Running Head: Culture Free Intelligence Test

**Culture Free Intelligence Test with Refugees - The Role of Education in Intelligence
Testing**

Caterina Pinto* & Anja Kühnel*

Author Note

Caterina Pinto, MSB Medical School Berlin, Germany, <https://orcid.org/0000-0002-6243-8345>

Anja Kühnel, MSB Medical School Berlin, Germany, <https://orcid.org/0000-0002-7060-7399>

* Both authors contributed equally to this work

Correspondence concerning this article should be addressed to Caterina Pinto,

E-Mail: caterina.pinto@icloud.com

Abstract

Culture free intelligence tests developed in western societies claim to measure the universal general factor of intelligence and for this reason are used in cross cultural and comparative studies but have shown large and unaccountably significant difference in IQ scores between western subjects and other ethnic and cultural groups.

We collected the IQ scores from 115 subjects in a natural quasi-experiment with culture free IQ tests. The sample consists of refugees from Middle East and Africa, mostly from Iran (50%) and Syria (19%). We examined the effect of the educational level (schooling, university attendance and field of study/work) on the IQ score obtained through a culture free intelligence test: The results show a strong positive correlation between the years of schooling and IQ scores. The duration of university attendance can explain 7% of the variance in the IQ scores. The educational level rank of the country of origin (percentage of alphabetization and compulsory schooling in the population) and the field of study/work correlate significantly with the IQ score. There is also a significant difference in IQ scores between participants who attended university and those who did not. Our results show that years of schooling and the educational level of the country of origin influence intelligence testing also in the so-called culture free intelligence tests. Thus, it is not the general intelligence factor which is tested but the formal education and test situation familiarity of the participants. We propose two solutions. First, non-Western test persons should be trained in advance and be familiarised to test situations to at least provide a fairer test situation before administering culture free intelligence tests. Second, culture specific intelligence tests should be constructed and validated for each specific culture and particular kind of socialization.

Keywords: IQ, culture free intelligence test, refugees

Culture Free Intelligence Test with Refugees - The Role of Education in Intelligence

Testing

The testing of intelligence has a long and controversial history. Intelligence testing was first developed in western societies to assess the abilities of individuals or their development (Maltby, Day & Macaskill, 2013). Intelligence tests are considered a reliable measuring instrument, as test-takers cannot get better IQ scores as allowed by their actual intelligence (Schmitt and Gerstenberg, 2014; Schuler, 2014). IQ tests have a better predictive validity (for example for school achievement and job performance) than self and external evaluation as many studies showed (Maltby, 2011; Holling et al., 2004 as cited in Stemmler et al., 2016; Rost, 2013). Boring (1923) defined intelligence “as the tests test it”. Unfortunately, what is measured with an intelligence test depends on the definition of intelligence - and its constituting abilities - used in test construction and not on intelligence in general or in a broad sense of the word (Maltby, Day & Macaskill, 2013).

The theory of a *general intelligence factor* (g) has been explored by Spearman between 1904 and 1921 through a factor analysis of the relationship between a variety of intelligence tests (Maltby et al., 2013). This theory was further developed by Cattell (in 1987) who stated that general ability consists of the factors *fluid intelligence* (g-factor) and *crystallized intelligence* (verbal, numerical, reasoning; see Rost, 2013). General intelligence can be defined as an independent ability, which allows to build specific skills and to acquire knowledge (Helfrich, 2013). Helfrich (2013) assumed that specific cognitive abilities (e.g. visual-spatial abilities, memory capacity and mathematical understanding) may be operationalized as deductive thinking and established as a comparable cross-cultural concept.

Nevertheless, there are a number of studies, which show an influence of culture and nurture on the measured g-factor: For example, the educational level manipulates the Intelligence Quotient (Helfrich, 2013). Flynn and Williams (both in 1998) suggested that western modern schooling trains the abilities measured in achievement tests (as cited in

Maltby et al., 2013). The mean IQ in a country correlates positively with the prevalence of formal education in that very population (Helfrich, 2013; Fox & Mitchum, 2012).

Furthermore, the knowledge about and the operating experience with intelligence tests allowed subjects to score higher (Maltby, Day & Macaskill, 2013). Helfrich (2013) also indicates a cultural influence on the biological roots of thinking, especially through language and the interaction between language and memory (memory span in particular; see also Fasfous, Puente, Pérez-Marfil, Peralta-Ramirez & Pérez-García, 2013).

Culture free intelligence tests were developed as an answer to these problems and to test intelligence without the cultural bias inherent in the earlier tests (e.g. Stemmler et al., 2016).

