.96	0.14	1.65	-1.23
	12	5.12	-3.79	-2.92	-1.75	-0.78	0.35	-1.78
	44	4.09	-3.14	-1.66	-0.52	0.31	1.25	-0.75

When Table 5 is examined, it is seen that the levels of the features/traits measured by the items according to the CTT vary between 2.51 (Sense Making Scale, item 4) and 5.64 (Hope and Optimism Scale, item 49) and the median is 4.24. It is seen that the vast majority of the items (38 items) have a negative skewness value and when all items are considered, the average skewness value is -4.32. When all these findings are evaluated together, it has been determined that both the items generally measure the feature/trait to be measured with the scales they belong to at a high level and all the items have a relatively high approval rate. In other words, it can be said that participating students have chosen the high-level end of the response categories.

According to the IRT, one less number of b parameters were estimated from the number of response categories of the items. Since the inventory has a six-point Likert response format, the number of b parameters estimated was five (b1 - b5). The b1 parameter estimated for an item is the ability (θ) level, which corresponds to the preference of the other five answer categories of the item to the first answer category, in other words, the choice of the second, third, fourth, fifth and sixth answer categories with a probability of 0.50. The b2 parameter is the ability (θ) level, which corresponds preferring the third, fourth, fifth and sixth answer categories with a probability of 0.50 instead of the first and second answer categories. The b3 parameter is the ability (θ) level, which corresponds to choosing the fourth, fifth and sixth answer categories with a probability of 0.50 instead of the first, second and third answer categories. With a similar logic, the b4 and b5 parameters also express the ability (θ) level, which corresponds to the preference of the relevant answer category and subsequent answer categories/category with a probability of 0.50 instead of the previous answer categories. When the item boundary parameter, that is, the b parameter values, are examined, it is seen that they mostly have negative values. Based on this, it can be said that the answers are mostly supported by the low level of the measured feature/trait ($\theta<0$) (Uyar et al., 2013).

In this context, when the levels of the feature/trait measured by the items according to IRT is examined, the arithmetic mean values of five b parameters estimated for each items vary between -3.47 (Sense Making Scale, item 27) and 0.30 (Curiosity Scale, item 2), and the median is -1.140. According to the IRT, the low levels of the features/traits measured by the items are an indication that the higher level response categories are selected, the higher levels of the features/traits measured by the items are also the indicators that the lower level response categories are selected. The average of the arithmetic means of the levels of the feature/trait measured by the items estimated within the scope of the study is -1.056. Usually the b parameter can take a value between ± 3, with probability 0.50 representing the required θ level of feature/trait for the approval of the item. A negative b value can be interpreted as the items are better at distinguishing those with a low level of the trait of interest from those with a moderate level (Flannery et al., 1995).

When the frequency distribution of the responses to the items is examined, it is seen that although the students prefer each of the answer options at varying rates, they generally choose the high-level response categories. For example, the distribution of the answers according to the response categories for the 46th item in the Mindful Agency Scale, of which the item score average is 4.23 according to the CTT is; 1 = 6 (0.60%), 2 = 24 (2.30%), 3 = 100 (9.50%), 4 = 187 (17.70%), 5 = 334 (31.70%), 6 = 403 (38.20%). A similar trend to this item was observed in the rest of the items.

When Table 5 is examined, another point that stands out is that some b1 and b2 parameters are less than -3. It was stated by Embretson and Reise (2000) that this may be due to the low number of respondents who preferred the first response categories of these items or the fact that the item could not accurately measure the desired feature/trait. Accordingly, when the distribution of

response categories is examined, it is seen that the students who prefer the first categories are much less than the other categories.

While introducing the scales of the CLARA Inventory, it was emphasized that the low or high score obtained from the "Learning Orientation Scale" reflects a rigid persistence in the sense of not deviating from what he/she knows at one end; and reflects a dependent fragility, a feeling of being vulnerable in the slightest challenging situation at the other. For this reason, while the highest and lowest values of the levels of the feature/trait measured by the items according to both theories were reported, the values of the "Learning Orientation Scale" were ignored in order not to be misleading.

In Table 6, descriptive statistics of the levels of the features/trait measured by CLARA-TR items, estimated based on the CTT and IRT, are presented.

Table 6. *Descriptive Statistics of the Levels of the Features/Traits Measured by CLARA-Tr Items Estimated According to CTT and IRT.*

Descriptive Statistics				
	CTT	Stand. Error	IRT	Stand. Error
Minimum	2.08		-4.28	
Maximum	5.64		1.87	
$\bar{X}$	4.21	0.11	-1.06	0.17
Median	4.24		-1.14	
SD	0.77		1.20	
Kurtosis	0.338	0.67	0.65	0.67
Skewness	-0.619	0.34	-0.30	0.34
Range	3.56		6.15	
Number of Items (k)	49		49	
Number of Students	1054		1054	

The correlation between the level of the features/traits measured by the items determined according to CTT and IRT was calculated with the Spearman Rank-Differences Correlation Coefficient and it was determined that there was a negative and highly significant relationship between these two values (r = -0.830, p <0.05). If individuals prefer higher response categories while answering the items, the item score average, i.e. the value of the level of the features/traits measured by the items, increases according to the CTT. According to the IRT on the other hand, the boundary location parameter value, which is accepted as the level of the feature/trait measured by the items, decreases. The boundary location parameter is the required feature/trait level for responders to react above the limit of a response category with a probability of 0.50 (Ostini & Nering, 2006), and when individuals prefer higher response categories, the boundary location parameter, or b parameter, takes lower values. According to this result, it can be interpreted that the feature/trait levels of the items determined according to the CTT and IRT are similar to each other, and this result is consistent with the previous study results in which polytomous items statistics based on two test theories are compared (Karakılıç, 2009; Koch, 1983; Nartgün, 2002; Uysal, 2015).

