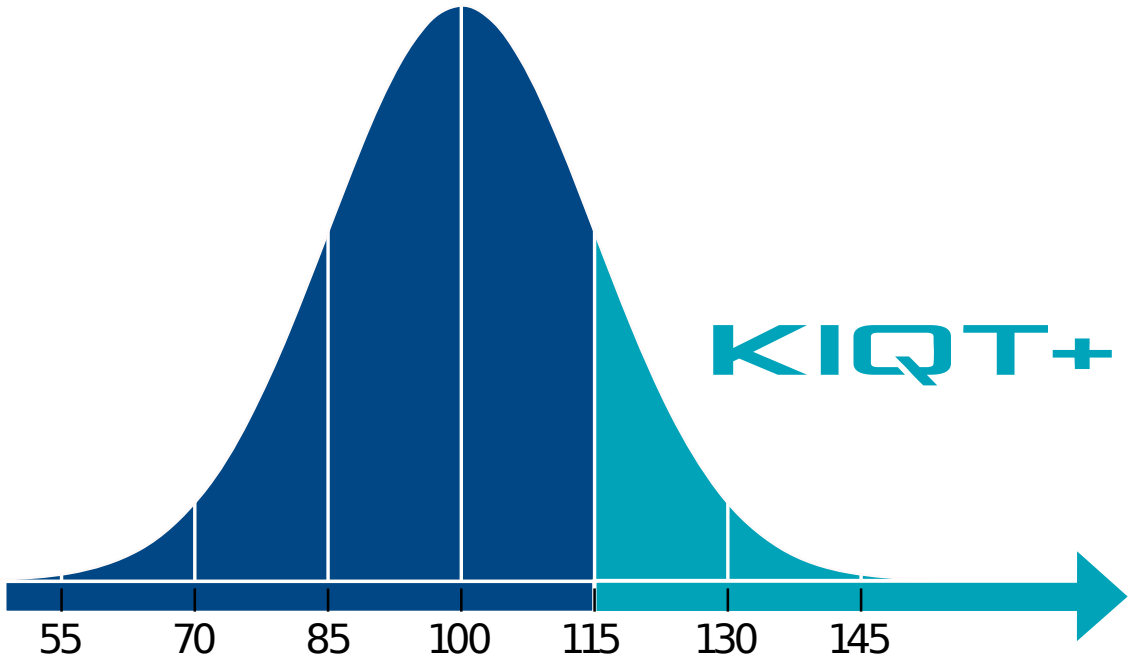

KIQT+

What is the KIQT+

The KIQT+ (Childrens IQ Test Plus, originally developed in the Netherlands as the Kinder IQ Test Plus) is an IQ test developed especially for the assessment of gifted and talented children. The test can be initiated by schools or parents when they suspect giftedness. The KIQT+ is most accurate within the IQ range of 115 (approximately 159 in 1000, or 1 in 6 children), to 170 (approximately 1 in 650.000 children).



The KIQT+ is designed by and for gifted and talented people. The motivation of the creators to design the KIQT+ is because there are so few IQ tests available for (highly) gifted children, there might be an inability for them to discover their true potential. With the KIQT+ , this is no longer a problem! Gifted and twice exceptional children are more than welcome to take part in this IQ test.

Standpoints of the KIQT+

The KIQT+ has four important standpoints: reliable, gifted proof, fair & efficient. Below we discuss each point and how each can be seen in the test.

Reliable

- The KIQT+ is intended for children ages five to ten with a (presumptive) IQ between 115 and 170. The KIQT+ is therefore designed as a more accurate measure than other IQ tests, without having to deal with the ceiling effects. See the study mentioned in reference [17] as an example.
- The KIQT+ uses a continuous age-norm that is calculated using the birth- and test date. SCALIQ believes that children continuously develop, rather than in three-to-four month windows. This approach prevents a child from being one day too old for the specific window, which could yield an IQ a few points lower than before. The KIQT+ therefore calculates the age exactly to the day.
- The KIQT+ provides information about the validity of the test for each child; that is, how well the child fits into the presumptions underlying the KIQT+ . The SCALIQ team will give an indication when the child's answer pattern indicates that calculating an IQ score would not be reliable. In case of such an inconsistent answer pattern, IQ scores are not given, since these would not be reliable. Given the extensive research that forms the basis of the development of the KIQT+ , the developers expect the frequency of such an answer pattern will be rare.
- The KIQT+ is an objective test. This is due to the fact that answers do not have to be interpreted by the examiner in order to comply to any termination rules. This prevents bias based due to (conscious or unconscious) expectations from the examiner about the intelligence of the child.
- First test rounds (N=54) show a very high internal consistency, a prerequisite for valid measurement (Cronbach's α = 0.952, Guttman's λ-2 = 0.956 (source: SCALIQ internal research). Considering the limit of 0.9, which the COTAN (Commission Test Affairs Netherlands) uses to deem tests appropriate for important decisions at the individual level, this result is very promising. According to the EFPA (European Federation of Psychologists' Associations) guidelines, these reliability coefficients can be deemed 'excellent'.