Culture Free Intelligence Tests. In cross-cultural studies, only the fluid intelligence (*g*) is usually taken into account when measuring the participants IQ scores. To exclude the influence of culture and language, most of the so-called *culture free* or *culture fair* intelligence tests are based on the Raven's Progressive Matrices (as cited in Helfrich, 2013). The assumption is, that these items analyze the ability to form perceptual relations and to reason by analogy independent of language and are therefore culture fair (Maltby et al., 2013). Helfrich (2013), however, reasons that these tests only measure the ability to work with specific visual patterns and their underlying rules rather than general intelligence. Studies by Dicks and Flynn showed how the industrialization (and digitalization) of the modern western societies influences the performance in matrices tests (as cited in Helfrich, 2013). In the course of the last century, people in these societies gained a high amount of practice in working with abstract visual patterns through formal schooling of at least eight to ten years. Procedural knowledge about problem solving, the ability to solve these problems as well as abstract thinking are trained in formal school education (Fox and Mitchum, 2012; Maltby et al., 2013; Bergold, Wirthwein, Rost & Steinmayr, 2017). Poortiga (as cited in Trommsdorff, 2007) found higher IQ scores of subjects from nations with an education

system similar to western schooling as compared to subjects from countries with limited or nonexistent educational offers. Accordingly, Bergold et al. (2017) showed that each additional year of schooling has a significant positive impact on IQ scores by training abilities tested in intelligence tests independent of the contents learned in schools. For example, Brunner et al. (as cited in Rost, 2013) found that the *g* factor is a main predictor of solving mathematical tasks and that the training of three-dimensional thinking in mathematics influences arithmetic inference. Bergold et al. (2017) found the same results for students in the natural science or technology track, as Rost (2013). These findings indicate the positive role of scientific knowledge and training in gaining higher IQ scores by testing deductive thinking and reasoning. Similar effects can be found for reading. While gaining expertise in reading, students learn to categorize abstract shapes, which increased their IQ scores in matrices tests (Fox and Mitchum, 2012). Schooling furthermore leads to a high level of experience with test situations in paper-pencil as well as computer-based formats. A lack of this experience decreases test performance (according to Van de Vijver, as cited in Trommsdorff, 2007; Fafous et al. 2013). Members of western societies are thus more experienced with these test formats and score higher (Helfrich, 2012).

As Intelligence is seen as a universal ability, IQ scores are expected to be distributed normally worldwide. This is not the case as shown in cross-cultural studies (e.g. Rost, 2013; Maltby, Day and Macaskill, 2011; Funke & Vaterrodt-Plünnecke, 2004). Different societies tend to increase distinctive cognitive abilities due to prioritization and usefulness, as they are relevant in this culture and represent an advantage for their members. Sternberg (2007) created the expression *successful intelligence* as “what is needed for success in life”. Success in industrialised modern western societies depends on a specific kind of cognitive abilities trained in schools, universities and vocational trainings. Precisely these abilities (for example abstract reasoning and critical thinking) are measured in western intelligence tests (Sternberg, 2007). Using western diagnostic test batteries on a non-western sample, thus, produces

serious consequences in the test results and their interpretation. Van de Vijver (as cited in Trommsdorff, 2007) argued that intelligence cannot be a cross-cultural construct, if it is defined in a specific culture. This would mean that intelligence is not a shared feature between different cultures and IQ scores cannot be compared (van de Vijver, as cited in Trommsdorff, 2007; Sternberg, 2007). Non-western subjects, therefore, should not be tested with western diagnostic tests, nor compared with a western sample. Stanczak, Stanczak and Awadalla (as cited in Fasfous et al., 2001) demonstrated this with a neuropsychological assessment of healthy and brain damaged adults from Sudan and the US. They found that the healthy Sudanese group scored similar to the North American brain-damaged group.

Jensen (as cited in Trommsdorff, 2007) analyzed the inherent intelligence construct behind different diagnostic instruments and found that none of the current tests can be called *culture fair* (as Cattell titled them in 1963), and thus suggested to rename them *culture reduced*. To test a specific subset of intelligence test items is not enough to allow for culture free testing or to test general intelligence. Cole (2019) even titled this the “Illusion of Culture-free Intelligence Testing”. He argues that IQ Tests were deeply embedded in the cultural experience (e.g. the western schooling environment) from the first IQ Test conducted by Binet and his colleagues. Cole argues that abilities are cultural devices and that therefore cross culture testing is extremely difficult (Cole, 2019).

To summarize, the cultural background does influence culture free intelligence tests through experience with formalized knowledge and the training of cognitive abilities in school settings. Furthermore, diagnostic test batteries often do not measure the general factor of intelligence directly but specific cognitive abilities, which are influenced by culture and socialization. Thus, culture, school attendance, familiarity with test situations and computer-based tests, and mathematical understanding influence the score achieved in an intelligence test and should be taken into consideration when interpreting the test results.

In the present paper, the IQ test results from refugees from the Middle East and North Africa were used. The aim of the study is to quantify the relationship between IQ scores (of culture free intelligence tests) and culture as well as the educational level of the participants. We hypothesize that *schooling* (years of school attendance and university attendance) is *strongly and positively related to the IQ scores gained with a culture free intelligence test*. Consequently, that *the educational level of the subject can predict these IQ scores (Hypothesis 1) and academics reach higher IQ scores than non-academics (Hypothesis 2)*.