3.3.3. *Research question 3.3. findings & comments – relationship between the study group's learning power levels estimated by CLARA-Tr*

Another sub-question to be answered within the scope of the third research question is "Is there a relationship between the study group's learning power levels estimated by CLARA-Tr based on CTT and IRT? For this sub-question, the relationship between the scores obtained by the students from eight Scales, which together make up the learning power profile, based on CTT and IRT was examined with the Pearson Product-Moments Correlation.

In this context, the relationship between the eight Scales that constitute the CLARA-Tr Inventory, the levels of features/traits measured according to CTT and IRT was examined. Based on the CTT, the raw scores of the students from each scale were transformed into a 100-point system with a simple formulation. In doing so, firstly, the arithmetic mean of the answers given by the students to the items in each scale was taken. Then, the base score that could be obtained from an item was subtracted from this average, and finally, this score was divided by five and multiplied by 100. According to the IRT, for each scale the trait levels of the students measured with that scale were estimated according to the Graded Response Model. Students' estimated scale scores belonging to eight scales according to CTT and IRT are given in Appendix 3. Descriptive statistics of these scores are presented in Table 7.

Table 7. *Descriptive Statistics of the Students' Scores Estimated According to CTT and IRT.*

Descriptive Statistics					
Scale		CTT	Stand. Error	IRT	Stand. Error
Belonging	Minimum	0.00		-1.547	
	Maximum	100.00		1.547	
	$\bar{X}$	54.22	0.91	0.134	0.025
	Median	53.33		0.072	
	SD	29.43		0.819	
	Kurtosis	-1.03	0.151	-0.542	0.151
	Skewness	-0.53	0.075	0.003	0.075
	Range	100.00		3.094	
	Number of Items (k)	3		3	
	Number of Students	1054		1054	
Collaboration	Minimum	0,00		-1.547	
	Maximum	100		1.547	
	$\bar{X}$	67.12	0.64	0.461	0.175
	Median	66.67		0.444	
	SD	20.67		0.569	
	Kurtosis	0.17	0.151	0.178	0.151
	Skewness	-0,55	0.075	-0.263	0.075
	Range	100		3.094	
	Number of Items (k)	3		3	
	Number of Students	1054		1054	
Creativity	Minimum	0.00		-2.436	
	Maximum	100		2.436	
	$\bar{X}$	65.48	0.53	0.597	0.021
	Median	65.00		0.524	
	SD	17.11		0.691	
	Kurtosis	-0.38	0.151	0.153	0.151
	Skewness	-0.17	0.075	0.307	0.075
	Range	100		4.872	
	Number of Items (k)	8		8	
	Number of Students	1054		1054	

Table 7. *Continues*

Curiosity	Minimum	0.00		-2.175	
	Maximum	100		2.175	
	$\bar{X}$	67.34	0.58	0.623	0.022
	Median	70.00		0.607	
	SD	18.75		0.702	
	Kurtosis	-0.45	0.151	-0.277	0.151
	Skewness	-0.36	0.075	0.035	0.075
	Range	100		4.350	
	Number of Items (k)	6		6	
	Number of Students	1054		1054	
Hope & Optimism	Minimum	0.00		-1.547	
	Maximum	100		1.547	
	$\bar{X}$	70.49	0.70	0.553	0.020
	Median	73.33		0.548	
	SD	22.80		0.645	
	Kurtosis	-0.26	0.151	-0.261	0.151
	Skewness	-0.59	0.075	-0.197	0.075
	Range	100		3.094	
	Number of Items (k)	3		3	
	Number of Students	1054		1054	
Mindful Agency	Minimum	0.00		-2.542	
	Maximum	100.00		2.542	
	$\bar{X}$	68.11	0.50	0.709	0.021
	Median	68.89		0.675	
	SD	16.16		0.682	
	Kurtosis	0.00	0.151	0.417	0.151
	Skewness	-0.37	0.075	0.204	0.075
	Range	100.00		5.084	
	Number of Items (k)	9		9	
	Number of Students	1054		1054	
Orientation to Learning	Minimum	0,00		-2.637	
	Maximum	100		2.289	
	$\bar{X}$	49.10	0.46	0.033	0.20
	Median	48.89		0.000	
	SD	14.91		0.644	
	Kurtosis	0.01	0.151	0.823	0.151
	Skewness	0.16	0.075	0.092	0.075
	Range	100		4.926	
	Number of Items (k)	10		10	
	Number of Students	1054		1054	
Sense Making	Minimum	0,00		-2.315	
	Maximum	100		2.315	
	$\bar{X}$	74.12	0.39	0.897	0.017
	Median	74.29		0.846	
	SD	12.59		0.546	
	Kurtosis	0.91	0.151	0.932	0.151
	Skewness	0.45	0.075	0.150	0.075
	Range	100		4.630	
	Number of Items (k)	7		7	
	Number of Students	1054		1054	

When the values in Table 7 and Appendix 3 were examined together, it was determined that there is a similarity between the students' levels of the features/traits participating in the study, which were estimated based on both theories for all scales. It was observed that a student who got a lower score from a scale according to CTT had a similarly low score estimated according to IRT.

Within the scope of the third research question of the study, the relationship between the scores obtained by the students from eight scales estimated according to CTT and IRT was examined with Pearson Product-Moments Correlation Coefficient. Correlation values are presented in Table 8.