Gifted proof

- The KIQT+ avoids ambiguity and does not include questions with multiple correct answers. The open questions in other intelligence tests can lead (highly) gifted children to overthink, or to take longer than necessary to respond because they might think: "The answer cannot be that easy."
- The KIQT+ is specifically designed for highly intelligent children. With the high ceiling of the KIQT+ , possible ceiling effects are avoided (only 1 in 650.000 children reach the maximum IQ score of 170). The KIQT+ was developed specifically for gifted and talented children. It is a valuable tool to investigate how a child may deal with challenges that

are at (or above) their level. This can provide insight into work attitude, motivation and possible fear of failure.

- The KIQT+ can deal with so-called ”carelessness errors” of gifted individuals. Gifted children have sufficient cognitive capacity to complete the easy tasks. It is expected that incorrect answers will only occur in the most difficult tasks. However, sometimes it may be observed that a fairly easy assignment is answered incorrectly, which may be due to carelessness or overthinking. If the IQ is calculated based on the total number of correct answers, an easy or difficult task will have a similar influence on the final score. Because the KIQT+ does not use a sum score model but rather uses an Item Response Theory model (for further explanation, see page 9) the test is much less sensitive to such inconsistencies in response patterns. As a result, the KIQT+ will award the child that made such errors at the start of the test with an (approximately) equal score as the child who successfully completed all the easy questions. This is because the answer pattern shows that the thinking level of the child is higher than the level of the mistake made.

Fair

- Time-pressure is avoided as much as possible when taking the KIQT+ , . Children who suffer from fear of failure, performance pressure, issues with focussing or simply do not perform as well under time pressure are therefore not unnecessarily disadvantaged by the KIQT+ . Children with learning disabilities also generally function worse on intelligence tests with time-pressure [5], so their intellect wil be better represented on the KIQT+ . Other countries also advise to decrease the time pressure when identifying gifted children [22] because this appears to be a score lowering obstacle [29].

- The KIQT+ is less biased against children with autism spectrum disorder, because verbal items are not being used [13]. Research shows that the performance of children without autism spectrum disorder is comparable for intelligence tests with and without a verbal component [6, 30], whereas, children with autism score lower on intelligence tests with a verbal component.

- The KIQT+ is less biased against children from lower socio-economic environments, children that have attended non-optimal educational settings, and children with a language delay or migration background. This is because the KIQT+ makes no assumptions about the childs present knowledge.

Efficient

- The (sub)scores achieved by a child on the KIQT+ are fully calculated by the SCALIQ algorithm. After both the raw data and the calculated scores have been checked by a SCALIQ psychometrician, the test user receives a digital report. This report contains all the information necessary for making a comprehensive intellectual report. The tester is saved a lot of valuable time.

- The test procedure of the KIQT+ is simple and is easy to master.

An intelligence test that really measures intelligence!

Each intelligence test uses one or more underlying theories of what intelligence is and how it can be measured. Today, the Cattell-Horn-Carroll (CHC) model of cognitive capacities is seen as the most comprehensive and empirically supported model for intelligence [19, 14]. The model is a fusion of the work of Raymond Cattell, John Horn and John Carroll [1, 16, 19, 28]. Due to the impressive amount of empirical evidence for this model in different research areas (developmental psychology, neuropsychology, etc.) the model is implemented in construing, interpreting, and categorizing intelligence tests. Most new and revised intelligence tests are based on the CHC model [12].