Since education standards differ among countries, we furthermore hypothesize that *country of origin is strongly correlated with IQ scores (Hypothesis 3)*.

As mentioned earlier different academic fields (mathematics, engineering, science, etc. vs. arts, the humanities) should also prepare subjects differently for IQ testing. Thus, *field of education/work should also influence IQ scores (Hypothesis 4)*.

To assess the influence of culture and formal schooling on intelligence test scores a sample of refugees from different countries were tested in Germany. Testing took place via the Charity Association “FairWelcome e.V.”. These Assessment Centers for refugees aim to assess, select and train participants for the German labor market and if necessary, for further vocational trainings, in a workshop setting. Maltby et al. (2013, cap. 11) recommend using intelligence tests to predict success at work. The Assessment Center evaluated participant’s job preferences as well as their intelligence levels. These results were then used for job-recruiting and individual job consultation. Three tests were used. In 2015 a culture fair intelligence test *iqtest*¹ certified by the International High IQ Society was used. The *iqtest* measured the IQ score with a ranking scale. In 2016, the culture free *IQ Test DK* was used which provided interval scaled IQ scores. Since 2017, thanks to a cooperation with the

¹ Retrieved from: <http://iqtest.dk/main.swf>

publisher and author of the *CTF-20-R Test* (culture fair test), Prof. Weiß (2006), the IQ scores could be measured with a cardinal scale test.

Method

Coding Procedure

Ranking Countries by Educational Level (1 to 3)

As suggested by the Global Education Monitoring Report from UNESCO², the sample is subdivided by nationality and by the alphabetization of the population. We ranked the country by the percentage of citizenship, which completed the secondary education, which means the pupils have attended for at least 4 years the primary school and continued to study after it. Countries with compulsory school attendance (mostly of 10 years) ranked higher.

All countries in which 50% or more of the population completed the secondary education level (mostly 10 school years) assigned three points. Countries ranging between 26% and 49% of alphabetization get two points. The remaining countries (in which less than 25% of the population completed secondary education) become one point (see Table 1).

Table 1 *Overview of the countries as well as the number of participants in the education level ranks. Countries are ranked according to the alphabetization level in the population. Percentage of alphabetization in the population in brackets.*

educational level rank	Country of origin	number of participants
1 (< 25%)	Afghanistan, Bangladesh, Iraq, Cameroon, Congo	21
2 (26% to 49%)	Egypt, Kosovo, Palestine, Sudan, Syria	34

² The *Global Education Monitoring Report* (the GEM Report, formerly known as the *Education for All Global Monitoring Report*) is an independent editorial, authoritative and evidence-based annual report published by UNESCO. The GEM Report is a singular, comprehensive, analytical and authoritative reference for the global follow-up and review of education. Retrieved from: <https://en.unesco.org/gem-report/>

3 (< 50%)	Iran, Libya, Ukraine	59
	total	114

Ranking by Professional Field and Educational Level (0 to 3)

We collected information about university attendance and field of profession due to the hypothesis of a strong positive correlation between the mathematical and natural science fields and IQ score (as seen in Bergold et al., 2017, Stemmler et al., 2016 and Rost, 2013).

Participants without a formal training or a profession, which did not require any specific knowledge or training, get 0 points in the field of profession ranking (Group of “unqualified workers” and participants without a profession).

The second rank (one point) was assigned to workers in non-academic fields such as mechanics, building and other kinds of crafts- and tradesmen ships (such as goldsmiths, dressmakers, plumbers), as well as drivers, electricians, office workers, salesmen (clothes, medicine, cars, etc.), cosmetician and hairdresser, employees in the field of catering, tourisms and fashion, police officers, etc. (Group of “hand-crafters”).

The third rank (two points) was given to people with an academic degree in arts as well as teachers (not including teachers in mathematics and science) alongside all academics and employees in the fields of humanities (law, psychology, literature, translation, journalism, pedagogy) and any kind of management and logistic, veterinarian and librarian jobs. (Group of “academics”).

We grouped all degrees in natural science, technologies, IT and mathematics to the fourth rank with three points. This group included technicians, computer engineers and other engineers, business administrators, financiers, economists and accountants, architects, urban planners and surveyors, agricultural scientists, biologists and microbiologists, pharmacists and mathematicians (also teachers), (Group of “scientists”). When subjects indicate two different

fields of profession (or an academic degree and a vocational training), the higher rank (three points) was assigned, (see Table 2).

The fourth rank is considered be the best predictor for IQ thanks to the training of formal logical thinking: e.g. Weiß (2006) showed the difference in IQ score between different school curricula. Pupils attending the mathematical focused curriculum got the higher IQ scores. Stemmler et. al. (2016) found strong positive correlation between the IQ scores and the profession's prestige. We ranked academical profession higher, scientific, technical and mathematical fields at the top.