Table 8. *Relationship Between Students CLARA-Tr Scores Estimated from CTT and IRT.*

Scales (CTT / IRT)	n=1054	
	r	*p*
Belonging	0.992	0.000
Collaboration	0.991	0.000
Creativity	0.983	0.000
Curiosity	0.986	0.000
Hope & Optimism	0.987	0.000
Mindful Agency	0.979	0.000
Orientation to Learning	0.975	0.000
Sense Making	0.973	0.000

When Table 8 is examined, it is seen that there is a positive, high and significant (r = 0.973-0.992, p <.01) relationship between the scores of the students obtained from CLARA-Tr's eight scales estimated based on CTT and IRT. Based on these correlation coefficients, it can be inferred that the scores estimated according to both theories are similar and comparable.

3.3.4. *Research question 3.4. findings – reliability of clara-tr*

The last answer will be sought within the scope of the third research question is "How is the reliability of CLARA-Tr according to the Classical Test Theory and Item Response Theory?" For this sub-question; the reliability of the instrument was determined by calculating the Cronbach alpha internal consistency coefficient value according to the CTT and the marginal reliability coefficient according to the IRT.

Each reliability levels of the eight scales in the Turkish form of CLARA were examined both according to CTT and IRT. Calculated Cronbach alpha internal consistency coefficients and marginal reliability coefficients are summarized in Table 9.

Table 9. *Cronbach Alpha Internal Consistency & Marginal Reliability Coefficients of CLARA-Tr.*

Scale	Number of Items (k)	Cronbach Alpha	McDonald's omega	Marginal Reliability
Belonging	3	0.871	0.871	0.874
Hope & Optimism	3	0.833	0.869	0.873
Mindful Agency	9	0.812	0.813	0.842
Creativity	8	0.790	0.795	0.814
Curiosity	6	0.785	0.806	0.850
Sense Making	7	0.754	0.759	0.756
Orientation to Learning	10	0.742	0.741	0.834
Collaboration	3	0.730	0.734	0.721

When Table 9 was examined, it was seen that the Cronbach alpha reliability coefficient for all scales varied between 0.871 and 0.730, and McDonald's omega coefficients varied between 0.871 and 0.734. For scales, reliability coefficient values above 0.70 are accepted as high reliability levels (Nunnally, 1978; Özdamar, 2013). According to these values, it can be said

that all scales are consistent within themselves and have a high level of reliability according to CTT. When the marginal reliability coefficients estimated according to IRT are examined, it is seen that these values change between 0.874 and 0.721. The marginal reliability coefficient is defined as the arithmetic mean of the reliability coefficients estimated separately for the different levels of the measured psychological feature/trait (Thissen, 1991; Flannery et al., 1995). In this respect, the marginal reliability coefficient is accepted as a reliability coefficient calculated for the whole of a measurement tool. The high value of this coefficient is an indication that the results obtained from the measurement tool used are reliable. It is seen that the reliability coefficient values estimated according to both theories presented in Table 9 are quite high and similar to each other. These values can be interpreted as all eight scales of CLARA-Tr can make reliable measurements.

4. DISCUSSION and CONCLUSION

In this study, it was aimed to adapt Crick Learning for Resilient Agency (CLARA) Inventory to Turkish culture, also to analyze and compare the psychometric properties of the Inventory in the adaptation process according to the methods and techniques of both Classical Test Theory (CTT) and Item Response Theory (IRT). The results achieved are listed below in items.

1. It was determined that there is a positive, high and significant relationship between the scale scores obtained from the English and Turkish forms of CLARA's language equivalence applications. Based on this finding, it was accepted that linguistic equivalence was provided between the original form of CLARA and its Turkish form.

2. Two separate confirmatory factor analyses were conducted on the data obtained from the pilot application to determine whether the eight-scale original structure of the Inventory was also verified by Turkish university students. It was decided that the original factor structure of the Inventory was also verified in Turkish undergraduate students, and that data on learning power could be collected from university students in a valid and reliable manner with the Inventory.

3. Before starting the analysis to determine the psychometric properties of CLARA-Tr according to CTT and IRT, it was tested whether the data obtained as a result of the pilot application provided the assumptions of both theories. As a result of meeting the assumptions, analyses were carried out regarding the research questions.

4. The item discrimination index (item corrected total score correlation and a parameter) of the items of CLARA-Tr was estimated according to the Classical Test Theory and Item Response Theory, and a decent level discrimination index values were obtained according to both theories.

5. The levels of the features/traits (arithmetic mean and b parameter) measured by the items that constitute the Inventory were determined according to both test theories. It has been determined that the items generally measure the features/traits to be measured with the scales they belong to at a high level and all items have a relatively high approval rate, in other words, the participants have responded to high-level categories. The relationship between the levels of the features/traits measured by the items determined according to CTT and IRT was examined and a highly significant negative relationship was found. According to this result, it has been interpreted that the levels of the features/traits measured by the items determined according to CTT and IRT are similar to each other.

6. The relationship between the estimated scores, based on both test theories, of the undergraduates' obtained from eight scales, which together constitute the learning power profile, was examined. In this context, a high level of relationship was found between undergraduates' learning power levels predicted according to both theories, and from this point of view, it was concluded that the scores obtained from the two theories were similar.

7. The reliability of the scales that constitute the CLARA-Tr has been examined by calculating the Cronbach alpha internal consistency coefficient according to the CTT, and by the marginal reliability coefficient according to the IRT. It has been observed that the reliability coefficient values estimated according to both theories are quite high and similar to each other. As a result of these values, it was concluded that all eight scales of CLARA-Tr make reliable measurements.

