Center for Open Access in Science ▪ https://www.centerprode.com/ojer.html
Open Journal for Educational Research, 2024, *8*(1), 1-16.
ISSN (Online) 2560-5313 ▪ https://doi.org/10.32591/coas.ojer.0801.01001e

Development and Validation of the Tablet-Based ERAL Nonverbal Intelligence Test 5-17 (ERAL-NIT)

Ercümend Ersanli

Ondokuz Mayıs University, Samsun, TURKEY
Faculty of Health Sciences

Ali Kılıçarslan

Alparslan Middle School, Tokat, TURKEY

Received: 9 April 2024 ▪ Revised: 2 July 2024 ▪ Accepted: 2 August 2024

Abstract

Intelligence has been extensively explored across various disciplines such as psychology, cognitive science, and neurology. Countless scholars have delved into understanding why certain individuals exhibit higher mental acuity and knowledge. Consequently, numerous studies aim to unveil the essence of intelligence and gauge human cognitive capacity. This study introduces and validates the ERAL Nonverbal Intelligence Test (ERAL-NIT) for children aged 5 to 17. Unlike many existing tests, ERAL-NIT is nonverbal, accommodating diverse linguistic backgrounds. Moreover, it incorporates parental assessments, enhancing its holistic approach. Administered via tablet with specialized software, ERAL-NIT provides comprehensive reports to practitioners and families. Its development integrates analyses of hemispheric specialization and executive functions of different brain areas. Content validity is ensured through expert input, while criterion validity is established by comparing results with established assessments and medical diagnoses. ERAL-NIT yields insights into pervasive developmental disorders, specific learning disabilities, attention deficits, and language disorders, thereby enriching diagnostic practices.

Keywords: nonverbal intelligence test, test development, test validation, test construction, neuro-cognitive IQ test.

1. Introduction

There has always been hot debate on the definition of intelligence. The roots of the term go back to ancient Greeks and Romans. First Cicero coined the word 'intelligence', and it was used as a man's mental capacity and intellectual abilities. Since those times, various practitioners have attached different meanings to the term. Although there has never been consensus on what constitutes intelligence or how to operationalize it, a variety of definitions have been suggested by different scholars. Mainly those definitions can be grouped under three categories: (i) the cognitive capacity to learn, reason, think and solve novel problems (Anderson, 2006; Bingham, 1937; Herrnstein & Murray, 2010), (ii) the total knowledge a person has acquired (Henmon, 1921), and (iii) the ability to adapt to new situations and to the environment successfully (Anastasi, 1992; Gardner, 1993; *An Interview with Dr. Simonton*, 2003; Pintner, 1921).

Correspondence: Ercümend Ersanli, Ondokuz Mayıs University, Faculty of Health Sciences, Samsun, TURKEY.

In a similar vein, Ackerman's (1996) meta theory of intelligence distinguishes between intelligence as knowledge and intelligence as process. His model suggests that intelligence is not a single one-dimensional construct, but rather consists of many factors. Ackerman suggests that intelligence as knowledge refers to people's existing knowledge structures, whereas intelligence as process refers to people's abilities to process information.

These various understandings of intelligence are particularly effective at demonstrating the need for different types of measures. There are multiple underlying theoretical constructs beneath those instruments. It is out of the scope of this study to discuss those theories fully. Yet, a brief overview of some of the major theories of intelligence is believed to shed light into the theoretical part of the current study.

General intelligence theories are classified under two categories; single factor theories and multiple factor theories. Spearman (1927) suggests that there is one factor or mental attribute, which he calls g or general intelligence, but it also requires some specific abilities. A current version of the general plus specific abilities theory is John Carroll's (Carroll, 1993) three-stratum (3S) theory of intelligence identifying abilities that correspond to surface level characteristics of mental tasks, broad abilities, highly abstract, and general abilities that affect all tasks requiring cognitive ability (Benson et al., 2018).

Today, one factor theories of intelligence are challenged by Multiple factor theories. The most widely accepted view is that intelligence has many facets and consists of a hierarchy of abilities. Sternberg (1985) in his Triarchic theory of intelligence suggests a cognitive process approach to understand intelligence. According to him it has three parts; analytic part- it involves mental processes and the ability to think abstractly, creative experiential part- coping with new experiences, automaticity in thinking, problem solving, and practical part- adaptation to culture and tacit knowledge that is learned in everyday life. Similarly, Sattler (2001) believes that there are many independent faculties that make up intelligence. Years ago, Thurstone (1943) listed seven distinct major mental abilities that make up intelligence as verbal comprehension, memory, and reasoning, ability to visualize spatial relationships, numerical ability, word fluency, and perceptual speed.

In the light of the above given theories of intelligence, there have been consistent efforts by researchers to build a single theory of intelligence. Yet, it is crystal clear that intelligence is a multidimensional concept; it is not just about measuring how good you are at certain aspects, but "...[it] correlates with other aspects of a person such as personality or motivation, and these factors are likely to make a difference to education and life outcomes, too" (Ball, 2018: 34).

