Abstract

Different methods of scoring the multifactor general knowledge test were evaluated to test what the best method of calculating scores is. The best methods of calculating scores were treating every checkbox as an item and adding up all of the scores of the items or creating IRT models of the answers and distractors separately. The test was highly reliable ($\omega = 0.93$), had a low median completion time (571 seconds), and a high ceiling (IQ = 149). One set of items (internet abbreviations) were found to have very low g-loadings, so we recommend removing them. The test also had age, national, and gender differences which replicate previous literature.

The test was clearly biased against non-Anglos, especially in the sections of aesthetic knowledge, cultural knowledge, literary knowledge, and technical knowledge. DIF testing could not detect this bias, calling into question its usefulness in identifying highly biased tests. Between sexes, DIF found that many items were biased against both genders, but the magnitude of the bias did not vary by either sex. We recommend scoring this test by adding up all of the distractors and answers, and the use of a cultural and linguistic translation on an international scale.

1. Introduction

General knowledge tests are highly g-loaded (Johnson et al., 2004; Voronin et al., 2015), can be completed quickly, and require little effort from participants. This makes it an ideal subtest for any test that intends to measure general intelligence, unfortunately, some batteries neglect to test it, like the Stanford-Binet V (Roid & Pomplun, 2011) and the General Aptitude Test Battery (Palmer et al., 1990).

Some research has intended to examine whether the manner in which a test is scored affects how valid it is. For example, it could be possible that a knowledge test that uses free responses is testing a different retrieval mechanism than the ones that use multiple choice mechanisms. The most well-replicated and consistent differences between multiple choice formats and free response formats are that the former are faster to test and have easier items (Polat, 2020; Sirota & Juanchich, 2018; Breuer et al., 2020). The difference in validity between the versions appears to be null (Sirota & Juanchich, 2018) or to even potentially favor multiple choice tests (Breuer et al., 2020).

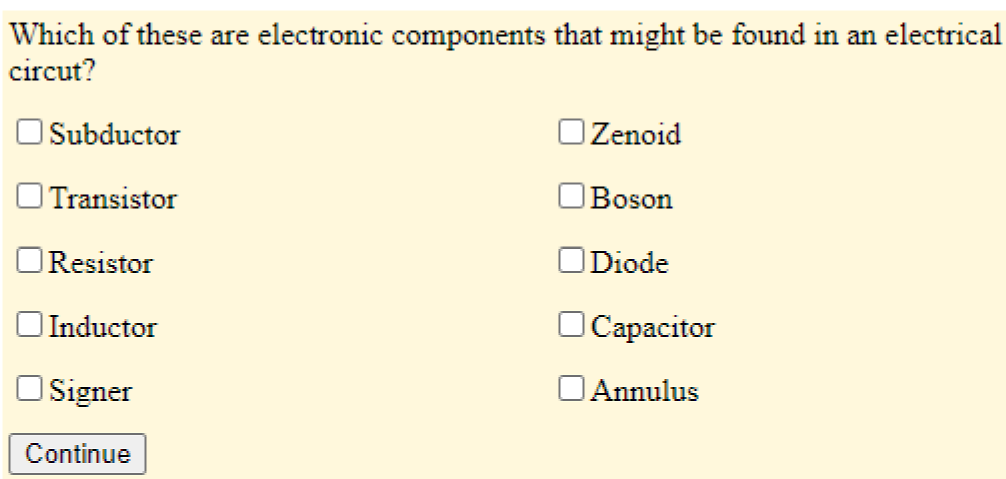
Typically, people are skeptical of online tests, as they are unsure whether they are valid measurements of intelligence. Some researchers have attempted to gather samples who take both an online and offline test (Young & Keith, 2020; Logos et al., 2021). Overall, offline and online IQ tests correlate at 0.57 on average, compared to an average correlation of .77 found between two offline tests (Jensen, 1980). Typically, offline tests return lower IQ estimates than online ones - The openpsychometrics FSIQ test underestimated scores by 5.7 points in college students and the VIQ test underestimated scores by 11.5 points in the same sample. There is also nothing innately different about an online and offline test besides the medium - what matters is the quality.

The primary goal of the study is to research the extent to which this test has predictive validity and is a valid measurement of general knowledge. As some characteristics (e.g. validity, reliability) are desired in a test, it was also tested whether scoring the test in different ways resulted in better outcomes. Nationality/sex bias was also tested for using differential item functioning (DIF). As data on the participant's country of origin was provided, it was tested whether there were international differences in scores on the general knowledge test.

2. Data

Data were taken from the openpsychometrics website, which contained a dataset of 19218 individuals who took the Multifactor General Knowledge Test. The test consists of 32 general knowledge questions in which a participant is asked which of 10 items satisfies a particular criterion (e.g., "Which of these are electronic components that might be found in an electrical circuit"). Five of the 10 items are correct for each question. An example question is displayed in Figure 1 for clarification.

Figure 1. Example item from this test.



Data regarding the individual's gender, age, English proficiency, time spent on each question, nationality, screen height, and screen width was also available. Information on screen height and width was then used to infer device type, as some resolutions (e.g. 360x640) are typical of mobile phones. This allows the users to be dichotomized into two different categories: desktop users and mobile users. By default, people were assumed to be desktop users, unless their particular resolution was very suggestive of the use of a mobile device. Data regarding the time taken to complete the test was collected along with the time spent on individual items.

Gender was coded as 1 = Male, 2 = Female, and 3 = Other. There were several individuals coded as ‘zeroes’ in the source dataset, which are presumably missing data, though the codebook does not explicitly state this. In the Male-Female comparisons, individuals who are coded as a 0 or a 3 are excluded from the analysis. Some individuals also reported unrealistic age values (twelve above 200, one of 102), so as a precaution, all individuals with an age of above 100 were set to missing values.

Individuals whose first language was not English were excluded from most of the analysis, as well as those who spent under one second on one question. This is because most questions take much more than one second to answer, so they are probably reflective of mistakes or low effort attempts. The data had also come with the removal of individuals who were under 13 and those who said they did not provide accurate answers.

3. Methods

There are several approaches that can be taken to score these items. Listed, they are:

1. Treating every individual checkbox as an item, which leads to a test of 320 items. Then, a numeric score out of 320 is calculated. (summed scores)
2. Treating every individual checkbox as an item, then doing an IRT analysis on the distractors and correct answers separately, then adding up the two IRT scores. The reason the IRT analysis is done on the distractors and correct answers separately is that IRT will falsely assume that the distractors are correct answers, no matter what direction they are coded in as. (160 + 160 IRT). One problem with this method is that it violates the local independence assumption, as answering one item from a question correctly corresponds to an increased likelihood in

doing this for all items in the question. In this paper, IRT analysis was conducted using the ‘mirt’ package (Chalmers, 2012).

3. Adding up all of the correct answers to the individual questions, then doing a graded IRT analysis of all 32 questions to generate one general score. (32 IRT)
4. Adding up all of the correct answers to the individual questions, then calculating the first general factor from the 32 questions. (32 FA)
5. Adding up all of the correct answers to the individual questions, then calculating the first principal component from the 32 questions. (32 PC)

In approach 2, 4 different approaches were taken when evaluating the IRT items themselves. Three of them involve the inclusion of various levels of logistic parameters, with one model having 2, one having 3, and another having 4. The difference between the three methods is the following:

- Two parameters (2PL): allows the discrimination and difficulty to vary by item.
- Three parameters (3PL): allows discrimination, difficulty, and lower asymptotes to vary by item.
- Four parameters (4PL): allows discrimination, difficulty, and lower/upper asymptotes to vary by item.
- Splines: uses splines instead of logistic parameters to evaluate the test.

These 5 methods will be graded on 4 criteria:

1. Reliability. This is the most important criteria for the test to abide by, as a more reliable test result will lead to a better understanding of the world. For the methods where there is not a convenient way to measure reliability, the odd and even items were correlated and then the spearman-brown formula was applied to calculate the estimated reliability. Otherwise, the omega reliability (denoted ω) was used to estimate reliability.

2. Correlation with other scoring methods. Better scoring methods will correlate more with other scoring methods as they are measuring more signal and less noise.
3. Sex bias. This is an unconventional criteria, given that researchers generally try to avoid biased tests. However, because men typically have more general knowledge than women ($d = 0.53$) (Tran et al., 2014), a more valid measurement of it would have a difference closer to that mean.
4. Age bias. Humans accumulate knowledge as they grow older, so a method that identifies a larger effect of age is also probably more valid.
5. Desktop advantage. Individuals who use mobile phones tend to be less intelligent than those who use other types of devices (Brown et al., 2023; Wilmer et al., 2017), so a method that returns larger advantages in favor of desktop users will probably be a more valid metric.
6. Nationality bias. When data of the mental ability of different nations is collected, some Nations are more intelligent than others on average (Lynn & Becker, 2019). A method that generates greater differences between nationalities is also more likely to be valid.