8. It can be said that with CLARA-Tr, obtained by adapting CLARA, a valid and reliable tool has been provided to the Turkish literature to be used in the future studies.

The recommendations made as a result of the adaptation process and the comparisons made according to different test theories in this process are presented below in items.

1. Within the scope of this study, it is revealed that CLARA-Tr, whose psychometric properties were examined by adapting into Turkish culture, make valid and reliable measurements according to both test theories; and in the light of this result, it can be said that researchers who aim to reveal the undergraduate students' learning power profiles will be able to use the Inventory.

2. Whether the item parameters of CLARA-Tr show invariance between different samples according to both test theories can be discussed in a separate study.

3. In order to compare with the results of this study, the Inventory can be applied to samples of different sizes and different characteristics.

4. Within the scope of this study, it was determined that the values of the psychometric properties of the Inventory estimated according to both theories were similar. In this context, studies can be carried out on the basis of both theories as currently applied in scale development studies. On the other hand, it is recommended that researchers who want to reach more explanatory information at the item and test level should especially prefer the IRT. The fact that IRT gives different error estimates at different levels of the psychological feature/trait to be measured, and that items which give information with higher precision can be selected, will enable researchers to develop scales suitable for their purposes.

The CLARA learning power profile tool was designed to enable an individual learner to develop their capacity for self-leadership in learning which is a crucial 21st Century life competence (Sala et al., 2020). It was, at the same time, a deliberate attempt on the part of researchers to challenge the dominant ‘performativity’ discourse in educational assessment (Broadfoot, 1998). The accuracy, reliability and validity of the measurement model as reported here provides the foundation for this personal, social and political development, supported most effectively through coaching relationships. Since the first learning power model was developed in 2002 there has been significant user led demand for the tool which has been and practiced extensively in education, community and corporate contexts around the world, for example (Crick and Bentley, 2020).

However, it also brings with it the inherent challenge of forging pathways to impact for research outputs, moving beyond academia into digital learning analytics and also into practice led improvement in different contexts. Such pathways to impact require new business models which can integrate the differing requirements, funding mechanisms and lifecycles of research, policy, practice and commercial enterprise. The digital capability for the assessment tool, built on a data architecture which has a ‘single view of the learner’, uses one data point to provide rapid feedback to the individual, the team and the organisation as well as raw data for ongoing research. This is beyond the traditional capacity of a single research or educational institution and requires ethical quality assurance derived from a not for profit entity, funding for user services as well as digital entrepreneurship in a world which tends towards an individualist and reductionist ideology and practice. Twenty-one years of experience have led to the current

business model – which also provided the basis for this research study. The next steps are to take CLARA-Tr and explore whether and how it can add value in practice, through Work Integrated Learning Design (WILD) in Türkiye.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the author(s).

Authorship Contribution Statement

This study has been generated from the doctoral dissertation named "Adaptation of CLARA Inventory to Turkish Culture and Investigation of the Factors Affecting Learning Power with Hierarchical Linear Models" which was submitted by **Hasan Fehmi OZDEMIR** under the supervision of **Omer KUTLU** at Ankara University, Institute of Educational Sciences. The adapted Inventory was originally developed by **Ruth CRICK** and revised by **Shaofu HUANG** and **Ruth CRICK**. They have also contributed to the translation, online data gathering and writing the original draft processes.

Orcid

Hasan Fehmi Ozdemir https://orcid.org/0000-0002-0705-8032
Omer Kutlu https://orcid.org/0000-0003-4364-5629
Shaofu Huang https://orcid.org/0000-0002-9887-2434
Ruth Crick https://orcid.org/0000-0003-1082-9560

REFERENCES

Arthur, James & Crick, Ruth & Samuel, Elspeth & Wilson, Kenneth & Mcgettrick, Bart. (2006). *Character Education: The Formation of Virtues and Dispositions in 16-19 Year Olds with particular reference to the religious and spiritual*. West Conshohocken, PA: John Templeton Foundation. Retrieved from: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.7540&rep=rep1&type=pdf

Broadfoot, P. (1998). Records of Achievement and the Learning Society: a tale of two discourses. *Assessment in Education, 5*(3), 447-477. https://doi.org/10.1080/0969595980050307

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz, Ş., & Demirel, F. (2014). *Bilimsel Araştırma Yöntemleri* [*Scientific Research Methods*]. Pegem Akademi.

Carr, M., & Claxton, G. (2002). Tracking the development of learning dispositions. *Assessment in Education*, 9 (1), pp. 9-37. https://doi.org/10.1080/09695940220119148

Claxton, G. (1999). *Wise Up: The Challenge of Lifelong Learning*. Bloomsbury.

Cohen, R.J., & Swerdlik, M.E. (2010). *Psychological testing and assessment*. McGrawHill Companies.

Crick, R., & Bentley, J. (2020). Becoming a resilient organisation: integrating people and practice in infrastructure services. *International Journal of Sustainable Engineering, 13*(6), 423-440. https://doi.org/10.1080/19397038.2020.1750738

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Harcourt, New York.

Cronbach, L.J. (1960). *Essentials of psychological testing*. Harper Collins Publishers.

Deakin Crick, R. (2005). *Learning Power: Dynamic Assessment for Learning*. The Leader: Secondary Heads Association.

Deakin Crick, R. (2006). *Learning power in practice: A guide for teachers*. Paul Chapman.

Deakin Crick, R. (2007). Learning how to learn: the dynamic assessment of learning power. *Curriculum Journal*, *18*(2), 135-153. https://doi.org/10.1080/09585170701445947

Deakin Crick, R. (2009a). Assessment in Schools - Dispositions. In: McGaw, B., Peterson, P., & Baker, E. (eds.) *The International Encyclopedia of Education* (3rd Edition). Elsevier.