1.1 *Measures of intelligence*

Based on the evidence that there are indeed different theories of intelligence, many different measures of intelligence have emerged in time. However, according to Naglieri, Das and Jarman, early measures of intelligence are too narrow and limited in their measurement of intelligence (Naglieri & Prewett, 1990). They mainly measure verbal ability/achievement, nonverbal functioning, memory functioning, and quantitative abilities. One major shortcoming of such measures is that they ignore human cognitive processing. For instance, Weschsler scale is based on ability and achievement, not process (Rijumol et al., 2010). Therefore, Aleksandr Luria, a neuropsychologist, proposed a model of cognitive measurement, which is called the Luria Model. In this model the cognitive abilities are divided into three functional units. The first is the maintenance of arousal and attention level. The second one is coding. It controls the input, recalling and storing of information. The third and the last of them is planning. It refers to programming, regulating, and verifying the cognitive activities. What is unique in this model is that although each of these units in the brain is responsible for particular functioning, they must

function together as a whole to operate properly. They function simultaneously and in an interconnected way on many areas of the brain as a whole. Thus, the model contrasts to approaches which assume localization (Henmon, 1921).

Together with the advances in the theory of intelligence it is inevitable that there is a shift from abilities and achievement to information processing. 1960s witnessed the examination of higher order mental processes described as the "cognitive revolution." One of the first attempts to move to a more comprehensive cognitive measurement is the Kaufman Assessment Battery for Children (K-ABC). Actually, the measurement is based on the Luria Model and cerebral specialization. It examines simultaneous and sequential processing. Yet, it falls short in encompassing all human cognitive processes such as planning and attention (Das, 1992).

Later, Naglieri and Das went well beyond traditional views of intelligence and proposed a cognitive information processing model (PASS) based on the neuropsychological principles initially suggested by Luria (Naglieri, 1999). The PASS model involves planning, attention, simultaneous, and successive cognitive processes as the building blocks of intelligence. This model defines intelligence as the sum of all these cognitive processes. It has both a neurological and cognitive basis. The Cognitive Assessment System (CAS) is based on PASS theory. CAS is defined as an individually administered battery measuring the cognitive processes; planning, attention, simultaneous and successive processes suggested by PASS theory (Kranzler & Keith, 1999). It consists of 12 subtests each of which refers to the PASS cognitive processes.

The development of diverse intelligence tests, some of which are outlined above, is mainly a consequence of neuropsychologists' tendency to use them in their neuropsychological assessments. As Gansler suggest the results of neuropsychological test batteries and intelligence tests often reveal similar results. According to them, one possible reason of the correlation in the results may be "the rise of cognitive neuroscience approaches to the study of intelligence" (Gansler et al., 2017: 2).

Cognitive psychology emerged as a reaction to Behavioristic psychology and it focuses on the human mind and all aspects of the human information processing. Scholars claim that cognition involves many aspects such as "perception, attention, categorization, learning, and memory, thinking, decision making, problem solving, and language use" (Beller & Bender, 2010; Medin, 2004). In these respects, ERAL-NIT adopts a cognitive perspective.

Besides, it is known that intelligence tests aim to measure "cognitive abilities as opposed to irrelevant factors related to culture or language" (Hooper & Bell, 2006). Therefore, ERAL-NIT is developed as a culture-fair cognitive IQ test. However, there are many assumptions on the relationship between culture and cognition. Bender and Beller state that "[cognitive] processing is independent of context or people's cultural background" (Bender & Beller, 2013: 44). Despite the universal aspects of cognition, they claim that "people do not simply reason, they *learn* to reason" (p. 47). This means that other people in one's societal environment have an effect on the cognition of that person at least to some extent. They state that the cognitive processes affected by societal influences are broad, ranging from visual perception to spatial cognition. In a similar vein, it is highlighted the positive relationship between cognitive abilities and cultural effects (Haun et al., 2006). Boroditsky and Gaby in their study on the cognitive conception *time* demonstrate that it can differ across cultures (Boroditsky & Gaby, 2010). Beller and Bender investigate *numerical cognition* and the effects of culture on it (Beller & Bender, 2010) and explore the *spatial* and *temporal cognition* across different cultures and identify culture-specific preferences. Haun and his colleagues investigate *spatial cognition* across cultures (Haun et al., 2006). The results indicate a difference not only in preference but also in competence. In a similar study, the *spatial cognition* and find out that it varies across cultures (Majid et al., 2004). In another study (Masuda & Nisbett, 2006), *perception* and *cognition* was investigated and the findings reveal that there are cultural variations in basic perceptual and cognitive processes.

Norenzayan and Nisbett investigate culture and causal cognition and find differences in the cognitive abilities across cultures (Norenzayan & Nisbett, 2000).

Based on the findings of above given studies, it is clear that culture is inherently embedded in cognitive processes of individuals. To this end, ERAL-NIT aims to minimize its effect on the cognitive processes of individuals by not involving any cultural element in its items. That is, in order to make the intelligence scores to be as accurate and culturally fair as possible, the test items do not reflect any cultural elements.

ERAL-NIT is also affected by neurological brain-based studies. It is believed that the frontal, parietal, temporal and occipital lobes of the brain function in an interconnected way to operate as a whole and each of them is responsible for certain mental higher order processes. In the current study, the researchers have identified 16 subareas for those higher order executive functions; receptive language, gross motor movement, numeric ability, attention, convergent thinking, figure-ground perception, expressive language, fine motor movement, social ability, memory, divergent thinking, visual manipulation, concept knowledge, visual perception, reasoning ability, and part-whole relationship.