DIF testing was used to assess whether certain items had sex biases, that is, whether certain items exhibit a gender difference in the probability of correct response when controlling for general ability. This was initially considered to test for bias in favor of certain national groups, but the method was found to be ineffective, so it was only used to assess bias in testing between German countries (Germany, Switzerland, and Austria) and Anglo countries (US, UK, Australia, New Zealand, Ireland). Bias-adjusted differences were computed by calculating the difference between the invariant and partially invariant fits, where a conservative threshold was used to detect bias ($p < .05$, bonferroni adjusted). To facilitate country comparisons, countries were grouped into the following categories:

- Anglo
- Latin American

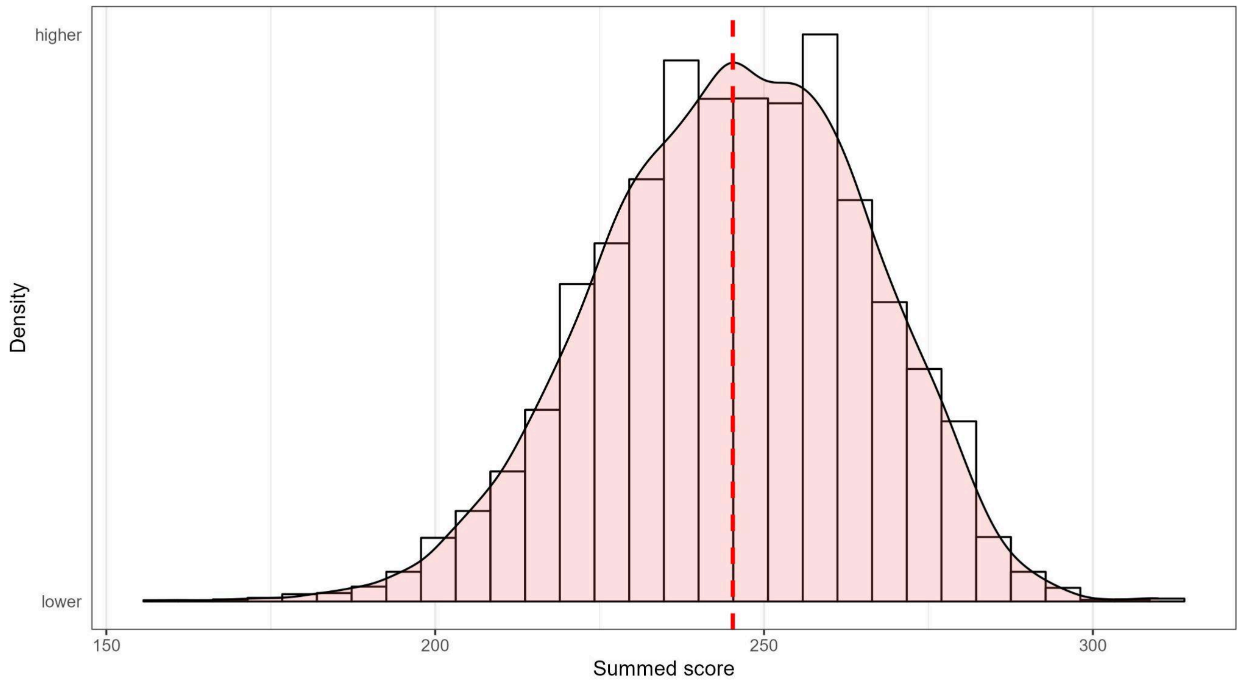
- Germanic
- Northern European
- Eastern European
- Balkan
- Caucasus
- MENA
- South Asian
- African
- East Asian
- South East Asian

To test the existence of a latent sex difference in intelligence, the method of correlated vectors was used. Given the large number of items, the method has a high power to detect a difference. Various methods of extracting g-loadings were considered which used different numbers of logistic parameters, ranging from 2 to 4.

Three methods were considered for the norming of the test - calculating the percentage of people who the individual scored at or higher than a given score, using a linear regression model which predicts the converted IQ score based on the summed score, or calculating the z-score based on the mean and the standard deviation. The first method works well when there is a very large sample size and there are departures from normality within the test. The 2nd and 3rd methods work well with a small sample size, but are sensitive to departures from normality within the scores of the test.

Given that the distribution of raw scores was almost identical to a normal distribution, based on the plot in Figure 2, and that the sample size of English speakers was very large (n = 13,696), all methods are acceptable ways to calculate norms. While there were small inconsistencies between the different methods at the tails of the test, all three methods of calculating IQ-based scores resulted in almost identical norms. A second set of age-specific norms was made to account for this problem.

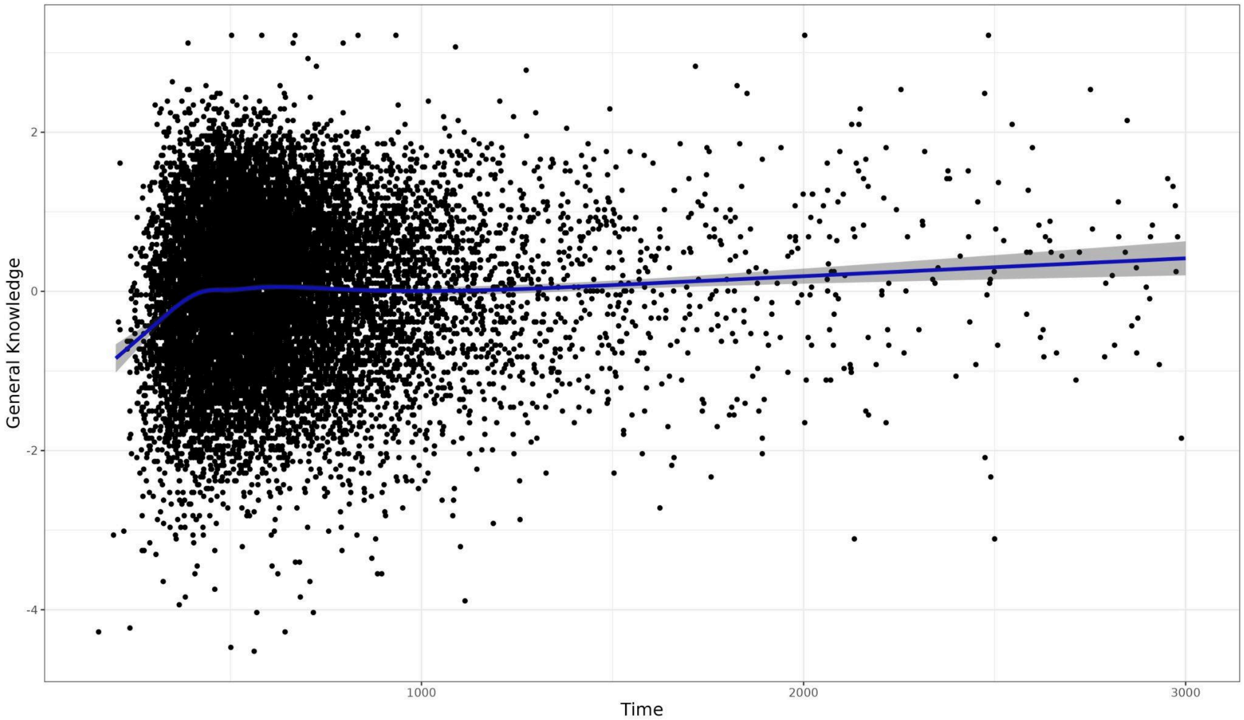
Figure 2. Density plot of the summed scores (internet abbreviations question excluded).



For the age based norms, z-score norming was used, as the sample sizes within the age cohorts are much smaller. Age norms were generated for specific ages from 13-30, and then for the age categories of 31-50 and 51-70. The predicted average score for every cohort was calculated using the restricted cubic splines. The predicted average standard deviation was also calculated, as the standard deviation is lower when the test takers are younger, also based on a restricted cubic splines method. Questions with very low g-loadings (< .15) were excluded from the calculation of norms, which only includes internet abbreviation (g-loading = .14).

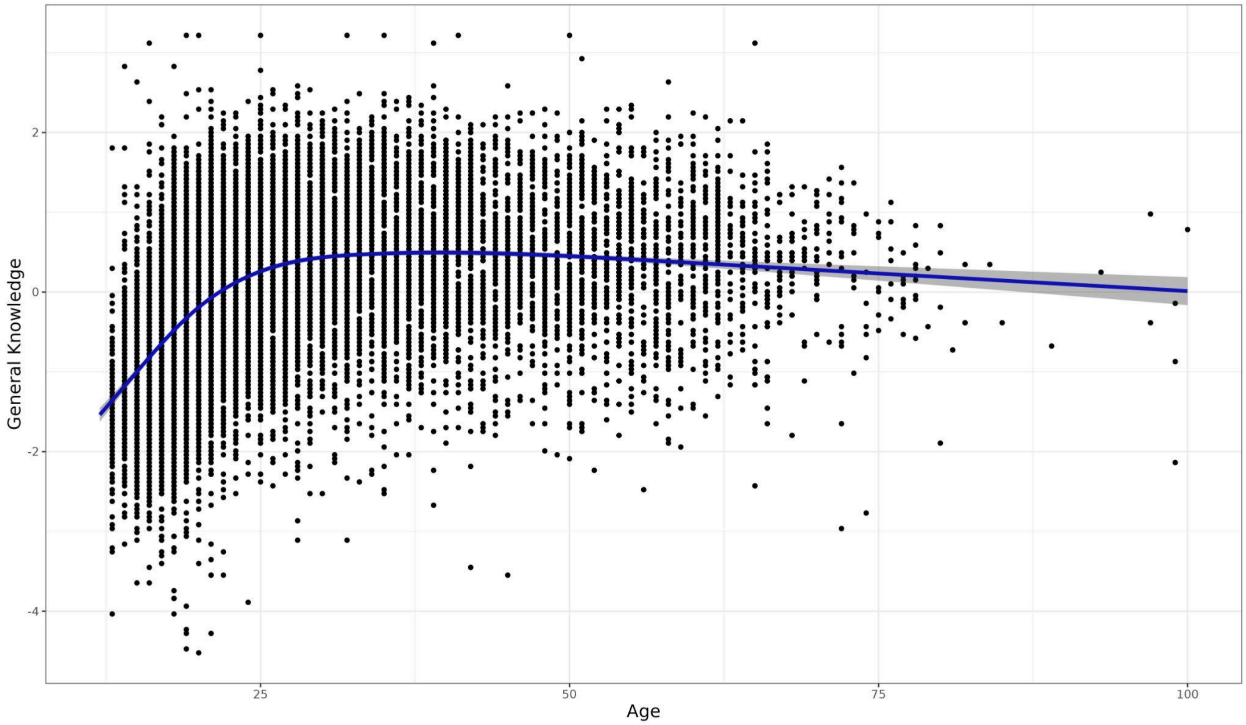
The effect of age and the time taken to do the test on the result of the test were also calculated to observe whether there was a notable age or effort effect. There was a small correlation between general knowledge and time taken to finish the test ($r = .049$, $p < .001$). This effect was largely due to people who took less than 6 minutes to take the test, based on the non-linear fit, which is displayed in Figure 3.