Deakin Crick, R. (2009b). Inquiry-based learning: reconciling the personal with the public in a democratic and archaeological pedagogy. *Curriculum Journal*, 20(1), 73-92. https://doi.org/10.1080/09585170902764021

Deakin Crick, R., Broadfoot, P., & Claxton, G. (2004). Developing an effective lifelong learning inventory: The ELLI Project. *Assessment in Education*, *11*(3), 248-272. https://doi.org/10.1080/0969594042000304582

Deakin Crick, R., & McCombs, B. (2006). The Assessment of Learner Centered Principles: An English Case Study. *Educational Research and Evaluation*, *12*(5), 423-444. https://doi.org/10.1080/13803610600697021

Deakin Crick, R., & Salway, A. (2006). *Locked Up Learning: learning power and young offenders*. ViTaL Partnerships Ltd.

Deakin Crick, R., & Wilson, K. (2005). Being a Learner: A Virtue for the 21st Century. *British Journal of Educational Studies*, *53*(5), 359-374. https://doi.org/10.1111/j.1467-8527.2005.00300.x

Deakin Crick, R., & Yu, G. (2008). Assessing learning dispositions: is the effective lifelong learning inventory valid and reliable as a measurement tool? *Educational Research*, *50*(4), 387-402. https://doi.org/10.1080/00131880802499886

Deakin Crick, R., Haigney, D., Huang, S., Coburn, T., & Goldspink, C. (2013). Learning power in the workplace: the effective lifelong learning inventory and its reliability and validity and implications for learning and development. *The International Journal of Human Resource Management*, *24*(11), 2255-2272. https://doi.org/10.1080/09585192.2012.725075

Deakin Crick, R., Huang, S., Shafi, A.A., & Goldspink, C. (2015). Developing resilient agency in learning: The internal structure of learning power. *British Journal of Educational Studies*, *63*(2), 121-160. https://doi.org/10.1080/00071005.2015.1006574

De Ayala, R.J. (2009). The theory and practice of item response theory. U.S.A.: Guilford Press.

Diker Coşkun, Y., & Demirel, M. (2012). Üniversite Öğrencilerinin Yaşam Boyu Öğrenme Eğilimleri [Lifelong Learning Dispositions of University Students]. *Hacettepe University Journal of Education*, 42, 108-120.

Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates, Inc.

Erkuş, A. (2013). *Davranış Bilimleri İçin Bilimsel Araştırma Süreci* [*Scientific Research Process for Behavioral Sciences*]. Seçkin Yayıncılık.

Flannery, W.P., Reise, S.P., & Widaman, F.K. (1995). An item response theory analysis of the general and academic scales of self-description questionnaire-II. *Journal of Research in Personality*, *29*(1), 168-188.

Gornall, S., Chambers, M.R., & Claxton, G. (2005). Building Learning Power in Action. TLO Ltd. Retrieved from: https://www.buildinglearningpower.com/wp-content/uploads/2014/11/BLPIA_A4_v1l-sample.pdf

Hambleton, R., & Swaminathan, R. (1985). *Fundamentals of Item Response Theory*. Sage Publishers.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory*. Sage.

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), 1-12.

Hambleton, R.K., Merenda, P.F., & Spielberger, C.D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Lawrence Erlbaum.

Harlen, W., & Deakin Crick, R. (2003a). *A systematic review of the impact of summative assessment and testing on pupils' motivation for learning*. Retrieved from: https://dspace.stir.ac.uk/bitstream/1893/19607/1/SysRevImpSummativeAssessment2002.pdf

Harlen, W., & Deakin Crick, R. (2003b). Testing and Motivation for Learning. *Assessment in Education*, *10*(2), 169-207. https://doi.org/10.1080/0969594032000121270

Hu, L.T., & Bentler, P.M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. http://dx.doi.org/10.1080/10705519909540118

International Test Commission, ITC. (2018). Guidelines for translating and adapting tests (Second edition). *International Journal of Testing*, *18*(2), 101-134.

Jöreskog, K.G., & Sörbom, D. (1993). *Lisrel 8: structural equation modeling with the simplis command language*. Erlbaum Associates Publishers.

Karakılıç, M. (2009). *Beden eğitimi dersi için hazırlanan tutum ölçeğinin psikometrik kuramlar açısından incelenmesi* [*Examination of the attitude scale prepared for the physical education lesson in terms of psychometric theories*] [Unpublished Doctoral Thesis]. Ankara University.

Kelecioğlu, H. (2001). Örtük Özellikler Teorisindeki b ve a Parametreleri İle Klasik Test Teorisindeki p ve r İstatistikleri Arasındaki İlişki [The relationship between b ve a parameters in latent trait theory and p and r statistiscs in classical test theory]. *Hacettepe University Journal of Education*, *20*(1), 104-110.

Kline, R.B. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.

Koch, W.R. (1983). Likert scaling using the graded response latent trait model. *Applied Psychological Measurement*, *7*(1), 15-32. https://doi.org/10.1177/014662168300700104

Masten, A. (2007). Resilience in developing systems: progress and promise as the fourth wave rises. *Developmental Psychopathology*, *19*(3), 921-930. https://doi.org/10.1017/S0954579407000442

MoNE (2014). Türkiye hayat boyu öğrenme strateji belgesi ve eylem planı 2014-2018 [Türkiye Lifelong Learning Strategy Document and Action Plan 2014-2018]. Retrieved from: http://www.resmigazete.gov.tr/eskiler/2014/07/20140716-8-1.pdf

MoNE (2018). *2023 Eğitim Vizyon Belgesi* [*2023 Education Vision Document*]. Retrieved from: http://2023vizyonu.meb.gov.tr/

Nartgün, Z. (2002). *Aynı Tutumu Ölçmeye Yönelik Likert Tipi Ölçek ile Metrik Ölçeğin Madde ve Ölçek Özelliklerinin Klasik Test Kuramı ve Örtük Özellikler Kuramına Göre İncelenmesi* [*The investigation of item and scale properties of likert type scale and metric scale measuring the same attitude according to classisical test theory and item response theory*] [Unpublished PhD Thesis]. Hacettepe University.