The literature suggests that receptive and expressive language is usually located in the left hemisphere, at the back part of the temporal lobe. Gross and fine motor movements include cerebellum (the development of fine motor skills plays a crucial role in school readiness). Frontal lobes control most cognitive functions. Reasoning ability and Numeric ability are located in frontal lobes, frontal lobes also appear to take charge of the brain's Attention ability, it controls relevant parts of the visual cortex, which receives sensory input. Temporal lobe region plays a major role in maintaining social ability. Memories are formed and stored in hippocampus, which is located in brain's temporal lobe. Hippocampus is also known to be responsible for creating and storing Concept knowledge. Occipital lobe is highly important in figure-ground perception, visual manipulation, visual perception and part-whole relationship. Parietal lobe region also plays a major role in these functions, it is the primary sensory area in which all of the sensory processing starts in the brain. It concerns primarily with the visual and spatial system. Divergent and Convergent thinking, the two processes central to cognition, refer to creative thinking and logical thinking respectively. It has been found out that there is close interaction between both hemispheres and the central parietal areas of both hemispheres are activated in those who display good performance in divergent thinking, and who use their imagination more effectively. Besides, convergent thinking is primarily right sided (Campbell & DeJong, 2005; *Stroke and cerebrovascular diseases*, 2019).

As can be seen, ERAL-NIT gives a clear profile of the cognitive capacity of the individuals.

ERAL-NIT also gives promising results in the identification of pervasive developmental disorder (PDD), specific learning disability (SLD), attention deficit disorder (ADD), and language use disorder (LUD).

1.2 *Justification for the study*

There are many reasons in developing ERAL-NIT; the first of them is that it is developed for individuals between the ages of 5 to 17. Especially in this age period individuals need guidance by their parents or teachers to reach their full potentials. ERAL-NIT serves as an effective measurement in this respect.

Additionally, although currently, many IQ tests are used in the world, most of these tests are verbal scales (Weschler 5, CAS, etc.). This may create a problem for the participants who are not literate or who cannot use the language in an effective way. Moreover, the rise in the

number of individuals from diverse cultural and linguistic backgrounds mainly because of the minority groups and immigrant population in all over the world create a need for nonverbal intelligence tests to provide fair assessment. To overcome this problem ERAL-NIT is a nonverbal test. Naglieri and Prewett suggest that in order to make a nonverbal intelligence test accurate and valid, it has to give a complete description of the cognitive processes of individuals who display physical limitations, language disorders, etc. as comprehensive as those of normal individuals (Naglieri & Das, 1990). In this respect, ERAL-NIT can be considered as valid and accurate since its results are also parallel with diagnoses such as PDD, SLD, ADD, and LUD.

Besides, ERAL-NIT does not merely evaluate individual performance of the participants. It also investigates parental evaluations and compares them with the child's overall performance. This gives way to parents to evaluate the accuracy of their own insights about their children. This allows the child to grow up in a healthier environment. As an initial step, parents state their own insights about general and more specific performance-based characteristics of their children (this takes maximum five minutes). At the end, in the light of the findings received from ERAL-NIT, it becomes clear how consistent the parents are in their evaluations about their children. The test shows parents whether their evaluations for the child's specific and general performances are compatible. The test compares the actual performances of the children with the parental evaluations on a chart and if there is any inconsistency, the test enables practitioners to make some suggestions for the parents.

Another reason beneath the need for the development of such a test is that unlike its many counterparts, ERAL-NIT takes rather a short period of time to conduct; approximately 25 minutes. This is especially important since it also encompasses early age groups who do have rather short attention spans.

It is a known fact that no matter how high the intercoder reliability of such tests, there is always a risk for mistakes emerging from practitioners. To minimize such problems, ERAL-NIT has been developed to be implemented on a tablet as a software program, which makes most of the data entry itself and provides a detailed report on the performance of the participants at the end of the evaluation process.

Furthermore, the conduction of ERAL-NIT does not require a clinic atmosphere and the participants who take this test do not feel a sense of failure.

Finally, most tests evaluate the individuals' performances at a particular time ignoring all the other possible factors such as feeling of hunger, security, etc. However, during the process there are many factors which may have an effect on the individuals' performance. Intelligence or mental capacity cannot be measured thoroughly by just focusing on what the individual can do within a certain time by himself or herself. In this vein, ERAL-NIT takes all these factors into consideration; the practitioners first observe the behaviors of the test takers (restless eye movement, shaking hands, swinging head, and body, etc.) and may ask some specific questions to the parents such as "Is she/he hungry?", "Is she/he tired?" etc. These observations and the responses they gathered have minimum effect on the total findings of the test.

1.2.1 *Purpose of the study*

The purpose of this study is to develop a neurocognitive IQ test for Turkish young people between 5 and 17 years old, and to investigate its validity by comparing the scores on that test with the scores on another well-known IQ test; CAS. In short, it is aimed to share the development process of this new test; ERAL-NIT and initial validation endeavors with researchers and practitioners. It is assumed that psychiatrists, psychologists, child development experts, and psychological counselors in advising families, and school age children and teenagers will use information from this study.

1.2.2 *Research questions*

To what extent is there evidence to support the content validity of ERAL-NIT (in the Turkish context)?