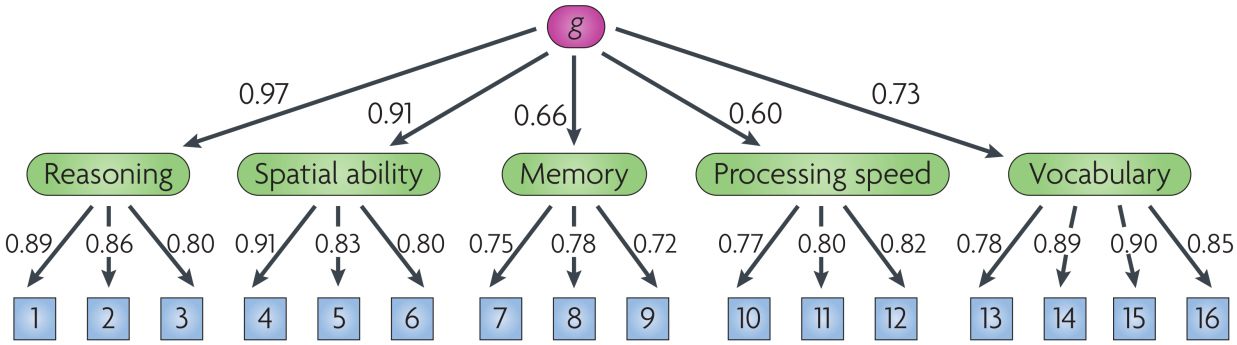


Figure 1: Factor analysis of 16 different cognitive tests conducted on nearly 7.000 people between 18 and 95 years old [8]. The 16 tests coincide in five "broad factors": Reasoning, Spatial abilities, Memory, Processing speed, and Vocabulary. All five broad factors have a positive correlation with the overlapping g-factor. Figure 1 based on original work by Salthouse [27].

Factor analysis of known intelligence tests that support the structure of the CHC model are regularly published (see, for example, [11]). Even if a multitude of tests are used, factor analysis continues to support the CHC model [8]. What is striking about the factor analysis of test data according to the CHC model is that the fluid reasoning factor always has the highest correlation with the underlying ("stratum 3") g-factor, followed by the visual / spatial factor, the verbal factor, working memory and finally processing speed. This finding is therefore consistent with the fact that visual reasoning tasks (such as the Raven Progressive Matrices) are often seen as intelligence tests with a high g-loading [27, 14].

SCALIQ believes that a good intelligence test should above all be good at measuring actual intelligence. This means that we do not test previously acquired knowledge (such as vocabulary) with the KIQT+ . We made this choice because the level of knowledge, and knowledge attainment opportunities varies per child. Although the acquisition of knowledge is correlated with intelligence, it also depends on the environment and learning opportunities. Despite the fact that most children go to school, there are major differences in the degree, type and amount of knowledge and information that is offered in childhood. For children with a migration background, children growing up in poverty, and children in other adverse situations, we can also not assume that they had the same opportunities for information as

children without such issues.

Previously acquired knowledge only represents something about the past, namely how well a child has absorbed the knowledge offered to him/her. This is also a result of the educational opportunities that he/she has enjoyed. Educational opportunities are not always appropriate, particularly for (highly) gifted children. A measurement of general intelligence (g) without relying on previously acquired knowledge therefore has a greater predictive value for this group of children for acquiring knowledge and skills in the future. When educational decisions (for example, placement in a gifted program, at a school for gifted children) are taken on the basis of an IQ score, it is important that the intellectual potential of the child is properly documented - independent of previously acquired knowledge and earlier education received.

In addition to the fact that previously acquired knowledge cannot be assumed to be the same for all children, the use of knowledge is not necessary for a reliable IQ measurement. As can be seen in the hierarchical model of intelligence, general intelligence (g) is the goal of intelligence measurements. For these reasons, SCALIQ chooses to only offer subtests that have a high g-loading and are not dependent on previously acquired knowledge.

Designing an intelligence test without verbal items

Although verbal items were still the main part of intelligence testing in the last century, advancing insight from research has led to many intelligence tests containing fewer language items. At SCALIQ we explicitly opt for an intelligence test with a reduced emphasis on language, whereby a verbal answer is not necessary. We do this for the following reasons:

- Verbal items by definition use linguistic knowledge acquired by the child. The acquisition of language - just like the acquisition of other knowledge and skills are influenced by both the aptitude and the environment of a child.
- The KIQT+ must be inclusive for the entire group of gifted and talented people. This means that it also should be applicable for children with a migration background, low socio-economic status and learning and developmental disorders such as dyslexia and autism spectrum disorder.

  For these groups, testing intelligence through linguistic / verbal items often means a great underestimation of their developmental potential. Research into children with autism shows that intelligence tests with a verbal component can significantly underestimate the intellect of these children [21, 30, 20, 6]: up to 14 IQ points for younger children [13]. It is also known that gifted children from the aforementioned groups are identified less quickly (or not at all).
- Language can be used to describe abstract concepts, with logical reasoning, classification or deduction on the abstract concept leading to the correct answer. Examples of this are verbally offering similarities (what is the similarity between ... and ...?), and verbal analogies (arm is related to hand and leg is related to ...?). The measurement presumption for these kinds of tasks concerns reasoning, classification and deduction, but whether the assignment is made successfully also depends on word knowledge or language level.