Figure 3. Relationship between time taken to complete the test and general knowledge. The blue line is modeled with a restricted cubic spline.



Age had a non-monotonic relationship with general knowledge. Individuals gradually increased their general knowledge until their mid-30s, and after that there was a slow decline in observed scores, as shown in Figure 4. This is consistent with data from other researchers, who find that crystalized ability gradually rises until it peaks in the mid 30s, after which it gradually starts to stagnate (Rohwedder & Willis, 2010).

Figure 4. Relationship between General Knowledge by age, modeled with a restricted cubic spline (ages of above 100 excluded in the analysis).



4. Factor structure

The factor structure of this test was assessed to compute sex differences and national differences in general knowledge. The structure can be determined with two methods: build an intuitive model of the test using confirmatory factor analysis, and use factor analysis to extract additional factors from the data. To facilitate the analysis, the 320 items from the checkboxes were aggregated into 32 questions.

To evaluate the number of factors necessary to model general knowledge, parallel analysis was used. The number of factors that are necessary to evaluate general knowledge was judged to be 7. The results of the parallel analysis are available in Figure 5, and the results of an oblimin rotated factor analysis with 7 factors are available in Table 1. While using rotation methods can undermine the size of the general factor in the data, this can be ameliorated by using it as a basis to form a hierarchical model.

Figure 5. Parallel analysis of the 32 questions in the dataset.

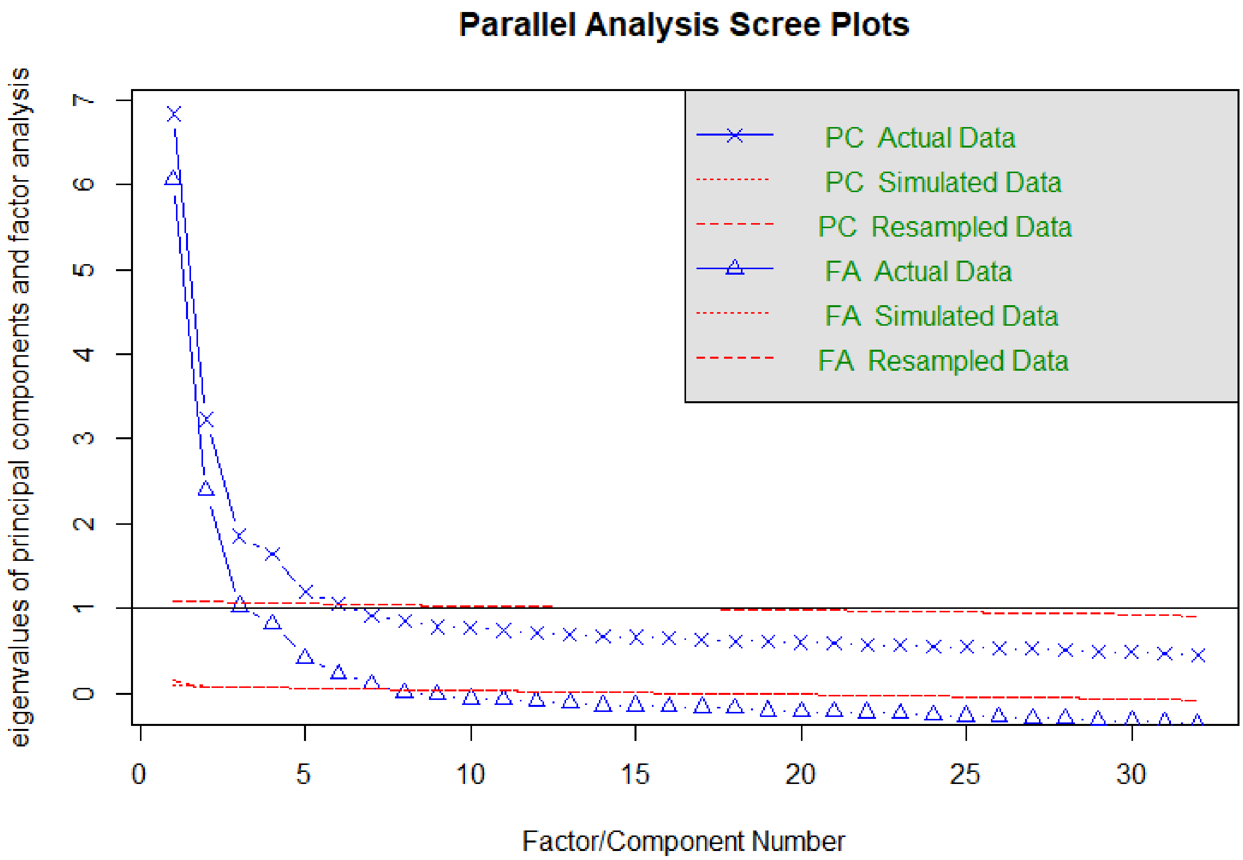


Table 1. Oblimin rotated factor analysis of the 32 questions.

Questions	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Loading (h2)
Q1	-0.13	0.05	0.24	0.11	0.47	0.02	0.06	0.42
Q2	-0.03	0.09	-0.03	0.31	0.44	-0.09	0.06	0.42
Q3	-0.02	-0.02	0.3	0	0.24	0.14	0.19	0.25
Q4	-0.02	-0.04	0.12	0.5	0.04	-0.21	0.19	0.36
Q5	-0.06	0.05	0.46	0.17	-0.1	0.15	0.14	0.35
Q6	-0.07	0.05	0.38	0.21	-0.05	0.08	0.13	0.29
Q7	0	-0.09	0.65	-0.06	0.1	0.03	-0.08	0.42
Q8	0.24	0.01	0.59	0	0.01	-0.1	0.01	0.43
Q9	-0.06	0.6	-0.03	0.08	0.08	-0.01	0.06	0.38
Q10	0.02	0.75	-0.1	0.04	0.06	0.02	0.03	0.58
Q11	0.05	0.67	0.08	-0.01	-0.03	0	-0.05	0.51

Q12	0.06	0.61	0.06	-0.12	-0.05	0.04	-0.08	0.43
Q13	0.6	0.05	0	0.05	-0.05	0.04	0.09	0.42
Q14	0.56	0.04	-0.06	0.03	0.03	0.01	0.11	0.36
Q15	0.61	0.07	-0.02	-0.03	0.03	0.03	-0.1	0.41
Q16	0.51	0.01	0	0.03	0.07	0.02	0.11	0.31
Q17	0.02	-0.01	-0.05	0.68	0.06	0.13	-0.05	0.53
Q18	0.11	0.06	0.15	0.17	0.03	0.42	-0.07	0.42
Q19	0.09	0.22	0.14	0.41	-0.04	-0.12	-0.07	0.3
Q20	0.07	0.02	0.27	0.2	0.14	0.23	-0.2	0.4
Q21	0.28	0.13	-0.01	-0.01	-0.12	0.42	0.04	0.4
Q22	0.41	0.24	0.06	-0.12	-0.12	0.02	0.14	0.35
Q23	-0.04	0.32	0.09	-0.09	0.11	0.07	0.05	0.15
Q24	0.02	0.3	0.49	0	0.08	0.03	-0.12	0.48
Q25	0.09	0.19	-0.02	0.02	0.46	0.03	-0.03	0.32
Q26	0.24	0.23	-0.05	-0.03	0.12	0.33	0	0.38
Q27	0.17	-0.04	0.11	0.4	0.16	-0.16	0.05	0.3
Q28	0.36	-0.11	-0.07	-0.17	0.21	-0.01	0.38	0.39
Q29	0.09	0.06	0.08	0.21	0.24	0.27	0	0.36
Q30	0.64	-0.02	0.15	0.06	-0.04	0.05	-0.12	0.49
Q31	-0.07	0.09	0.37	0.28	-0.05	0.21	0.15	0.42
Q32	-0.01	0.05	0.01	0.53	0.16	0.17	-0.09	0.48

Given that factor 7 was mostly associated with the internet abbreviations question, it was removed from the hierarchical confirmatory analysis.

A confirmatory factor analysis which models g as a second order latent variable was conducted based on these results, which was somewhat successful, yielding a CFI of .91, a SRMR of 0.063, and a RMSEA of 0.065. The model falls short of the traditional cutoff for CFI of 0.95, is within the SRMR cutoff of 0.8, and is close to the RMSEA cutoff of 0.06 (Hu & Bentler, 1999). However, the optimal cutoff can vary substantially depending on the model in question (Sivo et al., 2006), and the fit statistics are not poor given that the model is complex.

Correlated residuals were used in modeling due to the fact that in two cases some abilities continued to correlate moderately after the extraction of the general factor of knowledge. Best practice largely suggests that failing to include correlated residuals when necessary can lead to misleading results (Cole et al., 2007). A table of the correlation matrix of the residuals of knowledge subfactors is displayed in Table 2, which shows that computational knowledge and technical knowledge have a residual correlation of .21, while aesthetic knowledge and literary knowledge have a residual correlation of .22. The loadings for the items and latent factors are available in Figure 6 and the list of questions that each factor was associated with is available in the Appendix.

Figure 6. Confirmatory factor analysis of the Multifactor General Knowledge test. COK - computational knowledge, IK - international knowledge, CK - cultural knowledge, AK - aesthetic knowledge, LK - literary knowledge, TK - technical knowledge. DWLS (diagonally weighted least squares) estimation was used.