To what extent is there evidence to support the criterion validity of ERAL–NIT (in the Turkish context)?

2. Method

2.1 *Participants*

ERAL-NIT is developed using a nationally representative sample of 642 participants between the ages 5 to 17. Examinees in the sample are selected from 12 regions determined by the Turkish National Institute of Statistics' classification of statistical regional units. By this way, it is believed that the sample is representative of the general population. One city in each of these regions is chosen as the representative. Thus, a total of 12 cities are determined. Both the private and state school students in those cities are invited to join the research. Five schools are randomly selected in each city. Table 1 shows the frequencies of participants by cities.

Table 1. Frequencies of participants by cities

Cities	F	% of Total	Cumulative %
BURSA	82	12.8 %	12.8 %
SAMSUN	64	10.0 %	22.7 %
TRABZON	28	4.4 %	27.1 %
ANKARA	63	9.8 %	36.9 %
İSTANBUL	110	17.1 %	54.0 %
ANTALYA	82	12.8 %	66.8 %
GAZİANTEP	67	10.4 %	77.3 %
SİVAS	24	3.7 %	81.0 %
TEKİRDAĞ	34	5.3 %	86.3 %
İZMİR	41	6.4 %	92.7 %
MALATYA	30	4.7 %	97.4 %
ERZURUM	17	2.6 %	100.0 %

A stratified random sampling is used to select participants so that the size of the sample is proportional to the number of participants from age 5 to 17 in the population. Table 2 reveals the frequencies of participants by ages. The age distribution of the sample has a mean of 10.73 with a standard deviation of 3.60.

Table 2. Frequencies of participants by ages

Age	F	% of Total	Cumulative %
5	53	8.3 %	8.3 %
6	48	7.5 %	15.7 %
7	47	7.3 %	23.1 %
8	65	10.1 %	33.2 %
9	36	5.6 %	38.8 %
10	47	7.3 %	46.1 %
11	66	10.3 %	56.4 %
12	77	12.0 %	68.4 %
13	46	7.2 %	75.5 %
14	37	5.8 %	81.3 %
15	32	5.0 %	86.3 %
16	52	8.1 %	94.4 %
17	36	5.6 %	100.0 %

The number of participants reveal a balanced distribution in terms of gender. While 50.47% (n=324) of the participants are male, remaining 49.53% (n=318) are female.

Out of the total number of the participants parents of 574 are involved into the study. Before the conduction of the test, parents are asked to state their evaluations about their children's general and specific performances. Their comments are compared with the findings gathered from ERAL-NIT. Finally, apart from these, another group of 243 participants who have hospital diagnoses with SLD, PDD, LUD, and LA are also enrolled into the study.

2.1.1 *Instrument*

In this study the Cognitive Assessment System (CAS) is used to ensure the criterion validity of ERAL-NIT. The standard CAS battery consists of 12 subtests. The PASS processes are reflected in four scales and their respective subtests: Planning, Attention, Simultaneous, and Successive. A standard score is provided for each cognitive process along with a full-scale score. The internal reliability coefficients are high, Planning=.88; Attention=.88; Simultaneous=.93; Successive=.93; and Full Scale=.96 (Das et al., 1994; Naglieri, 1999).

Besides, the hospital diagnoses for SLD, PDD, LUD positive individuals are used to increase criterion validity of ERAL-NIT.

2.2 *Procedures*

2.2.1 *Development of ERAL-NIT*

As an initial step an approval is received from the research ethics committee from a state university. Before starting the development of ERAL-NIT, a group of 36 experts, including psychiatrists, psychologists, psychological counselors and child development specialists is asked what kind of a test is needed in the field of intelligence testing. The answers received from them are analyzed using Content Analysis Method and some features of ERAL-NIT are decided accordingly. As a result of expert opinions, it is decided that the test is non-verbal, it is a software program applied on a tablet, does most of the scoring, and the report itself, it does not include cultural elements, it is applied in a short time, it gives results about pervasive developmental disorder, special learning difficulty, attention deficit disorder, and language use disorder and it involves families.

The other stage is the generation of the items. With the help of 18 experts (three experts from the field of child development, three experts working in the field of test development, three psychologists, three psychological counselors, three psychiatrists, and three neurologists), items are generated in line with the neurological brain-based studies. By this way the content validity of the items is ensured. "Drawing a picture in which there is a human being, a tree, etc.," and "finding the rote to exit in a labyrinth" are some of the items in the test. These items provide raw score for the 16 sub areas for the higher order mental processes given before, for the total IQ score, and for the frontal, temporal, occipital and parietal lobe areas. The raw score is calculated based on the abilities of different ages stated in the related literature. The literature depicts what a normally developing five-year old's mental capacity can do, the raw scores are determined in parallel to these arguments. However, if the child performs better than expected she/he receives a higher raw score. In a similar vein, if the child's performance is below her/his age she/he receives a lower raw score. The IQ score, the scores for frontal, temporal, occipital, and parietal lobe areas and the scores for the 16 subareas of brain's mental functioning are calculated by the multiplication of the coefficients and these raw scores. It should be noted that the brain a unique organ and no matter the areas in the brain have certain functions they need to work as a whole, in a parallel way to operate thoroughly. Thus, all the raw scores have an effect on IQ score, the scores for the four lobe areas and the scores for the 16 subareas of brain's mental functioning via the coefficients.