Table 2 *Overview of the professions and number of participants in the field ranks. Professions are ranked as unqualified workers (0), non-academic/handcrafters (1), academic - humanities and arts (2), and academic - natural science, technics, IT and mathematics.*

Profession	Field rank			
	0	1	2	3
Without a profession, cleaning jobs	8			
Sellers, Workers and craftsmen (construction, sculptor, goldsmith, hydraulics, cutter, welder), assistant and workers (office, gastronomy and tourism), electrician, mechanic (vehicle, motorcars), camera operators, drivers (truck) fashion designer, police officers, health care & wellness (masseur, beautician, hair stylist) instructors (swimming, fitness, carpet weaver)		35		
Humanities scientist and art (English/Persian literature, education, music, psychology, law, translators), teachers (history, tailor), journalist, librarian, management (logistics, production technology, public)			21	

Engineer (Agricultural, Civil, Computer, Information Systems, Electrical, mechanical, production), Architects (town planner), biologist, accountant, Business Administration and finance, IT (engineering, administration, web design), nautical officer, surveyor, Biology (biochemistry, pharmacy, microbiology), Mathematicians (also teachers), technician (also telecommunication)	51
Number of participants	115

Coding the IQ-Tests by measurement (ranked and interval scaled IQs). The cognitive tests were administered using standardized task instruction presented via visual screen display on laptops. Each test was explained thoroughly in Persian, Arabic or English to assure a correct understanding of the tasks. Each cognitive test was preceded by a practice trial and participants were required to successfully complete the practice trial prior to the test trial for each measure. The tests result and participants data (demographic information) were collected for all participants who completed the IQ test and divided in two group: the ranked scaled (Group 1) and the interval scaled IQs (Group 2).

Group 1 (N = 115): Twenty-nine participants took the computerized online culture fair intelligence test certified by the International High IQ Society “iqtest” (n = 29). The test measured the IQ score with a ranking scale in 11 ranks and a subdivision into five categories: Low (between 1 and 2), below average (3 - 4), average (5 – 7), above average (8 - 9) and high (10 – 11). We assigned the rang order 1 for low score, 3 for below the average, 5 for average, 8 for above the average and 10 for high. The intelligence test consists of twenty-five questions in 30 minutes measuring the non-verbal ability to reason logically. The test avoids language biases and is based on the validated Raven’s progressive matrices test. The participants can compare their results with national fellows or with people from other countries. To provide comparable data, we reduce the entire sample at an IQ rank scale with following rank score

("<70" for the scores 1 to 2, "70-84" for 3 to 4, "85-111" for 5 to 7 and ">111" for 8-10). The ranks correspond approx. to the regular standard deviation in IQ score ($SD = 15$).

Group 2 ($n = 86$): Eighty-six participants received interval scaled IQ scores. This group was tested with the IQ Test DK or with the *CFT-R 20* Test (German translation by Weiß, 2006). The DK with 39 questions in 40 minutesⁱ is designed to test learning capability, memory, innovative thinking and the ability to simultaneously address several problems. The test measures the general intelligence - g-factor and is inspired by Raven's Advanced Progressive Matrices. The calculation of IQ scores is based on results from more than 250,000 people and provides interval scaled IQ scores. The computerized intelligence Test *CFT-R 20* is a cognitive battery measures domain of cognitive functions including logical and abstract reasoning and extract figural analogies. The test is based on the Theory of Cattell and his measure of the fluid intelligence. The *CFT-20* revisited in 2006 is the German Version. The *CFT-20* is divided in two parts, which can be used singularly. We used the short part with 56 items, which should take 40 minutes to be completed. The subgroups of the *CFT-20* include a subtest on continuing logical progression, classification, matrices and conditions. For the second group, the test instructions were prolongedⁱⁱ by translating them accurately into Arabic or Persian and with a short explanation on the test in general using the example exercise in the manual.

Sample description

The entire sample consisted of 115 subjects. The subjects live in Germany and asked for asylum. The participants could speak and understand German at least at the B1 level and were seeking a job or a vocational training. The basic demographic data were collected: gender (male, female), age (in years), country of origin, schooling (in school's years), university study (yes, no) and duration of study, vocational training (yes, no).

The sample consisted of 31 female (27%) and 84 male (73%) participants ($N = 115$). The Age ranged between 18 and 50 ($M = 31.6$, $SD = 7.0$, smallest mood = 28).

We collected the country of origin for 114 participants: Fifty-seven of the tested participants were refugees from Iran (49.6%) and 22 were Syrian (19%), (see Table 3 for all other countries ranks and frequencies).