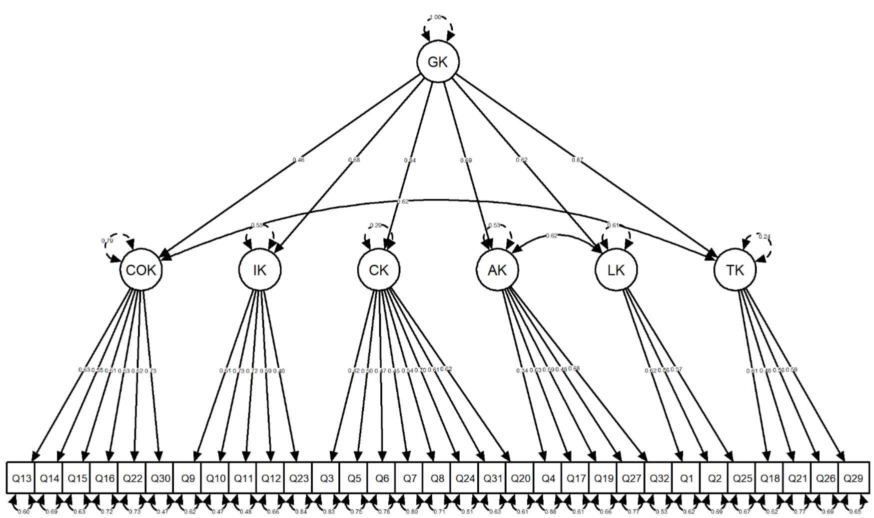


Table 2. Matrix of the residual correlations of the 6 knowledge subfactors of general knowledge. COK - computational knowledge, IK - international knowledge, CK -

cultural knowledge, AK - aesthetic knowledge, LK - literary knowledge, TK - technical knowledge. Correlations above 0.028 are significant at $p < .001$.

Ability	COK	IK	CK	AK	LK	TK
COK	0.68	0.10	-0.07	-0.18	-0.16	0.21
IK	0.10	0.64	-0.05	-0.07	0.00	0.02
CK	-0.07	-0.05	0.53	0.14	0.06	-0.05
AK	-0.18	-0.07	0.14	0.68	0.22	-0.08
LK	-0.16	0.00	0.06	0.22	0.74	-0.10
TK	0.21	0.02	-0.05	-0.08	-0.10	0.40

5. Results

Based on the results on Table 3, the best method to score the MGKT is to use the 160 + 160 IRT method (2PL) or the summed scores method. Due to the simplicity of the method, the summed scores method will be used for most of this study. However, when evaluating bias in the test, the 160 + 160 IRT (2PL) method will be used, as it is convenient to use to evaluate DIF.

Table 3. Comparison of the seven methods used to calculate general knowledge.

Method	Reliability	Loading on the general factor	Sex difference	Age correlation	Desktop advantage	National differences (averaged)
Summed Scores	0.93	0.989	-0.43	0.396	0.22	0.534
160 + 160 IRT (4PL)	0.91	0.94	-0.49	0.375	0.25	0.505
160 + 160 IRT (3PL)	0.93	0.963	-0.41	0.392	0.22	0.511
160 + 160 IRT (2PL)	0.93	0.958	-0.45	0.386	0.24	0.496
160 + 160 IRT (spline)	0.91	0.918	-0.45	0.371	0.22	0.505
32 IRT	0.89	0.983	-0.4	0.398	0.22	0.485
32 FA	0.88	0.991	-0.38	0.406	0.21	0.536
32 PCA	0.88	0.992	-0.38	0.407	0.21	0.531

† - only within those under the age of 25. Desktop advantage is the magnitude of the difference favoring desktop users over mobile phone users, the national differences coefficient is the average difference between every single country in the dataset.

Based on the results, the test contains a large amount of gender bias, where many items favor one gender over the other. Out of the distractors, 37 out of the 160 items displayed a pro-male bias, while 32 of the items had a pro-female bias. Within the answers, 48 of the 160 items had a pro-male bias, while 57 of the 160 items had a pro-female bias. Items with a pro-female bias typically were associated with literary knowledge, while items associated with a pro-male bias were typically associated with technological or international knowledge. The item probability functions by gender were calculated by using the leave one out method (LOO), where ability is calculated without taking a particular item into consideration, and then it was tested whether this measurement was biased in its ability to predict the excluded item. The mentioned probability functions are available in Figures 7 and 8.

Figure 7. Item Probability Functions of the distractors by gender.

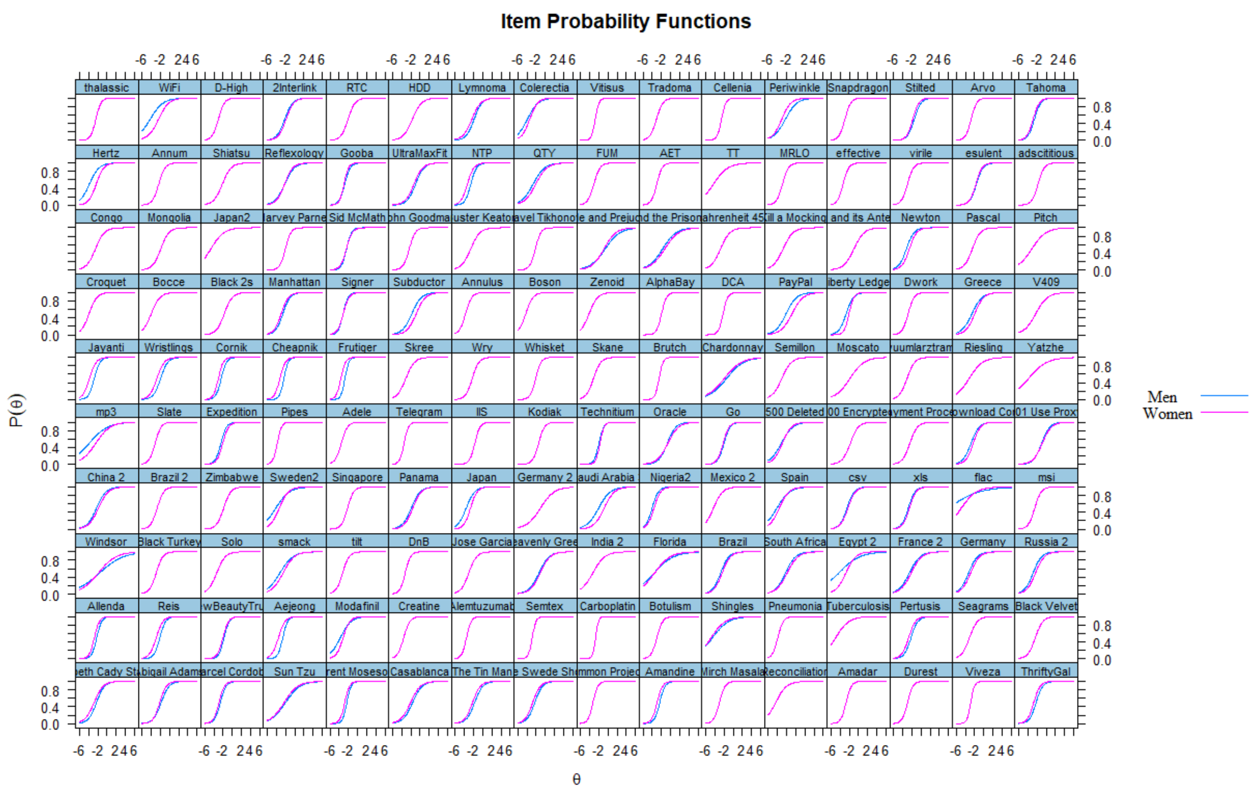
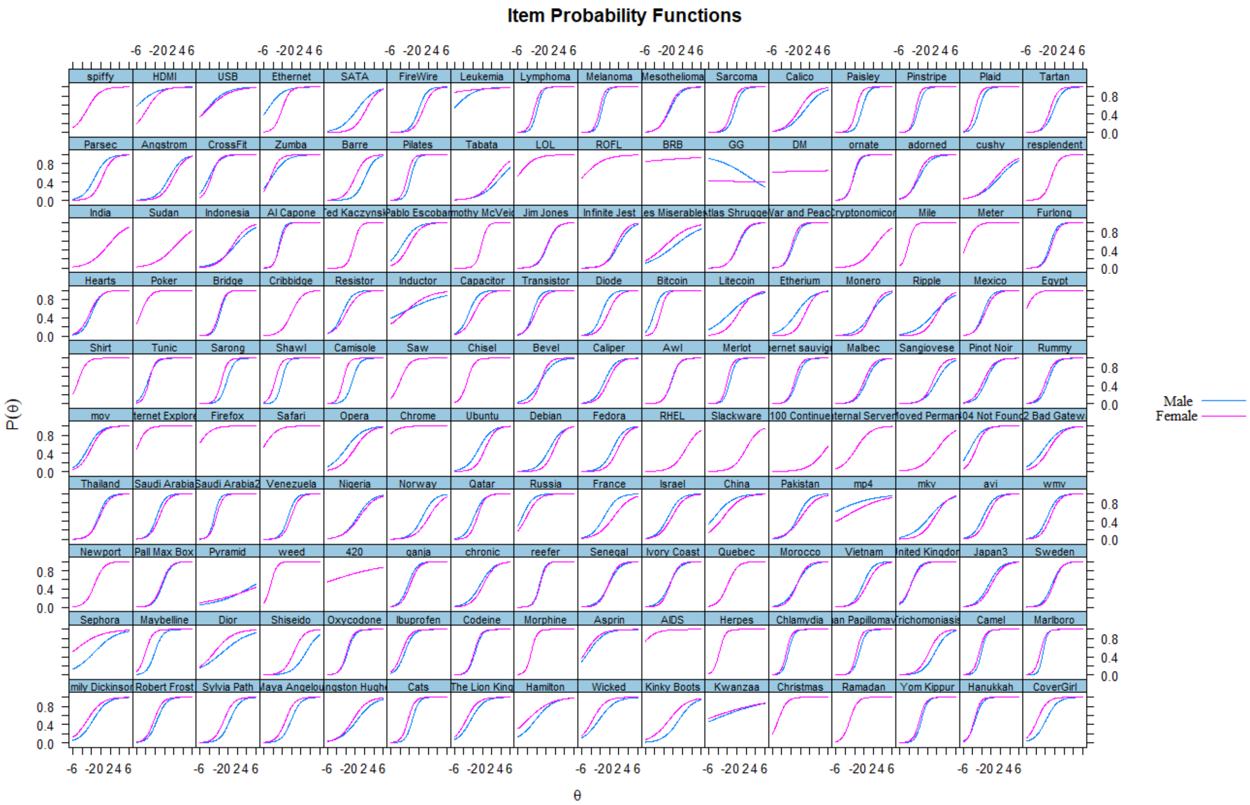


Figure 8. Item Probability Functions of the answers by gender.