Then, in order to calculate each sub-domain score, the coefficients of each item on the basis of sub-domains are first determined by the researchers, then the coefficients are revised by taking the opinions of five experts online with the Delphi method about the appropriacy of the item coefficients. The revised coefficients are again reviewed individually by the experts; for the coefficients that cannot be reached on a consensus, the arguments of the experts who have different opinions are sent to the other experts. Finally, the interviews among the experts go on until a consensus on the coefficients is provided and the coefficients are thus finalized. The same procedure is followed in order to calculate the score for the parietal, frontal, temporal and occipital lobes and the IQ score. At the end of this process, according to the expert opinion of the test to be measured, content validity is provided, and evidence is collected for the construct validity.

The practitioners use observation checklists to understand the appropriacy level of the participants before starting the process. The items in the checklist include body language of the participant (for exp; eye movements, shaking hands, etc.). It also involves some specific questions for the parents about their children (for exp: the level of hunger, tiredness, etc.). Finally, the checklist involves items related to the physical environment the test is conducted in (for exp: the lightening, the background noises, etc.). Research suggests that all these factors may have an effect on the performance of the participants (Afridi et al., 2019; Kraft et al., 2016).

The criterion validity studies of ERAL-NIT are ensured via CAS regarding attention and hospital diagnoses on SLD, PDD, and LUD. Apart from this, eight school counselors apply the two tests: ERAL-NIT and CAS to 642 participants. They use a detailed application manual prepared by the researchers and send the data to the researchers by mail. Criterion validity of ERAL-NIT is ensured by comparing the general IQ scores obtained from ERAL-NIT and CAS in this way.

Parents of 574 of these participants are asked to state their evaluations about general and more specific performance-based characteristics of their children. To this end, a five-point Likert type questionnaire is used. The items in the questionnaire are determined by asking the opinions of five experts. The researchers use the same Delphi method to score the items in the questionnaire. The data gathered is used to see whether parental evaluations are consistent within themselves. Besides, the researchers compare parents' evaluations with the actual performances of their children and investigate any possible consistency and/or discrepancy. The benefit of using such a method is to facilitate the accuracy of the feedback given to parents about their children.

Apart from those 642 participants, 243 participants diagnosed with SLD, PDD, LUD, and ADD are also enrolled into the study. The aim is to increase criterion validity of the test by comparing the hospital diagnoses and the findings gathered from ERAL-NIT and to reveal to what extent the findings of the test support the hospital diagnoses. The findings of ERAL-NIT and hospital diagnoses are compared to increase the validity of the test. However, since the number of the individuals who have diagnosed for ADD in hospitals is just four, not enough for valid statistical processing, ERAL-NIT's findings for attention are compared with CAS' findings.

3. Findings

Below is given the descriptive statistics results of the development of the ERAL-NIT.

Table 3. Descriptive statistics of ERAL-NIT scores (with and without observations) and CAS scores

	IQ_I*	IQ**	TP (CAS)***
Mean	101.739	103.898	104.095
Std. error mean	0.513	0.517	0.525
Median	101.621	104.392	105.000
Standard deviation	13.003	13.106	13.295
Minimum	48.533	48.533	46.000
Maximum	141.231	141.231	139.000
Skewness	-0.119	-0.189	-0.330
Kurtosis	0.520	0.528	0.459

* Initial IQ_I (without considering the psychologist's observations)
** IQ scores after taking the effect of observations
*** CAS scores

Figure 1. Distributions of ERAL-NIT scores (with and without observations) and CAS IQ (TP) scores

	Group	N	Mean	Median	SD	SE
ERAL-NIT IQ	M	324	102.67	103.15	13.59	0.76
	F	318	104.98	105.21	12.47	0.70
CAS IQ	M	324	103.03	103.00	3.96	0.78
	F	318	105.07	106.00	12.50	0.70

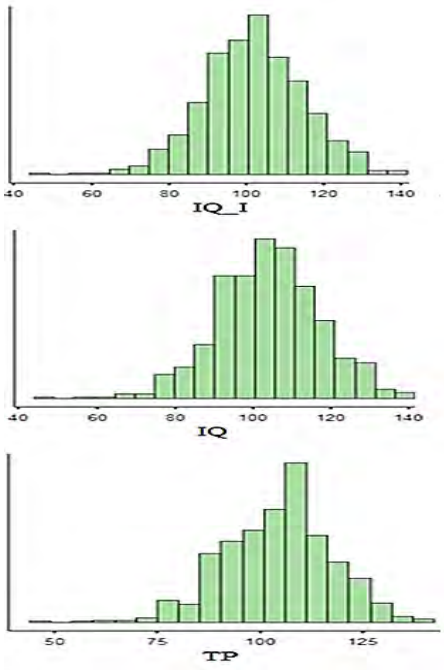

Table 3 and Figure 1 show that IQ scores without considering the practitioner's observations, IQ scores after taking the effect of observations and CAS scores have very similar descriptive statistics in terms of their means, standard deviations, skewness, kurtosis values; however, IQ scores' statistics after taking the effect of observations are closer to the CAS scores' statistics than that of IQ scores without considering the practitioner's observations. Pearson correlations among the three measurements indicate that although initial IQ scores without considering the practitioner's observations have a strong positive correlation with CAS scores, the relation between IQ scores after taking the effect of observations and CAS is stronger.