Table 3 *Detailed list of the number of participants in each educational level rank as well as their country of origin.*

educational level rank	number of participants	percentage of participants	country of origin (n)
1: < 25%	21	18.3	Afghanistan (13), Bangladesh (3), Iraq (2), Cameroon (2), Congo (1)
2: 26-49%	34	29.6	Syria (22), Egypt (9), Kosovo (1), Palestine (1), Sudan (1)
3: >50%	59	51.3	Iran (57), Libya (1), Ukraine (1)
total	114	99.1	

Schooling

In most of the cases the participants attended the primary and the secondary education level except for two subjects, who gave no information on their educational status. The average schooling duration is: 11.36 years ($SD = 2.1$). Eighty-six percent of the participants ($n = 99$) attended twelve years of school. Seven participants (6%) stated that they attended school for ten years (ten years of school is the standard duration for compulsory education). One subject attended school for eleven years. Few participants attended school less than ten years. Two subjects went to school for four years; and one subject each attended five, seven, eight and nine school years, (Table 1 shows the list of countries of origins as well as the ranking score of each country for the sample).

The attendance was counted regardless of successfully passing the final exams. The sample contained mostly well-educated participants. The *Statistical Report of the German Ministry of Migration and Refugees* (BAMF) in March 2019 collects data from the full sample

of refugees, which applied for Asylum in Germany. These data are comparable to our sample. This almost complete population statistics gives comparable data about the social demographical and educational data with our sample: Three-quarter of ca. 43,000 adults provided information on school attendance. Refugees from Syria made up the largest group. About half of adult asylum-seekers were aged from 18 to 29, approximately 60% were men. 31% named a secondary school as their highest educational institution while 11% had not received any formal schooling. Asylum-seekers from Iran have the highest level of education (Heß, 2019).

As mentioned above, the countries were grouped in a rang order following international statistical data (see *Global Education Monitoring Report* of UNESCO) on education and spread of literacy on the population, (see Table 1). According to this scale we found a mostly high educational level (rang 3) for 59 subjects in the sample (51%). That means: Most of our subjects came from countries with a system of free and compulsory schooling, where more than half of the population reached the secondary educational level (mostly 10 years of schooling). This part of the sample consists nearly exclusively of subjects coming from Iran ($n = 57$).

The second largest group in the sample was formed out of citizens from countries in which less than 50% complete the secondary education (with at least 11th grades). For 34 of the subjects the secondary educational level was completed. Most of the people which attended the school longer as the state requests as compulsory educational and got the 12th grades ($n = 22$) came from Syria (retrieved from UNESCO statistical data: since 2013 less than 50% complete the secondary education), the remaining nine came from Egypt and one each from Kosovo, Palestine and Sudan. In all these countries less than half of the population attends school until they reach the age of 17.

The remaining 20 subjects came from countries with very poor literacy in the population. In these countries, less than 25% of the population complete the secondary

education level (18% of the sample). That is the case predominantly for Afghanistan ($n = 13$), then for Bangladesh, Iraq, Kongo and Kameron (see Table 1).

University Attendance

Participants were asked to provide information on their university background. We collected the data from 114 participants. The majority ($n = 75$; 65,2%) attended university in their home country. 35% of the remaining subjects ($n = 40$) did not attend university. Forty-three of the university attendees (57%) completed their studies in four years (mean attendance time worldwide until a degree). Nineteen participants did not complete their studies and attended less than four years. The remaining 12 participants completed longer university courses (e.g. teachers) or took longer time to finish theirs studies (see Table 4).

Table 4 *Detailed list of the years of university attendance and the number of participants in each group.*

academic years	number of participants	percentage of participants
0	40	34,8
1	5	4,3
2	9	7,8
3	5	4,3
4	43	37,4
5	6	5,2
6	3	2,6
7	2	1,7
8	1	0,9
total	114	99,1

Vocational training and training-on-the-job

The participants were asked to give detailed information about their vocational training and work experiences. Only few countries seem to adopt standardized models in worker qualification (as is the case in Germany). Usually participants were trained on the job in order to be qualified for jobs on the hand-crafter's level.

Seventy-six participants (66%) did not attend a vocational training or a training-on-the job to learn a profession but 39 participants do (34%). Only 17 participants (6,7%) neither completed a training nor attended university. Twenty-three refugees (5%) completed a vocational training or had a training-on-the-job. Sixteen participants (7%) attended university and completed a vocational training.

Field of study and work (0 to 3)

The sample was divided in 4 subgroups according to the qualification level and to the working field of the participants. The score 0 was given to the “unqualified worker” ($n = 8$; 7%). To the group of the “hand-crafters” (score 1) belonged 35 participants (30%). The “academics” (score 2) reach 18% of the sample ($n = 21$) whereas 51 subjects (44%) were in the group of the “scientists” (score 3), the most numerous group.

Results

IQ scores of the sample

All participants ($N = 115$) got a ranked IQ score (Group 1). Seventy-three percent of the participant ($n = 84$) reached an IQ score within the mean range (between 85 – 111). 15.7 % ($n = 18$) reached higher IQ scores (>111) and 9.6% ($n = 11$) showed IQ scores between 70 and 84, two people (1,7%) scored less than 70. (See Table 5 for the distribution in rank scores of the sample).