From the previous it follows that verbal test items are not necessary to measure general intelligence (g). Just as with other IQ tests, the reasoning with, sorting, quantifying an qualifying of abstract concepts is central in the KIQT+ . SCALIQ chooses to present (parts of) these

abstract concepts to the child without using language skills or word knowledge

This does not mean that the KIQT+ is a test for non-verbal intelligence. Neither can the KIQT+ be seen as a fully nonverbal intelligence test, since parts of understanding the instruction still need a certain language level. The KIQT+ can therefore best be seen as a general intelligence test with a reduced emphasis on language and prior knowledge.

This choice aligns the KIQT+ into the wider movement of intelligence tests that choose to reduce the emphasis on verbal content and previously acquired knowledge, like the NNAT (Naglieri Nonverbal Ability Test), de (C)TONI ((Comprehensive) Test of Nonverbal Intelligence) and the Raven CPM/SPM/APM (Coloured, Standard en Advanced Progressive Matrices). In the identification process of gifted and talented children it is also frequently advised to look at those subtests and/or factor scores of intelligence tests that represent abstract reasoning, while the usefulness of short term memory and processing speed is de-emphasized [22].

Opting for an intelligence test without ambiguity

Whish animal does not belong in the following list:
Cow Rooster Pig Sheep

Is it Pig? Because it is not kosher?
Is it Cow? Because this is the only one with a 'w' in the word?
Is it Rooster? Because this is not a mammal?
Is it Cow? Because people can be allergic to cow's milk?
Or is it Pig anyway? Because the skin is completely visible in this animal?
Or is it Cow anyway? Because this is the only animal I havent seen at my local petting zoo?
Or is it Rooster? Because it only has two legs?
Or is it Cow anyway? Because it has four stomachs?
Or is it Sheep after all? Because it does not belong in any of the previous ways?

SCALIQ believes that ambiguity does not belong in an intelligence test. Although many gifted and talented children probably have an idea of what the correct answer might be, many also find it hard to take a guess [29]. In the development of the KIQT+ , SCALIQ uses mathematical models to provide insight into possible relationships between the stem (the assignment/question), the correct answer and the incorrect answers (distractors) and where necessary to adjust or remove items. This is to ensure that ambiguity, uncertainty and confusion occurs as little as possible.

Design

Specific attention was paid to design while developing the KIQT+ . The structure and design of the test was developed with the idea that scoring and carelessness errors can be avoided as much as possible. In addition, extensive attention has been paid to the prevention of visual overstimulation or lack of clarity, by avoiding unnecessary shape and / or color variations.

Example item (see figure 2)

- By offering only 1 item at a time, the assumption of logical relationships across items is avoided. In this way, the presence of other items in sight cannot form a distraction.
- To prevent visual overstimulation and distraction, each item contains as little unnecessary information as possible.

1. A recognizable item number at a fixed position ensures clarity and reduces the number of scoring errors.
2. The KIQT+ uses a fixed, simple color scheme that is suitable for color blind people. In this way, parts of the item that differ only on the color aspect are clearly differentiable. New and unnecessary colors are avoided.
3. The black frame makes it immediately clear which parts belong to the stem of the item.
4. Distractions are positioned at a large distance from the stem, so visual interference is minimized. Concentration on the stem of the item is promoted and response elimination (the systematic elimination of answers) is discouraged. Research shows that response elimination not only lowers the g-loading of an intelligence test, but can also jeopardize construct validity [2, 4].
5. Answer options are positioned separately from each other, so no series formation is implied.
6. A balanced number of distractors ensures a correct ratio between reducing the chance of guessing and the risk of providing information about the correct answer.
7. The use of multiple high-quality distractors formed according to established rules promotes the use of constructive matching, which increases the g-loading of items [2].
8. The response box and the answer options are designed asymmetrically by means of a visual anchor on the bottom left. In this way the suggestion that answer options could be rotated can be prevented. Almost a century ago it was mentioned in the literature that this should be considered [23].
9. The numbering of answer options is integrated at the visual level. This prevents errors compared to placing the answer numbering above, below or sideways of the answer options.