Nunnally, J.C. (1978). *Psychometric theory*. 2nd Edition, McGraw-Hill, New York.

Ostini, R., & Nering, M.L. (Eds.). (2006). *Polytomous item response theory models*. Sage Publication.

Özdamar, K. (2013). *Paket Programlar ile İstatistiksel Veri Analizi* [*Statistical Data Analysis with Package Programs*] (9. Edition). Nisan Kitabevi.

Özdemir, H.F., Toraman, Ç., & Kutlu, Ö. (2019). The use of polychoric and Pearson correlation matrices in the determination of construct validity of Likert type scales. *Turkish Journal of Education*, *8*(1), 180-195. https://doi.org/10.19128/turje.519235

Reise, S.P., & Yu, J. (1990). Parameter recovery in the graded response model using MULTILOG. *Journal of Educational Measurement*, *27*(2), 133-144. https://doi.org/10.1111/j.1745-3984.1990.tb00738.x

Rutter, M. (1985). Resilience in the face of adversity. Protective factors and resistance to psychiatric disorder. *British Journal of Psychiatry*, *147*(6), 598-611. https://doi.org/10.1192/bjp.147.6.598

Rutter, M. (2012). Resilience as a Dynamic Concept. *Development and Psychopathology*, 24 335-344. https://doi.org/10.1017/S0954579412000028

Sala, A., Punie, Y., Garkov, V., & Cabrera Giraldez, M. (2020). LifeComp: The European Framework for Personal, Social and Learning to Learn Key Competence. Luxembourg: Publications Office of the European Union. https://data.europa.eu/doi/10.2760/302967

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, *34*(4), 5-17. https://www.psychometricsociety.org/sites/main/files/file-attachments/mn17.pdf

Samejima, F. (1996). *Polychotomous responses and the test score*. Tennessee: The University of Tennessee.

Savaşır, I. (1994). Ölçek Uyarlamasındaki Sorunlar ve Bazı Çözüm Yolları [Problems in Scale Adaptation and Some Solutions]. *Turkish Journal of Psychology*, *9*(33), 27-32. https://www.psikolog.org.tr/tr/yayinlar/dergiler/1031828/tpd13004433199400000m000295.pdf

Scott, D., & Usher, R. (2011). *Researching education: Data, methods and theory in educational enquiry*. Continuum.

Sümer, N. (2000). Yapısal Eşitlik Modelleri: Temel Kavramlar ve Örnek Uygulamalar [Structural Equation Models: Basic Concepts and Applications]. *Turkish Psychological Articles*, *3*(6), 49-73. https://www.psikolog.org.tr/tr/yayinlar/dergiler/1031828/tpy1301996120000000m000225.pdf

Şimşek, Ö.F. (2007). *Yapısal Eşitlik Modellemesine Giriş: Temel İlkeler ve LISREL Uygulamaları* [*Introduction to Structural Equation Modeling: Fundamentals and Applications of LISREL*]. Ekinoks Yayınları.

The Assessment Reform Group. (2010, December). Retrieved from https://www.nuffieldfoundation.org/project/the-assessment-reform-group

Thissen, D. (1991). *Multilog user's guide*. Scientific Software

Tunca, N., Alkın Şahin, S., & Aydın, Ö. (2015). Öğretmen adaylarının yaşam boyu öğrenme eğilimleri. *Mersin Üniversitesi Eğitim Fakültesi Dergisi, 11*(2), 432-446. https://doi.org/10.17860/efd.92694

Uyar, Ş., Öztürk Gübeş, N., & Kelecioğlu H. (2013). PISA 2009 Tutum Anketi Madde Puanlarının Aşamalı Tepki Modeli ile İncelenmesi [Examinig the Pisa 2009 Attitude Questionnaire Item Scores with Using Graded Response Model]. *Journal of Research in Education and Teaching*, *2*(4), 125-134. http://www.jret.org/FileUpload/ks281142/File/15.uyar.pdf

Uysal, M. (2015). *Araştırma özyeterlik ölçeğinin psikometrik özelliklerinin klasik test kuramı ve madde tepki kuramına göre incelenmesi* [*An investigation of psychometric properties of research self-efficacy scale according to Classical Test Theory and Item Response Theory*] [Unpublished Master Thesis]. Gazi University.

Vieira, A.L. (2011). *Preparation of the analysis. Interactive LISREL in practice* (First Edition). Springer.