Table 5 *Overview of the number of participants in each IQ Rank Score.*

IQ rank	number of participants	percentage of participants
IQ < 70	2	1,7
IQ 70-84	11	9,6
IQ 85-111	84	73
IQ > 111	18	15,7
total	115	100 %

Participants in the second group were assigned interval scale IQ scores (Group 2, $n = 86$). The mean IQ score amounts 99.57 (average intelligence score), the median value is 99.57 with standard deviation $SD = 11.65$. The participant at the bottom of the IQ score gained 63 (min). The best score was 127 (max.). The sample seems not to be (perfectly) normal distributed.

Statistics and data analysis

First Hypothesis

Schooling (years of school attendance and university attendance) is strongly related to the IQ score. To test this hypothesis a spearman correlation between duration of schooling and the IQ rank score of the whole sample was done. To account for a longer time in formal schooling due to university attendance, one year of schooling was added for participants which attended university. The participants with longer years of schooling ranked significantly higher in IQ scores: $r_s(115) = .36, p < .001$. The correlation between schooling and the IQ scores in the entire sample is also positive and strong: $r_s(115) = .32, p < .001$.

Consequently, the educational level of the subject should predict the IQ score. For this analysis school attendance and university attendance was separated again. The model of linear

regression between IQ and the duration of school attendance was non-significant. A simple linear regression was calculated to predict the IQ score based on duration of university attendance. A significant regression equation was found ($F(1,83) = 7.469, p < .008$), with an R^2 of .083. Seven percent of the variance on the IQ score depends on the duration of university attendance. Thus, studying at university is a significant predictor for a higher IQ score ($t = 2.733, p = .008$).

Second Hypothesis

Academics reach higher IQ scores than non-academics. A t-Test was done to answer the second hypothesis. Therefore, only the group with interval scaled IQ scores was used ($n = 86$). Participants who did not attend university ($n = 33$) reached an average IQ score of 95.91 ($SD = 13.37$). The academics and scientists ($n = 53$) reached an average IQ score of 101.85 ($SD = 9.9$). This difference of 5.94 IQ points is significant ($t(84) = -2.35, p = .02, d = 0.32$).

To increase power we also compared academics and non-academics in the entire sample (ordinal scaled data, $N = 115$) with the Kruskal-Wallis-Test. The results showed a significant difference between the academics (middle rank = 63.17, $n = 17$) and non-academics (middle rank = 48.31): $H(1), 8.55, p = .003$. The effect size is 0.075.

Third Hypothesis

The country of origin is strongly correlated with the IQ scores. To test this hypothesis a Spearman rang correlation was used. We hypothesized that higher IQ scores could be for countries with a higher education standard, which is indeed the case, $r_s(114) = .17, p = .039$.

Fourth Hypothesis

Field of education and/or work influences the IQ scores. Spearman's rang correlation was able to confirm that a more academic/scientific correlates positively with higher IQ scores, $r_s(114) = .38, p < .001$.

Discussion

In a natural quasi-experiment refugees from different Middle East and African countries attended culture free intelligence tests. We found a significant positive correlation between the length of schooling and the results achieved in the culture free intelligence tests. As was shown above, formal education is beneficent to the results obtained in formal tests (amongst others see Helfrich, 2013; Flynn and Williams, as cited in Maltby et al., 2013; Fox & Mitchum, 2012). Thus; our sample shows the relationship between formal education and performance in test situations. Participants who were enrolled at university showed higher tests results than non-academics, furthermore, strengthening the relation between formal schooling and performance in formal testing situations. The results showed, that participants IQ scores were higher the more academics their profession. This relationship between education and testing could be interpreted in two ways. On the one hand, persons with a higher IQ might attend school for a longer time and thus are enabled to enroll at university and/or choose a more academic profession because of their higher cognitive abilities. On the other hand, higher IQ scores could be achieved because the subjects were schooled for a longer time and therefore have more training in formal testing situations, the tasks used in these tests, and the form of thinking in these tests. This second view supports studies proposing, that the distinct processes used in intelligence testing are trained through formal education (a. o. Fox and Mitchum, 2012; Maltby et al., 2013; Bergold, Wirthwine, Rost & Steinmayr, 2017; Poortiga, as cited in Trommsdorff, 2007; Brunner et al., as cited in Rost, 2013). This perspective might pose a problem for culture free intelligence tests, which were developed to rule out the influence of third factors (like schooling). If culture free intelligence tests are able to test intelligence independent of schooling, the IQ scores should not depend on the ratio of formal schooling in a country. Nevertheless, our data show this correlation.