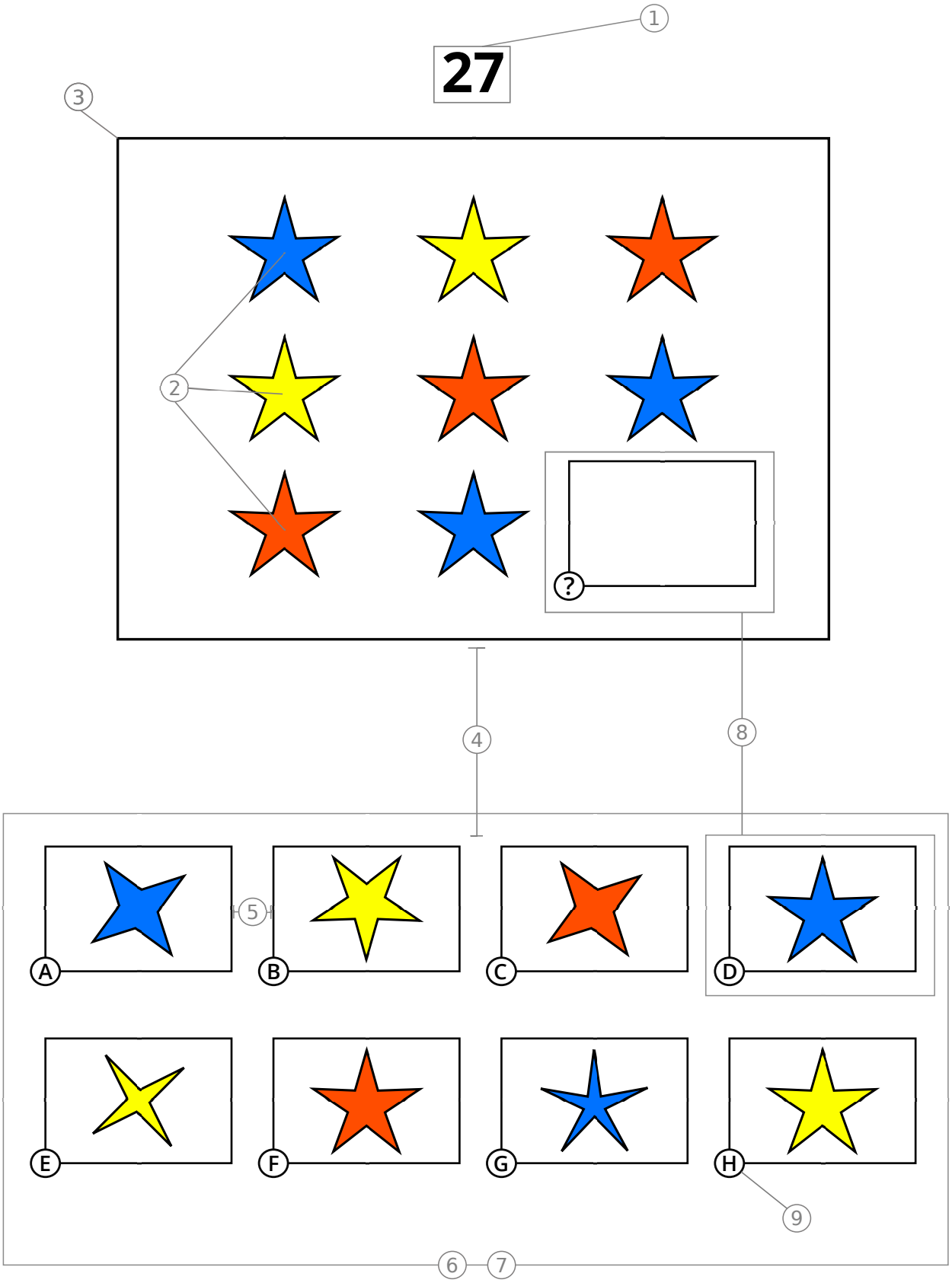


Figure 2: KIQT+ sample item.

Item Respons Theory (IRT)

Every IQ test has an underlying theoretical model to generate a measurement (substantiated estimate) of intelligence. The expectation in classical test theory is that more intelligent children will answer more questions correctly. In traditional test theory, sum scores are used for the test score. The number of correctly-made assignments is added, this number is then the child's raw score. Item Response Theory (IRT) works in a fundamentally different way.

Example

The easiest way to explain the difference between classical test theory and Item Response Theory is based on an example. In the table below we see a test of ten multiple choice questions with increasing difficulty. The test was scored according to both Classic Test Theory (CTT) and Item Response Theory (IRT).