APPENDIX

Appendix 1. An example of the MULTILOG program output (Belonging Scale)

MULTILOG--FOR MULTIPLE CATEGORICAL ITEM RESPONSE DATA--VERSION 7.0.3

MULTILOG for Windows 7.00.2327.2

Created on: 19 September 2018, 12:55:24

```
>PROBLEM RANDOM,

        INDIVIDUAL,

        DATA = 'C:\Users\kullanici\Desktop\ait\ait.DAT',

        NITEMS = 3,

        NGROUPS = 1,

        NEXAMINEES = 1054,

        NCHARS = 4;


  DATA FILE NAME IS
        C:\USERS\KULLANICI\DESKTOP\AIT\AIT.DAT


TYPE OF INPUT:
   INDIVIDUAL RESPONSE VECTORS


>TEST ALL,

      GRADED,

      NC = (6(0)3);

NUMBER OF CODES  6
```

123456
VECTOR OF CATEGORIES FOR CODE=1

111
VECTOR OF CATEGORIES FOR CODE=2

222
VECTOR OF CATEGORIES FOR CODE=3

333
VECTOR OF CATEGORIES FOR CODE=4

444
VECTOR OF CATEGORIES FOR CODE=5

555
VECTOR OF CATEGORIES FOR CODE=6

666
(4A1,T5,3A1)

MULTILOG--FOR MULTIPLE CATEGORICAL ITEM RESPONSE DATA--VERSION 7.0.3
MULTILOG for Windows 7.00.2327.2
Created on: 19 September 2018, 12:55:24

DATA PARAMETERS:
NUMBER OF LINES IN THE DATA FILE: 1054
NUMBER OF CATEGORICAL-RESPONSE ITEMS: 3
NUMBER OF CONTINUOUS-RESPONSE ITEMS, AND/OR GROUPS: 1
TOTAL NUMBER OF "ITEMS" (INCLUDING GROUPS): 4
NUMBER OF CHARACTERS IN ID FIELDS: 4
MAXIMUM NUMBER OF RESPONSE-CODES FOR ANY ITEM: 6
THE MISSING VALUE CODE FOR CONTINUOUS DATA: 9.0000
THE DATA WILL BE STORED IN MEMORY

ESTIMATION PARAMETERS:

THE ITEMS WILL BE CALIBRATED--
BY MARGINAL MAXIMUM LIKELIHOOD ESTIMATION
MAXIMUM NUMBER OF EM CYCLES PERMITTED: 25
NUMBER OF PARAMETER-SEGMENTS USED IS: 3
NUMBER OF FREE PARAMETERS IS: 18
MAXIMUM NUMBER OF M-STEP ITERATIONS IS 4 TIMES
THE NUMBER OF PARAMETERS IN THE SEGMENT
THE M-STEP CONVERGENCE CRITERION IS: 0.000100
THE EM-CYCLE CONVERGENCE CRITERION IS: 0.001000
THE RK CONTROL PARAMETER (FOR THE M-STEPS) IS: 0.9000
THE RM CONTROL PARAMETER (FOR THE M-STEPS) IS: 1.0000
THE MAXIMUM ACCELERATION PERMITTED IS: 0.0000
THETA-GROUP LOCATIONS WILL REMAIN UNCHANGED

QUADRATURE POINTS FOR MML,
AT THETA:
-4.500
-4.000
-3.500
-3.000
-2.500
-2.000
-1.500
-1.000
-0.500
0.000
0.500
1.000
1.500
2.000
2.500
3.000
3.500
4.000
4.500

MULTILOG for Windows 7.00.2327.2

READING DATA...

KEY-

CODE CATEGORY

1 111
2 222
3 333
4 444
5 555
6 666

FORMAT FOR DATA-

(4A1,T5,3A1)

FIRST OBSERVATION AS READ-

ID 0001
ITEMS 555
NORML 0.000

FINISHED CYCLE 25
MAXIMUM INTERCYCLE PARAMETER CHANGE= 0.00344 P(7)

ITEM SUMMARY

MULTILOG for Windows 7.00.2327.2

ITEM 1: 6 GRADED CATEGORIES
P(#) ESTIMATE (S.E.)
A 1 4.12 (0.17)
B(1) 2 -1.38 (0.06)

B(2) 3 -0.68 (0.03)
B(3) 4 -0.14 (0.03)
B(4) 5 0.18 (0.03)
B(5) 6 0.72 (0.04)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)
-3.0 - -1.6 0.021 0.048 0.109 0.245 0.538 1.131 2.165 3.499
-1.4 - 0.0 4.406 4.409 4.279 4.602 4.825 4.837 5.043 5.202
 0.2 - 1.6 5.090 4.850 4.754 4.311 3.126 1.819 0.918 0.430
 1.8 - 3.0 0.194 0.086 0.038 0.017 0.007 0.003 0.001

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN
CATEGORY(K): 1 2 3 4 5 6
OBS. FREQ. 104 171 201 126 192 260
OBS. PROP. 0.0987 0.1622 0.1907 0.1195 0.1822 0.2467
EXP. PROP. 0.1031 0.1627 0.1825 0.1176 0.1799 0.2541

ITEM 2: 6 GRADED CATEGORIES
P(#) ESTIMATE (S.E.)
A 7 4.46 (0.20)
B(1) 8 -1.35 (0.05)
B(2) 9 -0.76 (0.04)
B(3) 10 -0.16 (0.03)
B(4) 11 0.22 (0.03)
B(5) 12 0.70 (0.04)
@THETA: INFORMATION: (Theta values increase in steps of 0.2)
-3.0 - -1.6 0.013 0.031 0.074 0.180 0.428 0.981 2.076 3.724
-1.4 - 0.0 5.105 5.336 5.279 5.540 5.383 5.300 5.731 5.967
 0.2 - 1.6 5.944 5.739 5.606 4.927 3.317 1.758 0.810 0.349
 1.8 - 3.0 0.146 0.060 0.025 0.010 0.004 0.002 0.001

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN
CATEGORY(K): 1 2 3 4 5 6
OBS. FREQ. 106 141 218 152 174 263
OBS. PROP. 0.1006 0.1338 0.2068 0.1442 0.1651 0.2495
EXP. PROP. 0.1053 0.1353 0.1996 0.1404 0.1622 0.2571