The bias-adjusted (using a bonferroni correction) sex difference in general knowledge (d = 0.4683) hardly differed from the unadjusted difference (d = 0.4689).

The gender differences in each specific ability were calculated. Women tended to score higher in facets related to literary knowledge and aesthetic knowledge, while Men scored higher in facets related to computational knowledge and international knowledge. Cultural knowledge showed a small difference in favor of men, though it was negligible in size. The results of this analysis are available in Table 4.

Table 4. Observed gender differences in knowledge by facet of knowledge. * = $p < .05$, ** = $p < .01$, *** = $p < .001$. Positive values indicate advantages for males.

Ability	Gender difference
Computational Knowledge	-1.11***
Technical Knowledge	-0.64***
International Knowledge	-0.73***

Artistic Knowledge	0.65***
Literary Knowledge	0.34***
Cultural Knowledge	-0.14***
General Knowledge	-0.48***

The latent differences were generated by modeling each latent variable as a composite of the observed variables that underlie it using a structural equation model. They are roughly identical to the observed difference in all cases, except for general knowledge, where there is a much larger gender difference (d = -0.70), though this model had an extremely poor fit (CFI = .61), as shown in Table 5.

Table 5. Latent differences in knowledge by sex and facet of knowledge. Reference group is men. COK - Computational knowledge, TK - Technical knowledge, IK - International Knowledge, AK - Aesthetic knowledge, LK - Literary knowledge, CK - Cultural Knowledge, GK - General knowledge. * = $p < .05$, ** = $p < .01$, *** = $p < .001$. Positive values indicate advantages for males. Model fit statistics are from the structural equation models.

Gender	COK	TK	IK	AK	LK	CK	GK
Female	-0.99***	-0.42***	-0.82***	0.80***	0.50***	-0.13***	-0.70***
CFI	.96	.95	.95	.87	.90	.90	.61
RMSEA	0.065	0.080	0.080	0.12	0.14	0.077	0.26
SRMR	0.032	0.037	0.037	0.058	0.049	0.048	0.16

Given that the latent models failed to be useful in determining whether there was a sex difference in the general factor of knowledge, the method of correlated vectors was used to examine the latent sex difference. Analysis was conducted separately within the answers and distractors, and repeated for each number of logistic parameters. To separate the effects of pass rates and g-loadings, they were entered as separate terms in a regression that predicts female advantages on tests. Within the answers, the models with two and three logistic parameters supported a negative relationship between g-loadings and female advantages; the fourth model finds a positive relationship between g-loadings and female advantages. Within the distractors, the method with two logistic parameters found a positive relationship between g-loadings and distractors, but the other two methods found no relationship. Table 6 summarizes the results of this analysis.

Table 6. Correlation between each vector with female advantage and pass rates. Female advantage calculated using the odds-ratio. * = p < .05, ** = p < .01, *** = p < .001. 160 items were used in each regression. Positive parameters indicate that the parameter is associated with higher scores for women.

Method	unstandardized coefficient (g-loadings)	unstandardized coefficient (pass rates)
2PL answers	-1.5 (0.66)*	0.81 (0.39)*
3PL answers	-2.03 (0.45)***	0.34 (0.36)
4PL answers	1.53 (0.48)**	0.88 (0.38)*
2PL distractors	-0.32 (0.29)	0.83 (0.43)
3PL distractors	-0.11 (0.24)	0.6 (0.37)
4PL distractors	0.14 (0.24)	0.49 (1.3)

While non-English speakers had been excluded from this analysis until now, they have been reincluded into the analysis for the sole purpose of assessing national differences. This is because English speakers within foreign countries are not representative of their host nations. Given that the test may be biased, DIF between Anglos and Germans was computed to determine whether this was the case. Anglos scored higher than Germans in general knowledge (d = -0.46, scored with the 160 + 160 scoring method), while the adjusted difference increased to -0.54.

6. Discussion

The analysis suggests the best method to calculate scores for this particular test is to simply add up all of the items, as it is simple and shows high reliability (ω = .93). This is impressive in comparison to the rest of subtests available, for example, the WAIS information subtest has an internal consistency of 0.91 (Weiss et al., 2010). It is worth mentioning that a non-negligible amount of cheating and selective sampling could be occurring, which would lead to an underestimated ceiling.

The sex difference in general knowledge varied by method: extracting scores from the questions with factor analysis results in a difference of -0.38 (favoring men), while extracting scores from the items with IRT that uses four logistic parameters results in a difference of -0.49. The summed scores method, judged to be the best, yields a difference of -0.42.

Testing for bias using DIF found that most items exhibited bias in one direction, but that correcting for this bias hardly changed the sex difference. Men tended to have more technical, computational, and international knowledge, while women had more aesthetic and literary knowledge. This is largely consistent with previous research on sex differences in knowledge (Tran et al., 2014), which suggests that men tend to know more about fields related to science and geography, while women know more about fashion.

One possible reason why there is a gender difference in general knowledge is because there are gender differences in intelligence (Lynn & Kanazawa, 2011; Nyborg, 2005; Hunt, 2010), and general knowledge correlates with intelligence at about .8. While some studies find no sex difference or a sex difference in favor of women, this is an artifact of the fact that the male advantage only emerges after children fully develop (Alexopoulos, 1996; Lynn & Kanazawa, 2011). The likely causal factor behind this difference is brain size, which correlates with intelligence at about .28 (Cox et al., 2019). Given that the sex difference in brain size is about 1 - 1.5 standard deviations (Nyborg, 2005), and that men have larger brains than women even after controlling for differences in height (Ritchie et al., 2018). the predicted male-female standardized difference in intelligence is 0.28 to 0.42 standard deviations, which is higher than the difference found in most psychometric tests, which would be equivalent to a difference of 0.15-0.3 standard deviations.

Beyond this, the analysis suggested that desktop users scored 0.22 standard deviations above mobile phone users in general knowledge. Other studies have found that mobile phone users score 0.58 standard deviations lower than other users in mental ability, which is reduced to 0.25 after selection bias is controlled for using propensity score weighting (Brown et al., 2023). This finding has replicated in other studies as well (Wilmer et al., 2017).

The raw difference in ability between Germans and Anglos was -0.46, which increased to -0.54 after adjusting for bias using DIF. This is not plausible given that the IQ of German speaking countries (99.5) is almost equivalent to those of Anglo ones (98). This is probably because most of the items in the test were Anglo-favored, resulting in the unbiased items being improperly flagged as pro-German. In the case of sex differences, adjusting for DIF bias did not change the difference because the bias of the male and female favored items balances out.

Further investigation should be done regarding whether it's possible to adjust for the cultural bias present in tests of general knowledge. In addition, the fact that the gender difference in general knowledge can vary (d = -.38 to -.49) depending on the way the test is scored is somewhat concerning, and suggests that test scoring methods could be fine-tuned to alter group differences. This analysis also suggests that using DIF is not an optimal method for assessing bias in highly biased tests, and that different methods should be used to assess it. One of them could be the bias/intercept method (Jensen, 1980), where two different regression lines compare the relationship between a predictor (e.g. IQ) and an outcome variable (e.g. income) within two groups (e.g. Whites and Blacks). If the intercepts are significantly different, that means that the predictor variable overpredicts or underpredicts performance on that particular variable, if the slopes are different, that means that the predictor variable is more predictive in one group than the other.

There are also some statistical issues with the modeling practices conducted in this study. Cross-loadings were forcibly removed from the final model, even though the factor analysis suggested that questions tended to load on multiple factors, which can lead to model misspecification (Hsu et al., 2014; Ximénez et al., 2022).

7. Norming

Norms for this test are available in Tables 8 and 9, and details surrounding the methods used to calculate these norms are available in the methodology section.

Table 8. Norms of the MGKT by method used to calculate the norms.

Summed score	Percentile-based IQ	Linear regression-based IQ	z-score based IQ	Averaged estimate
180	56.2	52.1	51.9	53.4
181	56.4	52.9	52.7	54.0
182	56.9	53.6	53.4	54.6
183	58.4	54.4	54.2	55.6
184	58.7	55.1	54.9	56.2
185	58.9	55.9	55.6	56.8
186	59.4	56.6	56.4	57.5
187	59.9	57.4	57.1	58.1
188	60.6	58.1	57.8	58.9
189	61.2	58.9	58.6	59.6
190	62.0	59.6	59.3	60.3
191	62.2	60.4	60.0	60.9
192	62.5	61.1	60.8	61.5
193	63.1	61.9	61.5	62.2
194	63.8	62.6	62.3	62.9
195	64.5	63.4	63.0	63.6
196	65.1	64.1	63.7	64.3

197	65.7	64.9	64.5	65.0
198	66.3	65.6	65.2	65.7
199	66.9	66.4	65.9	66.4
200	67.4	67.1	66.7	67.1
201	67.9	67.9	67.4	67.7
202	68.8	68.6	68.1	68.5
203	69.6	69.4	68.9	69.3
204	70.3	70.1	69.6	70.0
205	70.9	70.9	70.4	70.7
206	71.6	71.6	71.1	71.4
207	72.4	72.4	71.8	72.2
208	73.1	73.1	72.6	72.9
209	73.8	73.9	73.3	73.7
210	74.5	74.7	74.0	74.4
211	75.1	75.4	74.8	75.1
212	75.7	76.2	75.5	75.8
213	76.4	76.9	76.2	76.5
214	77.2	77.7	77.0	77.3
215	78.0	78.4	77.7	78.0
216	78.5	79.2	78.5	78.7
217	79.3	79.9	79.2	79.5
218	79.9	80.7	79.9	80.2
219	80.7	81.4	80.7	80.9
220	81.4	82.2	81.4	81.7
221	82.1	82.9	82.1	82.4
222	82.8	83.7	82.9	83.1
223	83.5	84.4	83.6	83.8
224	84.2	85.2	84.4	84.6