	← Easy questions								Difficult questions →		CTT	IRT	IRT
	1	2	3	4	5	6	7	8	9	10	score	score	reliability
Anne	✓	✓	✓	✓	✓	✓	✓	*x*	*x*	*x*	7	7	excellent
Jason	✓	*x*	✓	✓	✓	✓	✓	*x*	*x*	*x*	6	6.9	Good
Emily	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10	10+	Excellent
Mike	-	-	-	✓	✓	✓	✓	✓	✓	✓	6	10+	Good
Linda	✓	✓	*x*	✓	✓	*x*	*x*	✓	*x*	*x*	5	5	Average
Fred	-	*x*	✓	*x*	*x*	✓	*x*	✓	✓	✓	5	-	Unreliable

Anne answered 7 out of 10 questions correctly and therefore gets a score of 7 within the classical test theory framework. If we look at which questions she has answered correctly, we see that she has answered the 7 easiest questions correctly and none of the 3 most difficult questions. The (increasing) difficulty of the questions seems to match Anne's intellectual skills, in short it seems that Anne has been measured correctly by this test. The IRT score in this case is the same as the CTT score. Furthermore, the IRT reliability column indicates that Anne's response pattern fits perfectly with the expectation of the IRT model.

Jason has answered 6 out of 10 questions correctly and therefore receives a CTT score of 6. It is striking that Jason answered one of the easiest questions, question 2, incorrect. Has he been sloppy? It seems unrealistic that Jason is less smart than Anne. After all, on the more difficult questions, questions 3 to 10, he had the same answers as Anne! IRT considers the pattern of the answers in combination with the difficulty of the questions and gives Jason a score of 6.9. The reliability of the IRT score is still regarded as good, because 8 out of 10 answers from Jason meet the expected response pattern of a child with a real score of 6.9.

Emily answered all the questions correctly, so her CTT score is 10. If there had been even more difficult questions, could she have answered them correctly? We do not know that, because the test is clearly too easy for Emily. The IRT model indicates in this case that the actual score is probably above 10. In the "IRT reliability" column, the IRT model also indicates that it is fairly certain, because Emily answered all the questions correctly.

Mike did not answer the first three questions. His CTT score is therefore 7. The IRT model does not; however, directly regard unfilled questions as errors. Because the completed

questions were all correctly answered by Mike - and these were also the most difficult questions - the IRT model gives Mike a score of 10+. The IRT model is somewhat less certain about this, because it is not clear why the first 3 questions were not answered.

Linda's response pattern is somewhat inconsistent. Although she correctly answered question 8, this can of course be a coincidence in a multiple choice test. The IRT model indicates that the inconsistent response pattern gives reason to doubt Linda's score.

Fred correctly answered the three most difficult questions. The chance that he has correctly guessed all of this is very small. Besides that, Fred did not answer question 1, and incorrectly answered questions 2, 4, 5 and 7. The chance that all of these are careless mistakes is very small. The IRT model here indicates that Fred's response pattern cannot be interpreted well enough to give a reliable score.

Norming

Traditionally, intelligence tests are normed by using a large randomized reference group that is stratified over variables such as gender, age, level of education, degree of urbanization and geographical location. It is important to realize that this procedure did not come into existence just like that, and that every aspect of it must be carefully monitored in order to eventually obtain a representative sample from which a careful norm can be constructed. With the norming procedure for the KIQT+ , careful consideration has been given to how correct norms can be obtained. Because the KIQT+ is not a test for average children, but is explicitly intended for children from 1 standard deviation above average, strictly following a traditional norming procedure would cause some insurmountable problems.

When testing the KIQT+ within a random sample of children, approximately 16% of the children will fall within the target group for the KIQT+ . This also means that for 84% of the children the KIQT+ is (far) too difficult, and therefore their test data cannot be used in the norm group. A possible solution for this problem would be to use a selection method, for example an earlier IQ test. The disadvantage of using a selection procedure; however, is that the norm group becomes directly dependent on the selection procedure used and can therefore no longer be called random. An additional disadvantage is that, assuming that about 1000 children are needed in the norm group, this means that more than 6000 children should be subjected to such a selection test!

In addition to problems in the random selection of the norm group, based on the normal distribution of IQ scores, it is to be expected that in a randomly selected norm group there will be many more children with an IQ around 120 than children with an IQ around 160. This would make the standardization of the test very unbalanced and less information available about how the more difficult items discriminate between children with a higher IQ. A solution to this problem would of course be to stratify IQ scores (as measured by the selection test) and to consciously recruit more children with a higher IQ score in order to guarantee a balanced standardization of the more difficult test items. Unfortunately, this would mean that, to recruit enough children with an IQ of more than 3 standard deviations above average, many more children than the aforementioned 6,000 would have to be subjected to a selection procedure. In addition, there are few or no tests available to date that can accurately determine children's intelligence up to about 5 standard deviations above average.