ITEM 3: 6 GRADED CATEGORIES
P(#) ESTIMATE (S.E.)
A 13 2.29 (0.11)
B(1) 14 -1.51 (0.08)
B(2) 15 -0.49 (0.05)
B(3) 16 0.21 (0.05)
B(4) 17 0.66 (0.05)
B(5) 18 1.27 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)
-3.0 - -1.6 0.162 0.246 0.369 0.535 0.746 0.982 1.204 1.366
-1.4 - 0.0 1.444 1.461 1.467 1.495 1.540 1.579 1.605 1.626
0.2 - 1.6 1.647 1.660 1.659 1.641 1.602 1.524 1.386 1.184
1.8 - 3.0 0.943 0.704 0.500 0.342 0.227 0.149 0.096

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN
CATEGORY(K): 1 2 3 4 5 6
OBS. FREQ. 120 252 243 142 140 157
OBS. PROP. 0.1139 0.2391 0.2306 0.1347 0.1328 0.1490
EXP. PROP. 0.1158 0.2314 0.2198 0.1344 0.1422 0.1564

ITEM 4: GRP1, N[MU: 0.00 SIGMA: 1.00]
P(#);(S.E.): 20; (0.00) 21; (0.00)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)
-3.0 - -1.6 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000
-1.4 - 0.0 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000
0.2 - 1.6 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000
1.8 - 3.0 1.000 1.000 1.000 1.000 1.000 1.000 1.000

TOTAL TEST INFORMATION

@THETA: INFORMATION:
-3.0 - -1.6 1.195 1.325 1.552 1.960 2.712 4.094 6.446 9.589
-1.4 - 0.0 11.956 12.206 12.025 12.637 12.749 12.717 13.379 13.795

0.2 - 1.6 13.681 13.249 13.019 11.879 9.045 6.101 4.114 2.962
1.8 - 3.0 2.283 1.851 1.563 1.369 1.239 1.154 1.098

@THETA: POSTERIOR STANDARD DEVIATION:
-3.0 - -1.6 0.915 0.869 0.803 0.714 0.607 0.494 0.394 0.323
-1.4 - 0.0 0.289 0.286 0.288 0.281 0.280 0.280 0.273 0.269
0.2 - 1.6 0.270 0.275 0.277 0.290 0.333 0.405 0.493 0.581
1.8 - 3.0 0.662 0.735 0.800 0.855 0.898 0.931 0.954
MARGINAL RELIABILITY: 0.8741

NEGATIVE TWICE THE LOGLIKELIHOOD= -5344.0
(CHI-SQUARE FOR SEVERAL TIMES MORE EXAMINEES THAN CELLS)

NORMAL PROGRAM TERMINATION

START DATE: 09-19-2018
START TIME: 12:58:28
END TIME: 12:58:29

Appendix 2. Sample scale items of CLARA and CLARA-Tr

Mindful Agency

I know I can find a way of solving a problem if I have enough time to think.
(Düşünmek için yeterli zamanım olursa, karşılaştığım sorunu çözmenin bir yolunu bulabilirim.)
I think about everything that I will need before I begin a task.
(Bir işe girişmeden önce ihtiyaç duyacağım her şey hakkında düşünürüm.)

Hope and Optimism

I know I am changing and growing over time.
(Zamanla değiştiğimi ve geliştiğimi biliyorum.)
I am getting better at learning all the time.
(Öğrenme işinde sürekli daha iyiye gidiyorum.)

Sense Making

I make connections between what I am learning and what I have learned before.
(Yeni öğrendiğim şeylerle, önceden öğrendiklerim arasında bağlantı kurarım.)
I often look back and think about what I have learned.
(Öğrenmiş olduğum şeyler hakkında sıkça geçmişi hatırlar ve düşünürüm.)

Creativity

Sometimes good ideas just come into my head.

(Bazen, güzel fikirler ansızın aklıma geliverir.)

I tend to use my imagination to help me learn.

(Öğrenmeme yardımcı olması için hayal gücümü kullanma eğilimindeyimdir.)

Curiosity

I prefer learning something when I have to try really hard to understand it.

(Gerçekten çok çaba harcayarak anlayabileceğim şeyleri öğrenmeyi tercih ederim.)

I am more stimulated by interesting questions than easy answers.

(İlginç sorular, kolay cevaplara göre beni daha çok teşvik eder.)

Collaboration

I enjoy solving problems together with other people.

(Sorunları diğer insanlarla birlikte çözmekten hoşlanırım.)

I find it helps me to learn if I can talk about it with colleagues.

(Arkadaşlarımla, zorlayıcı sorunlar hakkında ayrıntılı bir şekilde tartışmayı severim.)

Belonging

There is at least one person close to me who has helped me to learn.

(Ben öğrenirken yardım etmiş olan bana yakın en az bir kişi var.)

I have at least one person close to me who I can turn to for guidance in my learning.

(Öğrenirken beni yönlendirmesi için başvurabileceğim, bana yakın en az bir kişi var.)

Orientation to Learning

I find it difficult to know what to do when I get stuck.

(Bir konuya takılıp kaldığımda ne yapacağımı bilmekte zorlanırım.)

Because I dislike feelings of confusion and uncertainty I generally steer clear of learning something new.

(Kafa karışıklığı ve belirsizlik duygularını sevmediğimden, genellikle yeni bir şey öğrenmekten kaçınırım.)

Appendix 3. Students' estimated scale scores belonging to eight scales according to CTT and IRT

Appendix 3. has been given as a separate document due to the number of pages.