If culture free intelligence tests are doing what they are designed to do, there should be a normal distribution of IQ scores as well as no difference between western and no-western test

results. Nevertheless, there is a marked difference between IQ scores found in our sample and a German norm sample of school pupils. The author of the *CFT 20-R* compared the different school curricula in Germany (see Table 6 and Appendices) and found different IQ scores for pupils attending the different curricula (Weiß, 2006). The pupils attending the “Realschule” (10 years of compulsory education) reached a mean IQ score of 103, and pupils attending the Gymnasium (12 or 13 years of schooling) reached a mean IQ score of 109. In our sample people who attended 12 years of school reached a mean of 100.9. Thus, our sample did not reach the scores of a German school sample. It is unreasonable to assume that the refugee sample is less intelligent than the German school sample.

Table 6 *Comparison of mean IQs of a german student sample (Weiß, 2006) and the participants in the current study (with interval scaled IQ scores) according to the years of schooling.*

School curricula	years of schooling	Mean IQ german student sample	Mean IQ refugees
Realschule	10	103	85,5 (n = 7)
Gymnasium	13 (Germany) 12 (most other countries)	109	100,9 (n = 75) 12 years of schooling

One problem, which occurred at testing, was unfamiliarity of the subjects with standardized test situations. The test instructions were translated, and participants were provided with a pre-testing situation to prepare them. Despite these preparations, 11% of the subjects scored one or two standards deviations below 100 IQ points. Many of the participants had never partaken in a formal (intelligence) test and had little or none experience using a computer (as the test was computer-based). As shown in the introduction this lack of experience with the test situation influences the test result (Van de Vijver, as cited in Trommsdorff, 2007; Fasfous et al. 2013) and might be the reason for the lower test results as compared to the German sample. Refugees and immigrants who need to enter the European labor market should thus be

trained in advance for these particular assessment and testing situations. This could at least provide a slightly fairer test situation. It is hard to imagine that the reduced IQ score, which were partly below 80, could represent the actual intelligence of the participants. Fleeing your home country in order to find refuge in a different one requires a high level of cognitive flexibility as well as high abilities to solve problems on quite different kinds of level, in order to adapt to the ever-changing situations. Given the fact, that the participants in the sample were representative for the refugee's population in Germany after 2015 (regarding age, gender and education) they can help us to gain insights into this population as well as the bias of a cultural free measurement of immigrants' abilities in European countries.

To summarize, the culture free intelligence tests used in this study showed theoretical and practical problems. These tests claim to measure the general factor of intelligence but their implementation (most of them are based on the Raven's Advanced Progressive Matrices) is considered the main concern regarding universality (Fasfous et al., 2013; Helfrich, 2013). We found significant positive correlations between the duration of the schooling and the IQ scores. The percentage of education and literacy in the country from which the subjects fled predicted the IQ scores with higher scores for a higher percentage of compulsory education and literacy in the home country. Furthermore, subjects who attended university showed higher IQ test results than subjects who did not attend and the field of study was also shown to correlate with IQ scores. Our results clearly show the relevance of academic schooling in intelligence testing. Interestingly, when correlating the attendance of a vocational training (or on-the-job-training) no significant correlation with IQ scores could be found ($r_s(115) = .006, p = .47$).

Culture free intelligence tests should therefore be used with caution. IQ scores do not equate *intelligence*; this difference should be made clear. The test results could list the cognitive abilities measured in the particular IQ test together with the test score and its confidence interval. This might avoid a misconception of the IQ score by the users and help to avoid generalizations about the participants' abilities.

Declaration of Conflicts of Interest

The Authors declare that no conflicts of interest exist.

References

- Boring, E.G. (1923). Intelligence as the tests test it. *New Republic*, 35(6), 35–36. Retrieved from https://brocku.ca/MeadProject/sup/Boring_1923.html on 20.03.2020.
- Bergold, S., Wirthwein, L., Rost, D. H., & Steinmayr, R. (2017). What happens if the same curriculum is taught in five instead of six years? A quasi-experimental investigation of the effect of schooling on intelligence. *Cognitive Development*, 44, 98-109. doi:10.1016/j.cogdev.2017.08.012
- Cole, M. (2019). *The Illusion of Culture-free Intelligence Testing*. Retrieved from: <http://communication.ucsd.edu/MCA/Paper/Cole/iq.html> on 20.03.2020
- Fasfous, A. F., Puente, A. E., Pérez-Marfil, M. N., Cruz-Quintana, F., Peralta-Ramirez, I., & Pérez-García, M. (2013). Is the Color Trails culture free? *Archives of Clinical Neuropsychology*, 28(7), 743-749. doi:10.1093/arclin/act062
- Faul, F., Erdfelder, E., Buchner, A. *et al.* Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149– 1160 (2009). <https://doi.org/10.3758/BRM.41.4.1149>
- Fox, M. C., & Mitchum, A. L. (2012). A knowledge-based theory of rising scores on "culture-free" tests. *Journal Of Experimental Psychology: General*, 142(3), 979-1000. doi:10.1037/a0030155
- Funke, J. (Ed.). (2009). *Was ist Intelligenz?*. München: Beck.
- Helfrich, H. (2013). *Kulturvergleichende Psychologie*. Wiesbaden: Springer VS.
- Heß, B. (2019). BAMF-Kurzanalyse. *Ausgabe 03/2019 der Kurzanalysen des Forschungszentrum Migration, Integration und Asyl des Bundesamtes für Migration und Flüchtlingen*, 03/2019, 1. Retrieved from: http://www.bamf.de/SharedDocs/Anlagen/EN/Publikationen/Kurzanalysen/kurzanalys_e3-2019-qualifikation-asylstantragstellende-1hj-2018.html?nn=1449080