Table 4. Correlation matrix depicting the comparison of ERAL-NIT scores (with and without observations) and CAS scores

	IQ_I	IQ	TP (CAS IQ)
IQ_I	—		
IQ	0.97*	—	
TP (CAS IQ)	0.87*	0.91*	—

* $p < .001$

Based on the descriptive statistics and the correlation matrix given in Table 4, the researchers decide to utilize IQ scores after taking the effect of observations and corresponding sub-scores for the rest of the study.

Before doing ERAL-NIT and CAS comparison, a test of normality is conducted on ERAL-NIT and CAS scores. Shapiro Wilk statistics ($p<.05$ for ERAL-NIT and $p<.01$ for CAS) indicate that both variables have non-normal distributions.

Table 5. Descriptive statistics of ERAL-NIT and CAS by gender

	Group	N	Mean	Median	SD	SE
ERAL-NIT IQ	M	324	102.67	103.15	13.59	0.76
	F	318	104.98	105.21	12.47	0.70
CAS IQ	M	324	103.03	103.00	13.96	0.78
	F	318	105.07	106.00	12.50	0.70

Table 5 suggests that ERAL-NIT and CAS scores have similar descriptive statistics for both genders. The two scores for males have slightly lower mean and median and higher standard deviations.

Figure 2. Graphical representation of ERAL-NIT and CAS(TP) scores by gender

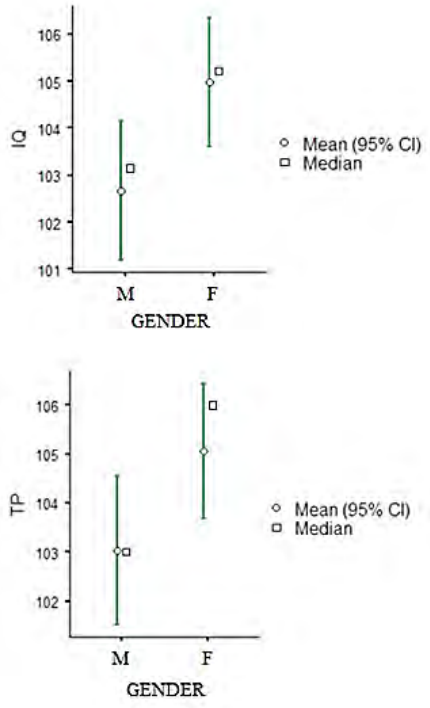

Figure 2 represents the differences on both test scores by gender. Based on the Mann-Whitney test results, it is concluded that the differences in IQ ($U=45874$; $p<.05$) and CAS ($U=46601$; $p<.05$) scores are significant in favor of girls. This finding is in line with some of the

studies in the literature. Palejwala et al. (2015) in their study related to gender differences and cognitive abilities find a female advantage for children aged 2-7. In a similar vein, Keith et al. (2008) in their study related to sex differences and latent cognitive abilities suggest that females show a consistent advantage on latent processing speed.

Table 6. ERAL-NIT positive and negative findings among those diagnosed as SLD positive by the hospital

Level	Count	Total	Proportion	P
POSITIVE	80	81	0.98	< .001
NEGATIVE	1	81	0.02	< .001

Table 6. shows the binomial test results on the SLD positive participants based on the tests conducted by the hospitals. Results reveal that the proportion of participants identified SLD positive by the ERAL-NIT is significantly higher than the proportion of the participants identified SLD negative. 83% of the SLD positive participants based on the hospitals' diagnostics are identified as SLD positive by ERAL-NIT. In addition, as seen in the contingency table below, an analysis of the frequencies regarding the participants with special needs per hospital records are generally matching with the diagnosis by ERAL-NIT. An attention drawing finding is that the 98% of SLD positives in terms of ERAL-NIT scores are identified as having other diagnosis such as language use disorder.

Table 7. Frequencies of ERAL-NIT and CAS in terms of ADD

ERAL-NIT ADD		NEGATIVE	POSITIVE	CAS ADD
NEGATIVE	Observed	527	28	555
	% within column	92 %	39%	86 %
POSITIVE	Observed	43	44	87
	% within column	8 %	61 %	14 %
Total	Observed	570	72	642
	% within column	100 %	100 %	100 %

Table 7. suggests that, 92% of the 570 participants who are ADD negative according to CAS are also ADD negative according to ERAL-NIT. 44 (61%) of the 72 participants who are ADD positive for CAS are also positive for ERAL-NIT. On the other hand, 28 participants who are ADD positive for CAS are negative for ERAL-NIT and 43 participants who are ADD negative for CAS are positive for ERAL-NIT. In short, CAS and ERAL-NIT results are very similar in detecting negatives, but the rate of obtaining similar results in determining positives decreases. A Chi-Square test revealed that there is a statistically significant relationship between the findings of ERAL-NIT and CAS in terms of ADD (χ^2 =156.58 (1); p< .001). These findings related to ADD also support the criterion validity of ERAL-NIT.

Table 8. ERAL-NIT positive and negative findings among those diagnosed as PDD positive by the hospital

Level	Count	Total	Proportion	P
POSITIVE	28	31	0.903	< .001
NEGATIVE	3	31	0.097	< .001

As Table 8 suggests, out of the 31 participants who are diagnosed as having PDD in the hospitals, 28 (90.3%) are identified as PDD positive by the ERAL-NIT. The correlation between them is statistically significant (P<.01).