225	84.9	85.9	85.1	85.3
226	85.6	86.7	85.8	86.0
227	86.3	87.4	86.6	86.8
228	87.1	88.2	87.3	87.5
229	87.9	88.9	88.0	88.3
230	88.5	89.7	88.8	89.0
231	89.3	90.4	89.5	89.7
232	90.0	91.2	90.2	90.5
233	90.7	91.9	91.0	91.2
234	91.4	92.7	91.7	91.9
235	92.1	93.4	92.5	92.7
236	92.8	94.2	93.2	93.4
237	93.4	94.9	93.9	94.1
238	94.1	95.7	94.7	94.8
239	94.8	96.4	95.4	95.5
240	95.5	97.2	96.1	96.3
241	96.1	97.9	96.9	97.0
242	96.8	98.7	97.6	97.7
243	97.5	99.4	98.3	98.4
244	98.2	100.2	99.1	99.2
245	98.9	100.9	99.8	99.9
246	99.6	101.7	100.6	100.6
247	100.4	102.4	101.3	101.4
248	101.1	103.2	102.0	102.1
249	101.8	103.9	102.8	102.8
250	102.4	104.7	103.5	103.5
251	103.1	105.5	104.2	104.3
252	103.8	106.2	105.0	105.0

253	104.5	107.0	105.7	105.7
254	105.1	107.7	106.4	106.4
255	106.0	108.5	107.2	107.2
256	106.7	109.2	107.9	107.9
257	107.5	110.0	108.7	108.7
258	108.2	110.7	109.4	109.4
259	109.0	111.5	110.1	110.2
260	109.7	112.2	110.9	110.9
261	110.5	113.0	111.6	111.7
262	111.4	113.7	112.3	112.5
263	112.2	114.5	113.1	113.2
264	112.9	115.2	113.8	114.0
265	113.8	116.0	114.5	114.8
266	114.6	116.7	115.3	115.5
267	115.5	117.5	116.0	116.3
268	116.2	118.2	116.8	117.1
269	116.9	119.0	117.5	117.8
270	117.9	119.7	118.2	118.6
271	118.8	120.5	119.0	119.4
272	119.6	121.2	119.7	120.2
273	120.4	122.0	120.4	120.9
274	121.2	122.7	121.2	121.7
275	122.2	123.5	121.9	122.5
276	123.1	124.2	122.6	123.3
277	124.2	125.0	123.4	124.2
278	125.1	125.7	124.1	125.0
279	126.2	126.5	124.9	125.8
280	127.3	127.2	125.6	126.7

281	128.4	128.0	126.3	127.6
282	129.4	128.7	127.1	128.4
283	130.6	129.5	127.8	129.3
284	131.5	130.2	128.5	130.1
285	132.6	131.0	129.3	131.0
286	133.6	131.7	130.0	131.8
287	134.5	132.5	130.7	132.6
288	135.5	133.2	131.5	133.4
289	136.2	134.0	132.2	134.1
290	137.4	134.8	133.0	135.0
291	138.3	135.5	133.7	135.8
292	139.3	136.3	134.4	136.7
293	140.2	137.0	135.2	137.5
294	141.3	137.8	135.9	138.3
295	142.1	138.5	136.6	139.1
296	143.2	139.3	137.4	139.9
297	144.6	140.0	138.1	140.9
298	145.4	140.8	138.8	141.7
299	145.7	141.5	139.6	142.3
300	145.7	142.3	140.3	142.8
301	145.7	143.0	141.1	143.3
302	145.9	143.8	141.8	143.8
303	146.9	144.5	142.5	144.6
304	146.9	145.3	143.3	145.1
305	147.3	146.0	144.0	145.8
306	147.4	146.8	144.7	146.3
307	147.3	147.5	145.5	146.8
308	147.7	148.3	146.2	147.4

309	149.3	149.0	146.9	148.4
310	149.3	149.8	147.7	148.9

Table 9. Norms of the MGKT by age group.

Score sum	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31-50	51-70
180	66.5	63.6	60.8	58.0	55.4	53.0	50.9	48.9	46.9	45.1	43.6	42.3	41.1	40.2	39.4	38.8	38.3	37.9	34.5	30.8
181	67.4	64.5	61.7	58.9	56.3	53.9	51.7	49.8	47.7	46.0	44.4	43.1	42.0	41.0	40.3	39.6	39.1	38.7	35.3	31.7
182	68.3	65.4	62.6	59.8	57.2	54.8	52.6	50.6	48.6	46.8	45.3	43.9	42.8	41.9	41.1	40.5	40.0	39.6	36.2	32.6
183	69.2	66.3	63.4	60.7	58.0	55.6	53.4	51.5	49.4	47.6	46.1	44.8	43.7	42.7	41.9	41.3	40.8	40.4	37.1	33.5
184	70.1	67.2	64.3	61.5	58.9	56.5	54.3	52.3	50.3	48.5	46.9	45.6	44.5	43.6	42.8	42.2	41.6	41.2	37.9	34.5
185	71.0	68.1	65.2	62.4	59.8	57.3	55.1	53.1	51.1	49.3	47.8	46.5	45.3	44.4	43.6	43.0	42.5	42.1	38.8	35.4
186	71.9	69.0	66.1	63.3	60.6	58.2	55.9	54.0	51.9	50.2	48.6	47.3	46.2	45.2	44.5	43.8	43.3	42.9	39.7	36.3
187	72.8	69.9	67.0	64.2	61.5	59.0	56.8	54.8	52.8	51.0	49.5	48.1	47.0	46.1	45.3	44.7	44.2	43.8	40.5	37.2
188	73.7	70.8	67.9	65.0	62.3	59.9	57.6	55.7	53.6	51.8	50.3	49.0	47.9	46.9	46.1	45.5	45.0	44.6	41.4	38.1
189	74.7	71.7	68.7	65.9	63.2	60.7	58.5	56.5	54.5	52.7	51.1	49.8	48.7	47.8	47.0	46.3	45.8	45.4	42.3	39.0
190	75.6	72.6	69.6	66.8	64.1	61.6	59.3	57.3	55.3	53.5	52.0	50.6	49.5	48.6	47.8	47.2	46.7	46.3	43.2	40.0
191	76.5	73.5	70.5	67.7	64.9	62.4	60.2	58.2	56.1	54.3	52.8	51.5	50.4	49.4	48.7	48.0	47.5	47.1	44.0	40.9
192	77.4	74.3	71.4	68.5	65.8	63.3	61.0	59.0	57.0	55.2	53.6	52.3	51.2	50.3	49.5	48.9	48.4	47.9	44.9	41.8
193	78.3	75.2	72.3	69.4	66.7	64.1	61.9	59.8	57.8	56.0	54.5	53.2	52.0	51.1	50.3	49.7	49.2	48.8	45.8	42.7
194	79.2	76.1	73.2	70.3	67.5	65.0	62.7	60.7	58.6	56.9	55.3	54.0	52.9	51.9	51.2	50.5	50.0	49.6	46.6	43.6
195	80.1	77.0	74.0	71.2	68.4	65.9	63.6	61.5	59.5	57.7	56.2	54.8	53.7	52.8	52.0	51.4	50.9	50.5	47.5	44.6
196	81.0	77.9	74.9	72.0	69.3	66.7	64.4	62.4	60.3	58.5	57.0	55.7	54.6	53.6	52.8	52.2	51.7	51.3	48.4	45.5
197	81.9	78.8	75.8	72.9	70.1	67.6	65.2	63.2	61.2	59.4	57.8	56.5	55.4	54.5	53.7	53.1	52.5	52.1	49.2	46.4
198	82.8	79.7	76.7	73.8	71.0	68.4	66.1	64.0	62.0	60.2	58.7	57.4	56.2	55.3	54.5	53.9	53.4	53.0	50.1	47.3
199	83.7	80.6	77.6	74.6	71.8	69.3	66.9	64.9	62.8	61.1	59.5	58.2	57.1	56.1	55.4	54.7	54.2	53.8	51.0	48.2
200	84.6	81.5	78.5	75.5	72.7	70.1	67.8	65.7	63.7	61.9	60.4	59.0	57.9	57.0	56.2	55.6	55.1	54.7	51.8	49.1
201	85.5	82.4	79.4	76.4	73.6	71.0	68.6	66.6	64.5	62.7	61.2	59.9	58.8	57.8	57.0	56.4	55.9	55.5	52.7	50.1
202	86.4	83.3	80.2	77.3	74.4	71.8	69.5	67.4	65.4	63.6	62.0	60.7	59.6	58.7	57.9	57.2	56.7	56.3	53.6	51.0
203	87.3	84.2	81.1	78.1	75.3	72.7	70.3	68.2	66.2	64.4	62.9	61.6	60.4	59.5	58.7	58.1	57.6	57.2	54.5	51.9
204	88.2	85.1	82.0	79.0	76.2	73.5	71.2	69.1	67.0	65.3	63.7	62.4	61.3	60.3	59.6	58.9	58.4	58.0	55.3	52.8
205	89.1	86.0	82.9	79.9	77.0	74.4	72.0	69.9	67.9	66.1	64.6	63.2	62.1	61.2	60.4	59.8	59.3	58.8	56.2	53.7
206	90.0	86.9	83.8	80.8	77.9	75.3	72.9	70.8	68.7	66.9	65.4	64.1	63.0	62.0	61.2	60.6	60.1	59.7	57.1	54.6
207	90.9	87.8	84.7	81.6	78.8	76.1	73.7	71.6	69.6	67.8	66.2	64.9	63.8	62.9	62.1	61.4	60.9	60.5	57.9	55.6