Partly for the foregoing reasons, it has been decided to norm the KIQT+ in a different and new way. The norming of the KIQT+ consists of an amalgamation of traditional standardization techniques, mathematical properties of the Item Response Theory and new research into the predictability of the score distribution of cognitive tasks over age. All this to, despite the obvious challenges in norming the KIQT+ , come to an accurate and reliable norm.

Invariance of item parameters in Item Response Theory

One of the benefits of Item Response Theory is that, when the chosen model fits the data, the item parameters are invariant with respect to population differences. This property of IRT is also called "parameter invariance", "population invariance", "group invariance" or "sample invariance" [18, 9, 7, 15, 3]. In addition to the fact that this invariance property is a consequence of the theoretical background of IRT, more recent research shows that this property

also holds in practice if some weak requirements are met [26, 25, 24]. The consequence of the invariance property of IRT for the norming of the KIQT+ is that only a few constants need to be empirically determined in order to arrive at the correct item parameters for the target population.

Anchor points

Determining the necessary constants (”anchor points”) is part of the KIQT+ norming process where traditional norming techniques demonstrate their reliability and value. These samples will of course be randomly selected and stratified over some variables such as gender, age, educational level (of parents), degree of urbanization and geographical location.

Predictability of score distribution for cognitive tasks over age

Previous research by SCALIQ (source: SCALIQ - internal research) shows that the score distribution for cognitive tasks over age behaves extremely predictably. In this study, norms from various intelligence tests, such as the Raven Matrices (CPM, SPM, SPM + and APM), some sub-tests from the WISC-V-NL and the RAKIT-2 and the NNAT, from different dates between 1982-2018 and from various countries including the Netherlands, the United States, Poland, China, the United Kingdom, Slovenia, Germany, Taiwan, France, Romania, Argentina, New Zealand, Australia, India, Switzerland and Qatar, were combined. In these norms we looked at so-called ”equivalence points”: imagine a child of 6.5 years from England gets a score of 25 on the CPM, and thus achieves the 90th percentile within the 1984 norm. However, within the same 1984 norm, the child would only achieve the 75th percentile if he or she had been 8 years old at the time. The equivalence point consists of 4 data points, two sets of an age and a z-score that ”belong together” in this way.

After constraining the regression model that if the age is the same, the corresponding z-score must also be the same and vice versa it appeared that with 3 of the 4 data points the variance in the 4th data point can be explained by 97.5% (correlation between prediction and actual data points 0.987) . Because norm tables often contain raw (sub) test scores and / or are subject to rounding errors, this research result can be called highly convincing for the predictability of the score distribution of (sub) tests for intelligence over age.
For the standardization and norming of the KIQT+ , we will further develop and expand these research results. For example, with the invariance property of the IRT and the anchor points obtained through traditional norming techniques, we guarantee accurate and reliable standardization and norming of the KIQT+ .

Contact

Do you have any questions about the KIQT+ ? Do you want to participate in the standardization or norming of the test (as a parent or professional) or do you want to start using it? Contact us via info@scaliq.com.

Read more at www.scaliq.com and subscribe to the newsletter. In this way, you can stay informed of further developments.

Bibliography

[1] V. C. Alfonso, D. P. Flanagan, and S. Radwan. The Impact of the Cattell-Horn-Carroll Theory on Test Development and Interpretation of Cognitive and Academic Abilities. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, pages 185–202. The Guilford Press, New York, NY, US, 2005.

[2] M. E. Arendasy and M. Sommer. Reducing response elimination strategies enhances the construct validity of figural matrices. *Intelligence*, 41(4):234–243, July 2013.

[3] F. B. Baker. *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation, College Park, Md., 2nd ed edition, 2001.

[4] N. Becker, F. Schmitz, A. Falk, J. Feldbrgge, D. Recktenwald, O. Wilhelm, F. Preckel, and F. Spinath. Preventing Response Elimination Strategies Improves the Convergent Validity of Figural Matrices. *Journal of Intelligence*, 4(1):2, Feb. 2016.

[5] C. Cornoldi, D. Giofr, A. Orsini, and L. Pezzuti. Differences in the intellectual profile of children with intellectual vs. learning disability. *Research in Developmental Disabilities*, 35(9):2224–2230, Sept. 2014.

[6] M. Dawson, I. Soulires, M. Ann Gernsbacher, and L. Mottron. The Level and Nature of Autistic Intelligence. *Psychological science*, 18(8):657–662, Aug. 2007.

[7] R. J. De Ayala. *The theory and practice of item response theory*. Methodology in the social sciences. Guilford Press, New York, 2009. OCLC: ocn270230503.

[8] I. J. Deary, L. Penke, and W. Johnson. The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, 11(3):201–211, Mar. 2010.