Table 9. ERAL-NIT positive and negative findings among those diagnosed as LUD positive by the hospital

Level	Count	Total	Proportion	P
POSITIVE	82	89	0.921	< .001
NEGATIVE	7	89	0.079	< .001

As Table 9 depicts, out of the 89 participants who are diagnosed as having LUD in the hospitals, 82 (92.1%) are identified as LUD positive by the ERAL-NIT. The correlation between them is statistically significant.

Finally, as expressed before, parents are asked to assert their evaluations about what their children can perform in general and specific areas. A five-point Likert type questionnaire is used for this. Appendix A reveals the reliability of the questionnaire.

Table 10. Descriptive statistics of the IQ scores received from CAS, ERAL-NIT, and parental evaluations for children's general and specific performances

	CAS IQ	ERAL-NIT IQ	**PARENTS' G***	**PARENTS' S****
Mean	103.82	102.63	111.61	117.20
Median	105.00	102.58	113.47	119.79
Standard deviation	13.39	13.03	22.15	21.31
Minimum	46.00	48.36	31.67	38.18
Maximum	134.00	137.86	206.31	169.75
Shapiro-Wilk W	0.99	0.99	0.99	0.97
Shapiro-Wilk p	< .001	0.043	< .001	< .001

* IQ score for parental evaluations for children's performances in general
**IQ score for parental evaluations for children's performances in specific areas

Table 10 reveals descriptive statistical data gathered from 399 participants. It can be clearly seen that, means of parental evaluations for children's performances in general and specific areas and standard deviations are significantly higher than those of ERAL-NIT and CAS. This finding is in parallel to the literature. The literature on parents' beliefs about the intelligence of their children mainly suggest that parents believe their children are significantly brighter than they are, this means both parents have a tendency to overestimate their children's IQ (Chamorro-Premuzic et al., 2009; Furnham et al., 2002). Parental evaluations are considerably important since they affect how children are treated. Knowing what their children can do or cannot do plays a crucial role in having a healthy relationship among the members of the family.

Besides, according to the Shapiro-Wilk normality test, four variables do not have normal distribution. Thus, the Spearman Brown correlation coefficient is calculated. Table 11 shows the results.

Table 11. The Correlations between IQ scores of CAS, ERAL-NIT, and Parental Evaluations for Children's General and Specific Performance

	CAS IQ	ERAL-NIT IQ	PARENTS' G	PARENTS' S
CAS IQ	—			
ERAL-NIT IQ	0.80***	—		
PARENTS' G	0.28***	0.27***	—	
PARENTS' S	0.23***	0.18**	0.64***	—

*** p < .001

The above Table 11 reveals that data received from both ERAL-NIT and CAS has statistically significant positive but weak correlations with the parental evaluations for children's performances in general and specific areas.

4. Discussion

This study aims to develop and validate ERAL Nonverbal Intelligence Test. It is developed within the Turkish context. The content validity of the test items is ensured by asking the opinions of 18 experts from the fields of psychology, psychiatry, test development, child development, psychological counseling and neurology. The experts suggest what individuals can do between the ages 5 to 17 under the light of related literature. The coefficients are identified via the Delphi Method. The multiplication of these coefficients with the raw scores gives the IQ score, the scores for frontal, temporal, occipital and parietal lobe areas and the scores for the 16 subareas of brain's mental functioning.

The other equally important part of test development is ensuring the criterion validity of the test. To this end, the researchers compare the findings of ERAL-NIT with another well-known IQ test; CAS in a number of ways. Firstly, the IQ results gathered from ERAL-NIT are compared with those of CAS. The results suggest a statistically significant correlation between the findings of these two tests. Besides, it should be noted that when the observations of the practitioners are included in the evaluation, there appears to be an increase in the correlation between the ERAL-NIT's results and CAS' results. Secondly, the two tests are compared in terms of gender. Both of the findings are found to be compatible and in favor of girls. Finally, ERAL-NIT's findings for ADD are compared with CAS' findings. The results again suggest a positive correlation between these two tests in terms of ADD.

The other evidence for the validity of ERAL-NIT emerges from the comparisons of the findings of ERAL-NIT with the hospital diagnoses for SLD, PDD, and LUD. The findings reveal statistically significant positive correlations between these diagnoses and the findings of ERAL-NIT.

The fact that parental evaluations for children suggest a statistically significant relationship with the IQ scores of ERAL-NIT and CAS indicates that ERAL-NIT is a valid measurement tool.

Based on the above given data analyses and the findings, it can be concluded that ERAL-NIT is a valid and comprehensive cognitive IQ test which has its roots in brain-based science. It has many implications for educators, psychologists, psychiatrists, psychological counsellors, and child developmental specialists:

It is a nonverbal IQ test developed for 5 to 17-year-olds;

It reveals participants' cognitive functioning in a comprehensive way, and gives complementary suggestions for the areas that need to be developed;

The software system enables practitioners give feedback and report immediately after the application;

The software system decreases the practitioners' possible mistakes during the application;

It takes approximately 25 minutes, which means it is quite practical to use;

It involves parental evaluations for their children's general and more specific performances;

It helps parents treat their children more accurately, and keep more healthy relationships with their children;

It guides practitioners identify participants with Specific Learning Disability, Pervasive Developmental Disorder, Lack of Attention and Language Use Disorder.