208	91.8	88.7	85.5	82.5	79.6	77.0	74.6	72.4	70.4	68.6	67.1	65.7	64.6	63.7	62.9	62.3	61.8	61.4	58.8	56.5
209	92.7	89.5	86.4	83.4	80.5	77.8	75.4	73.3	71.2	69.4	67.9	66.6	65.5	64.5	63.8	63.1	62.6	62.2	59.7	57.4
210	93.6	90.4	87.3	84.3	81.4	78.7	76.2	74.1	72.1	70.3	68.7	67.4	66.3	65.4	64.6	64.0	63.4	63.0	60.5	58.3
211	94.5	91.3	88.2	85.1	82.2	79.5	77.1	74.9	72.9	71.1	69.6	68.3	67.1	66.2	65.4	64.8	64.3	63.9	61.4	59.2
212	95.5	92.2	89.1	86.0	83.1	80.4	77.9	75.8	73.7	72.0	70.4	69.1	68.0	67.0	66.3	65.6	65.1	64.7	62.3	60.1
213	96.4	93.1	90.0	86.9	83.9	81.2	78.8	76.6	74.6	72.8	71.3	69.9	68.8	67.9	67.1	66.5	66.0	65.6	63.1	61.1
214	97.3	94.0	90.8	87.7	84.8	82.1	79.6	77.5	75.4	73.6	72.1	70.8	69.7	68.7	67.9	67.3	66.8	66.4	64.0	62.0
215	98.2	94.9	91.7	88.6	85.7	82.9	80.5	78.3	76.3	74.5	72.9	71.6	70.5	69.6	68.8	68.2	67.6	67.2	64.9	62.9
216	99.1	95.8	92.6	89.5	86.5	83.8	81.3	79.1	77.1	75.3	73.8	72.5	71.3	70.4	69.6	69.0	68.5	68.1	65.8	63.8
217	100.0	96.7	93.5	90.4	87.4	84.6	82.2	80.0	77.9	76.2	74.6	73.3	72.2	71.2	70.5	69.8	69.3	68.9	66.6	64.7
218	100.9	97.6	94.4	91.2	88.3	85.5	83.0	80.8	78.8	77.0	75.5	74.1	73.0	72.1	71.3	70.7	70.2	69.8	67.5	65.6
219	101.8	98.5	95.3	92.1	89.1	86.4	83.9	81.7	79.6	77.8	76.3	75.0	73.9	72.9	72.1	71.5	71.0	70.6	68.4	66.6
220	102.7	99.4	96.1	93.0	90.0	87.2	84.7	82.5	80.5	78.7	77.1	75.8	74.7	73.8	73.0	72.3	71.8	71.4	69.2	67.5
221	103.6	100.3	97.0	93.9	90.9	88.1	85.5	83.3	81.3	79.5	78.0	76.7	75.5	74.6	73.8	73.2	72.7	72.3	70.1	68.4
222	104.5	101.2	97.9	94.7	91.7	88.9	86.4	84.2	82.1	80.4	78.8	77.5	76.4	75.4	74.7	74.0	73.5	73.1	71.0	69.3
223	105.4	102.1	98.8	95.6	92.6	89.8	87.2	85.0	83.0	81.2	79.6	78.3	77.2	76.3	75.5	74.9	74.4	73.9	71.8	70.2
224	106.3	103.0	99.7	96.5	93.4	90.6	88.1	85.9	83.8	82.0	80.5	79.2	78.0	77.1	76.3	75.7	75.2	74.8	72.7	71.1
225	107.2	103.8	100.6	97.4	94.3	91.5	88.9	86.7	84.7	82.9	81.3	80.0	78.9	77.9	77.2	76.5	76.0	75.6	73.6	72.1
226	108.1	104.7	101.4	98.2	95.2	92.3	89.8	87.5	85.5	83.7	82.2	80.8	79.7	78.8	78.0	77.4	76.9	76.5	74.4	73.0
227	109.0	105.6	102.3	99.1	96.0	93.2	90.6	88.4	86.3	84.5	83.0	81.7	80.6	79.6	78.9	78.2	77.7	77.3	75.3	73.9
228	109.9	106.5	103.2	100.0	96.9	94.0	91.5	89.2	87.2	85.4	83.8	82.5	81.4	80.5	79.7	79.1	78.5	78.1	76.2	74.8
229	110.8	107.4	104.1	100.9	97.8	94.9	92.3	90.0	88.0	86.2	84.7	83.4	82.2	81.3	80.5	79.9	79.4	79.0	77.1	75.7
230	111.7	108.3	105.0	101.7	98.6	95.8	93.2	90.9	88.8	87.1	85.5	84.2	83.1	82.1	81.4	80.7	80.2	79.8	77.9	76.6
231	112.6	109.2	105.9	102.6	99.5	96.6	94.0	91.7	89.7	87.9	86.4	85.0	83.9	83.0	82.2	81.6	81.1	80.7	78.8	77.6
232	113.5	110.1	106.7	103.5	100.4	97.5	94.9	92.6	90.5	88.7	87.2	85.9	84.8	83.8	83.0	82.4	81.9	81.5	79.7	78.5
233	114.4	111.0	107.6	104.3	101.2	98.3	95.7	93.4	91.4	89.6	88.0	86.7	85.6	84.7	83.9	83.3	82.7	82.3	80.5	79.4
234	115.4	111.9	108.5	105.2	102.1	99.2	96.5	94.2	92.2	90.4	88.9	87.6	86.4	85.5	84.7	84.1	83.6	83.2	81.4	80.3
235	116.3	112.8	109.4	106.1	102.9	100.0	97.4	95.1	93.0	91.3	89.7	88.4	87.3	86.3	85.6	84.9	84.4	84.0	82.3	81.2
236	117.2	113.7	110.3	107.0	103.8	100.9	98.2	95.9	93.9	92.1	90.6	89.2	88.1	87.2	86.4	85.8	85.3	84.9	83.1	82.1
237	118.1	114.6	111.2	107.8	104.7	101.7	99.1	96.8	94.7	92.9	91.4	90.1	89.0	88.0	87.2	86.6	86.1	85.7	84.0	83.1
238	119.0	115.5	112.0	108.7	105.5	102.6	99.9	97.6	95.6	93.8	92.2	90.9	89.8	88.9	88.1	87.4	86.9	86.5	84.9	84.0
239	119.9	116.4	112.9	109.6	106.4	103.4	100.8	98.4	96.4	94.6	93.1	91.7	90.6	89.7	88.9	88.3	87.8	87.4	85.7	84.9
240	120.8	117.3	113.8	110.5	107.3	104.3	101.6	99.3	97.2	95.4	93.9	92.6	91.5	90.5	89.8	89.1	88.6	88.2	86.6	85.8