[9] C. DeMars. *Item response theory*. Series in understanding statistics. Measurement. Oxford University Press, Oxford ; New York, 2010. OCLC: ocn429474027.

[10] L. S. DeThorne and B. A. Schaefer. A Guide to Child Nonverbal IQ Measures. *American Journal of Speech-Language Pathology*, 13(4):275, Nov. 2004.

[11] S. C. Dombrowski, G. L. Canivez, and M. W. Watkins. Factor Structure of the 10 WISC-V Primary Subtests Across Four Standardization Age Groups. *Contemporary School Psychology*, 22(1):90–104, Mar. 2018.

[12] D. P. Flanagan and P. L. Harrison, editors. *Contemporary intellectual assessment: theories, tests, and issues*. Guilford Press, New York, 2nd ed edition, 2005.

[13] S. N. Grondhuis, L. Lecavalier, L. E. Arnold, B. L. Handen, L. Scahill, C. J. McDougle, and M. G. Aman. Differences in verbal and nonverbal IQ test scores in children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 49:47–55, May 2018.

[14] R. J. Haier. *The neuroscience of intelligence*. Cambridge fundamentals of neuroscience in psychology. Cambridge University Press, New York, NY, 2017.

[15] R. K. Hambleton, H. Swaminathan, and H. J. Rogers. *Fundamentals of item response theory*. Number 2 in Measurement methods for the social sciences series. Sage Publications, Newbury Park, Calif, 1991.

[16] J. L. Horn and N. Blankson. Foundations for Better Understanding of Cognitive Abilities. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, pages 41–68. The Guilford Press, New York, NY, US, 2005.

[17] C. Kaplan. Ceiling Effects in Assessing High-IQ Children With the WPPSI–R. *Journal of Clinical Child Psychology*, 21(4):403–406, Dec. 1992.

[18] W. J. v. d. Linden, editor. *Handbook of item response theory*. Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences Series. CRC Press, Taylor & Francis Group, Boca Raton, 2015.

[19] K. S. McGrew. The Cattell-Horn-Carroll (CHC) theory of cognitive abilities: Past, present and future. In *D. Flanagan, & Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues-Second Edition*, pages 136–202. Guilford Press, New York, 2005.

[20] S. D. Na and T. G. Burns. Wechsler Intelligence Scale for Children-V: Test Review. *Applied Neuropsychology: Child*, 5(2):156–160, Apr. 2016.

[21] A.-M. Nader, V. Courchesne, M. Dawson, and I. Soulires. Does WISC-IV Underestimate the Intelligence of Autistic Children? *Journal of Autism and Developmental Disorders*, 46(5):1582–1589, May 2016.

[22] NAGC. Use of the WISC-V for Gifted and Twice Exceptional Identification - Position Statement. Technical report, National Association for Gifted Children, Aug. 2018.

[23] L. S. Penrose and J. C. Raven. A NEW SERIES OF PERCEPTUAL TESTS: PRELIMINARY COMMUNICATION. *British Journal of Medical Psychology*, 16(2):97–104, Nov. 1936.

[24] A. A. Rupp. Quantifying Subpopulation Differences for a Lack of Invariance Using Complex Examinee Profiles: An Exploratory Multigroup Approach Using Functional Data Analysis. *Educational Research and Evaluation*, 11(1):71–97, Feb. 2005.

[25] A. A. Rupp and B. D. Zumbo. A Note on How to Quantify and Report Whether Irt Parameter Invariance Holds: When Pearson Correlations are Not Enough. *Educational and Psychological Measurement*, 64(4):588–599, Aug. 2004.

[26] A. A. Rupp and B. D. Zumbo. Understanding Parameter Invariance in Unidimensional IRT Models. *Educational and Psychological Measurement*, 66(1):63–84, Feb. 2006.

[27] T. Salthouse. Localizing age-related individual differences in a hierarchical structure. *Intelligence*, 32(6):541–561, Dec. 2004.

[28] W. Schneider and K. S. McGrew. The Cattell-Horn-Carroll Model of Intelligence. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, pages 41–68. The Guilford Press, New York, NY, US, 2005.

[29] L. K. Silverman. Assesment of giftedness. In *Handbook of giftedness in children: psychoeducational theory, research, and best practices*. Springer Science+Business Media, New York, NY, 2018.

[30] I. Soulires, M. Dawson, M. A. Gernsbacher, and L. Mottron. The level and nature of autistic intelligence II: what about Asperger syndrome? *PloS One*, 6(9):e25372, 2011.

KIQT+ Whitepaper
Version 9- 8- 2019