In conclusion, ERAL-NIT suggests a great potential and may serve as an effective alternative to other IQ tests. Yet, it should be noted that further studies need be conducted. In this vein, it is recommended to apply ERAL-NIT on larger groups of people and evaluate the results accordingly. It is also suggested that similar to parental evaluations it can also be used to understand the accuracy of teachers' evaluations with regard to their students. This may help them have a better idea about their students and their capabilities.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public commercial, or not-for-profit sectors.

The authors declare no competing interests.

References

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227-257.

Anastasi, A. (1992). What counselors should know about the use and interpretation of psychological tests. *Journal of Counseling & Development*, *70*(5), 610-615.

Anderson, M. (2006). Intelligence. In *MS Encarta online encyclopedia*.

Ball, P. T. I. t. N. S.-A.-. (2018). The IQ trap. *New Statesman*, *13*(19), 31-35.

Beller, S., & Bender, A. (2010). *Allgemeine Psychologie–Denken und Sprache*. Hogrefe Verlag GmbH & Company KG.

Bender, A., & Beller, S. (2013). Cognition is... fundamentally cultural. *Behavioral Sciences*, *3*(1), 42-54.

Benson, N. F., Beaujean, A. A., McGill, R. J., & Dombrowski, S. C. (2018). Revisiting Carroll's survey of factor-analytic studies: Implications for the clinical assessment of intelligence. *Psychological Assessment*, *30*(8), 1028.

Bingham, W. V. (1937). *Aptitudes and aptitude testing*. Harpers.

Boroditsky, L., & Gaby, A. (2010). Remembrances of times East: Absolute spatial representations of time in an Australian aboriginal community. *Psychological Science*, *21*(11), 1635-1639.

Campbell, W. W., & DeJong, R. N. (2005). *DeJong's the neurologic examination*. Lippincott Williams & Wilkins.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Das, J. (1992). Beyond a unidimensional scale of merit. *Intelligence*, *16*(2), 137-149.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Allyn & Bacon.

Gansler, D. A., Varvaris, M., & Schretlen, D. J. (2017). The use of neuropsychological tests to assess intelligence. *The Clinical Neuropsychologist*, *31*(6-7), 1073-1086.

Gardner, H. (1993). Frames of Mind–the theory of multiple intelligences Fontana. *Gardner, W. and Abu Libde, A.(1995). Professional development schools: How well do they travel*, 303-315.

Haun, D. B., Rapold, C. J., Call, J., Janzen, G., & Levinson, S. C. (2006). Cognitive cladistics and cultural override in Hominid spatial cognition. *Proceedings of the National Academy of Sciences*, *103*(46), 17568-17573.

Henmon, V. A. C. (1921). *The measurement of intelligence. School and Society*, *13*, 151-158.

Herrnstein, R. J., & Murray, C. (2010). *The bell curve: Intelligence and class structure in American life*. Simon and Schuster.

Hooper, V. S., & Bell, S. M. (2006). Concurrent validity of the universal nonverbal intelligence test and the Leiter international performance scale–revised. *Psychology in the Schools*, *43*(2), 143-148.

An Interview with Dr. Simonton. (2003). http://www.indiana.edu/intell.

Kranzler, J. H., & Keith, T. Z. (1999). Independent confirmatory factor analysis of the Cognitive Assessment System (CAS): What does the CAS measure? *School Psychology Review*, *28*(1), 117-144.

Majid, A., Bowerman, M., Kita, S., Haun, D. B., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in cognitive sciences*, *8*(3), 108-114.

Masuda, T., & Nisbett, R. E. (2006). Culture and change blindness. *Cognitive science*, *30*(2), 381-399.

Medin, D. L., Ross, B. H., & Markman, A. B. (2004). *Cognitive Psychology*. John Wiley.

Naglieri, J., & Prewett, P. (1990). Nonverbal intelligence measures: A selected review of instruments and their use. *Handbook of psychological and educational assessment of children: Intelligence and achievement*, *1*, 348-368.

Naglieri, J. A. (1999). *Essentials of CAS Assessment*. Wiley.

Naglieri, J. A., & Das, J. (1990). Planning, attention, simultaneous, and successive (PASS) cognitive processes as a model for intelligence. *Journal of Psychoeducational Assessment*, *8*(3), 303-337.

Norenzayan, A., & Nisbett, R. E. (2000). Culture and causal cognition. *Current directions in psychological science*, *9*(4), 132-135.

Pintner, R. (1921). Intelligence and its measurement: A symposium. *Journal of Educational Psychology*, *12*, 195-216.

Rijumol, K., Thangarajathi, S., & Ananthasayanam, R. (2010). The PASS Theory of Cognitive Processing. *Journal on Educational Psychology*, *3*(4), 51-55.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications*. Jerome M Sattler Publisher.

Spearman, C. (1927). The abilities of man, New York: MacMillan.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. CUP Archive.

Stroke and cerebrovascular diseases (2019). (t. ed, Ed.). McGraw-Hill.

Thurstone, L. L. (1943). *Primary mental abilities*. Chicago: University of Chicago Press.