International High IQ Society, (o. J.). *The culture fair intelligence test*. Retrieved from:

<https://www.123test.com/culture-fair-intelligence-test/> und

<http://iqtest.dk/main.swf>

Kultusminister Konferenz – KMK. (2017). *Basic Structure of the Educational System in the Federal Republic of Germany. Diagram*. Retrieved from: https://www.kmk.org/fileadmin/Dateien/pdf/Dokumentation/en_2017.pdf on 20.03.2020

Maltby, J., Day, L., & Macaskill, A. (Eds.). (2011). *Differentielle Psychologie, Persönlichkeit und Intelligenz*. München: Pearson Studium.

Maltby, J., Day, L., & Macaskill, A. (2013). *Personality, Individual Differences and Intelligence*. Pearson Education.

Poortinga, Y. H. (2007). Dateninterpretation in der kulturvergleichenden Psychologie. In G. Trommsdorff (Ed.), *Theorien und Methoden der kulturvergleichenden Psychologie* (pp. 289-335). Göttingen: Hogrefe Verlag

Rost, D. H. (Ed.). (2013). *Handbuch Intelligenz*. Weinheim: Beltz.

Schmitt, M. & Gestenberg, F. (Eds.). (2014). *Psychologische Diagnostik kompakt: mit Arbeitsmaterial zum Download*. Weinheim: Beltz.

Stemmler, G., Hagemann, D., Amelang, M., & Spinath, F. (Eds.). (2016). *Differentielle Psychologie und Persönlichkeitsforschung*. Stuttgart: Kohlhammer.

Sternberg, R. J. (2007). Who Are the Bright Children? The Cultural Context of Being and Acting Intelligent. *Educational Researcher*, 36(3), 148-155.
doi:10.3120/0013189X07299881

Schuler, H. (Ed.). (2014). *Psychologische Personalauswahl: Eignungsdiagnostik für Personalentscheidungen und Berufsberatung*. Göttingen: Hogrefe.

Van de Vijver, F. J. R. (2007). Methodologische und methodische Probleme des Kulturvergleichs. In G. Trommsdorff (Ed.), *Theorien und Methoden der kulturvergleichenden Psychologie* (pp. 337-382). Göttingen: Hogrefe Verlag.

Van der Maas, H.L.J.; Kan, K.-J.; Borsboom, D. (2014). Intelligence Is What the Intelligence Test Measures. Seriously. *Journal of Intelligence*, 2, 12-15; doi:10.3390/jintelligence2010012

Weiß, R.H. (2006). *Grundintelligenztest Skala 2. CFT 20-R (Revision)*. Göttingen: Hogrefe.

Appendices

Excursus: Explanation about the German school curriculum and grades.

The German school system consists of two levels. The primary schools are the first level. They are mandatory and encompass the first four or six years of schooling depending on the federal country (*Bundesland*). The second level includes five different curricula in which the pupils are able to obtain vocational or academic entrance qualifications (see Table 7, on the next page): (1) special school for children with disabilities or learning impairments (so called “*Förderschule*”) until the 10th school year (with the leaving certificate “*Mittlerschulabschluss*-MSA” or “*Realschulabschluss*”); (2) the so called “*Berufsschulen*” are vocational school at the upper secondary level generally providing part-time instruction in general and vocational subjects until 12th or 13th grades. The pupils can obtain a qualifications equivalent to the leaving certificate “MSA” and only under certain conditions the leaving certificate „*Fachhochschulreife*“ (12th grades) can be acquired on completion of a course lasting a minimum of two years. (3) The “*Hauptschule*” is at the lower secondary level and provides a basic general education until the 9th year. (4) The “*Realschule*” is also at lower secondary level but provides pupils with a more extensive general education and the opportunity to go on to courses of education at the upper secondary level that lead to vocational or higher education entrance qualifications. At the highest level (5) “*Gymnasium*” pupils are able to reach the general higher education entrance qualification (“*Allgemeine Hochschulreife*”) which entitles the holder to admission to all subjects at all higher education institutions and is usually obtained by passing the “*Abitur*” examination (Kultusminister Konferenz – KMK, 2017).

Table 7 *School Curricula in Germany*

School curricula	Curriculum grades
1) Förderschule	10 th grades
2) Berufliche Schule	13 th grades
3) Hauptschule	5 to 9 th grades
4) Realschule	5 to 10 th (or up to 13 th)
5) Gymnasium	13 th grades