241	121.7	118.1	114.7	111.3	108.1	105.2	102.5	100.1	98.1	96.3	94.7	93.4	92.3	91.4	90.6	90.0	89.5	89.0	87.5	86.7
242	122.6	119.0	115.6	112.2	109.0	106.0	103.3	101.0	98.9	97.1	95.6	94.3	93.1	92.2	91.4	90.8	90.3	89.9	88.4	87.6
243	123.5	119.9	116.5	113.1	109.9	106.9	104.2	101.8	99.7	98.0	96.4	95.1	94.0	93.0	92.3	91.6	91.1	90.7	89.2	88.6
244	124.4	120.8	117.3	114.0	110.7	107.7	105.0	102.6	100.6	98.8	97.3	95.9	94.8	93.9	93.1	92.5	92.0	91.6	90.1	89.5
245	125.3	121.7	118.2	114.8	111.6	108.6	105.8	103.5	101.4	99.6	98.1	96.8	95.7	94.7	93.9	93.3	92.8	92.4	91.0	90.4
246	126.2	122.6	119.1	115.7	112.4	109.4	106.7	104.3	102.3	100.5	98.9	97.6	96.5	95.6	94.8	94.2	93.6	93.2	91.8	91.3
247	127.1	123.5	120.0	116.6	113.3	110.3	107.5	105.1	103.1	101.3	99.8	98.5	97.3	96.4	95.6	95.0	94.5	94.1	92.7	92.2
248	128.0	124.4	120.9	117.5	114.2	111.1	108.4	106.0	103.9	102.2	100.6	99.3	98.2	97.2	96.5	95.8	95.3	94.9	93.6	93.2
249	128.9	125.3	121.8	118.3	115.0	112.0	109.2	106.8	104.8	103.0	101.5	100.1	99.0	98.1	97.3	96.7	96.2	95.8	94.4	94.1
250	129.8	126.2	122.7	119.2	115.9	112.8	110.1	107.7	105.6	103.8	102.3	101.0	99.9	98.9	98.1	97.5	97.0	96.6	95.3	95.0
251	130.7	127.1	123.5	120.1	116.8	113.7	110.9	108.5	106.5	104.7	103.1	101.8	100.7	99.8	99.0	98.3	97.8	97.4	96.2	95.9
252	131.6	128.0	124.4	120.9	117.6	114.5	111.8	109.3	107.3	105.5	104.0	102.7	101.5	100.6	99.8	99.2	98.7	98.3	97.0	96.8
253	132.5	128.9	125.3	121.8	118.5	115.4	112.6	110.2	108.1	106.4	104.8	103.5	102.4	101.4	100.7	100.0	99.5	99.1	97.9	97.7
254	133.4	129.8	126.2	122.7	119.4	116.3	113.5	111.0	109.0	107.2	105.7	104.3	103.2	102.3	101.5	100.9	100.4	99.9	98.8	98.7
255	134.3	130.7	127.1	123.6	120.2	117.1	114.3	111.9	109.8	108.0	106.5	105.2	104.1	103.1	102.3	101.7	101.2	100.8	99.7	99.6
256	135.2	131.6	128.0	124.4	121.1	118.0	115.1	112.7	110.7	108.9	107.3	106.0	104.9	104.0	103.2	102.5	102.0	101.6	100.5	100.5
257	136.2	132.5	128.8	125.3	122.0	118.8	116.0	113.5	111.5	109.7	108.2	106.8	105.7	104.8	104.0	103.4	102.9	102.5	101.4	101.4
258	137.1	133.3	129.7	126.2	122.8	119.7	116.8	114.4	112.3	110.5	109.0	107.7	106.6	105.6	104.9	104.2	103.7	103.3	102.3	102.3
259	138.0	134.2	130.6	127.1	123.7	120.5	117.7	115.2	113.2	111.4	109.8	108.5	107.4	106.5	105.7	105.1	104.5	104.1	103.1	103.2
260	138.9	135.1	131.5	127.9	124.5	121.4	118.5	116.1	114.0	112.2	110.7	109.4	108.2	107.3	106.5	105.9	105.4	105.0	104.0	104.2
261	139.8	136.0	132.4	128.8	125.4	122.2	119.4	116.9	114.8	113.1	111.5	110.2	109.1	108.1	107.4	106.7	106.2	105.8	104.9	105.1
262	140.7	136.9	133.3	129.7	126.3	123.1	120.2	117.7	115.7	113.9	112.4	111.0	109.9	109.0	108.2	107.6	107.1	106.7	105.7	106.0
263	141.6	137.8	134.1	130.6	127.1	123.9	121.1	118.6	116.5	114.7	113.2	111.9	110.8	109.8	109.0	108.4	107.9	107.5	106.6	106.9
264	142.5	138.7	135.0	131.4	128.0	124.8	121.9	119.4	117.4	115.6	114.0	112.7	111.6	110.7	109.9	109.3	108.7	108.3	107.5	107.8
265	143.4	139.6	135.9	132.3	128.9	125.7	122.8	120.3	118.2	116.4	114.9	113.6	112.4	111.5	110.7	110.1	109.6	109.2	108.3	108.7
266	144.3	140.5	136.8	133.2	129.7	126.5	123.6	121.1	119.0	117.3	115.7	114.4	113.3	112.3	111.6	110.9	110.4	110.0	109.2	109.7
267	145.2	141.4	137.7	134.1	130.6	127.4	124.5	121.9	119.9	118.1	116.6	115.2	114.1	113.2	112.4	111.8	111.3	110.9	110.1	110.6
268	146.1	142.3	138.6	134.9	131.5	128.2	125.3	122.8	120.7	118.9	117.4	116.1	115.0	114.0	113.2	112.6	112.1	111.7	111.0	111.5
269	147.0	143.2	139.4	135.8	132.3	129.1	126.1	123.6	121.6	119.8	118.2	116.9	115.8	114.9	114.1	113.4	112.9	112.5	111.8	112.4
270	147.9	144.1	140.3	136.7	133.2	129.9	127.0	124.4	122.4	120.6	119.1	117.8	116.6	115.7	114.9	114.3	113.8	113.4	112.7	113.3
271	148.8	145.0	141.2	137.5	134.0	130.8	127.8	125.3	123.2	121.5	119.9	118.6	117.5	116.5	115.8	115.1	114.6	114.2	113.6	114.2
272	149.7	145.9	142.1	138.4	134.9	131.6	128.7	126.1	124.1	122.3	120.7	119.4	118.3	117.4	116.6	116.0	115.5	115.0	114.4	115.2
273	150.6	146.8	143.0	139.3	135.8	132.5	129.5	127.0	124.9	123.1	121.6	120.3	119.1	118.2	117.4	116.8	116.3	115.9	115.3	116.1

274	151.5	147.6	143.9	140.2	136.6	133.3	130.4	127.8	125.8	124.0	122.4	121.1	120.0	119.0	118.3	117.6	117.1	116.7	116.2	117.0
275	152.4	148.5	144.7	141.0	137.5	134.2	131.2	128.6	126.6	124.8	123.3	121.9	120.8	119.9	119.1	118.5	118.0	117.6	117.0	117.9
276	153.3	149.4	145.6	141.9	138.4	135.0	132.1	129.5	127.4	125.6	124.1	122.8	121.7	120.7	120.0	119.3	118.8	118.4	117.9	118.8
277	154.2	150.3	146.5	142.8	139.2	135.9	132.9	130.3	128.3	126.5	124.9	123.6	122.5	121.6	120.8	120.2	119.6	119.2	118.8	119.7
278	155.1	151.2	147.4	143.7	140.1	136.8	133.8	131.2	129.1	127.3	125.8	124.5	123.3	122.4	121.6	121.0	120.5	120.1	119.6	120.7
279	156.0	152.1	148.3	144.5	141.0	137.6	134.6	132.0	129.9	128.2	126.6	125.3	124.2	123.2	122.5	121.8	121.3	120.9	120.5	121.6
280	157.0	153.0	149.2	145.4	141.8	138.5	135.4	132.8	130.8	129.0	127.5	126.1	125.0	124.1	123.3	122.7	122.2	121.8	121.4	122.5
281	157.9	153.9	150.0	146.3	142.7	139.3	136.3	133.7	131.6	129.8	128.3	127.0	125.9	124.9	124.1	123.5	123.0	122.6	122.3	123.4
282	158.8	154.8	150.9	147.2	143.5	140.2	137.1	134.5	132.5	130.7	129.1	127.8	126.7	125.8	125.0	124.4	123.8	123.4	123.1	124.3
283	159.7	155.7	151.8	148.0	144.4	141.0	138.0	135.4	133.3	131.5	130.0	128.7	127.5	126.6	125.8	125.2	124.7	124.3	124.0	125.2
284	160.6	156.6	152.7	148.9	145.3	141.9	138.8	136.2	134.1	132.4	130.8	129.5	128.4	127.4	126.7	126.0	125.5	125.1	124.9	126.2
285	161.5	157.5	153.6	149.8	146.1	142.7	139.7	137.0	135.0	133.2	131.7	130.3	129.2	128.3	127.5	126.9	126.4	126.0	125.7	127.1
286	162.4	158.4	154.5	150.6	147.0	143.6	140.5	137.9	135.8	134.0	132.5	131.2	130.1	129.1	128.3	127.7	127.2	126.8	126.6	128.0
287	163.3	159.3	155.3	151.5	147.9	144.4	141.4	138.7	136.7	134.9	133.3	132.0	130.9	130.0	129.2	128.5	128.0	127.6	127.5	128.9
288	164.2	160.2	156.2	152.4	148.7	145.3	142.2	139.5	137.5	135.7	134.2	132.8	131.7	130.8	130.0	129.4	128.9	128.5	128.3	129.8
289	165.1	161.1	157.1	153.3	149.6	146.2	143.1	140.4	138.3	136.5	135.0	133.7	132.6	131.6	130.9	130.2	129.7	129.3	129.2	130.7
290	166.0	161.9	158.0	154.1	150.5	147.0	143.9	141.2	139.2	137.4	135.8	134.5	133.4	132.5	131.7	131.1	130.6	130.1	130.1	131.7
291	166.9	162.8	158.9	155.0	151.3	147.9	144.8	142.1	140.0	138.2	136.7	135.4	134.2	133.3	132.5	131.9	131.4	131.0	130.9	132.6
292	167.8	163.7	159.8	155.9	152.2	148.7	145.6	142.9	140.8	139.1	137.5	136.2	135.1	134.1	133.4	132.7	132.2	131.8	131.8	133.5
293	168.7	164.6	160.6	156.8	153.0	149.6	146.4	143.7	141.7	139.9	138.4	137.0	135.9	135.0	134.2	133.6	133.1	132.7	132.7	134.4
294	169.6	165.5	161.5	157.6	153.9	150.4	147.3	144.6	142.5	140.7	139.2	137.9	136.8	135.8	135.0	134.4	133.9	133.5	133.6	135.3
295	170.5	166.4	162.4	158.5	154.8	151.3	148.1	145.4	143.4	141.6	140.0	138.7	137.6	136.7	135.9	135.3	134.7	134.3	134.4	136.2
296	171.4	167.3	163.3	159.4	155.6	152.1	149.0	146.3	144.2	142.4	140.9	139.6	138.4	137.5	136.7	136.1	135.6	135.2	135.3	137.2
297	172.3	168.2	164.2	160.3	156.5	153.0	149.8	147.1	145.0	143.3	141.7	140.4	139.3	138.3	137.6	136.9	136.4	136.0	136.2	138.1
298	173.2	169.1	165.1	161.1	157.4	153.8	150.7	147.9	145.9	144.1	142.6	141.2	140.1	139.2	138.4	137.8	137.3	136.9	137.0	139.0
299	174.1	170.0	165.9	162.0	158.2	154.7	151.5	148.8	146.7	144.9	143.4	142.1	141.0	140.0	139.2	138.6	138.1	137.7	137.9	139.9
300	175.0	170.9	166.8	162.9	159.1	155.5	152.4	149.6	147.6	145.8	144.2	142.9	141.8	140.9	140.1	139.4	138.9	138.5	138.8	140.8
301	175.9	171.8	167.7	163.8	160.0	156.4	153.2	150.5	148.4	146.6	145.1	143.8	142.6	141.7	140.9	140.3	139.8	139.4	139.6	141.8
302	176.8	172.7	168.6	164.6	160.8	157.3	154.1	151.3	149.2	147.5	145.9	144.6	143.5	142.5	141.8	141.1	140.6	140.2	140.5	142.7
303	177.8	173.6	169.5	165.5	161.7	158.1	154.9	152.1	150.1	148.3	146.8	145.4	144.3	143.4	142.6	142.0	141.5	141.0	141.4	143.6
304	178.7	174.5	170.4	166.4	162.6	159.0	155.7	153.0	150.9	149.1	147.6	146.3	145.2	144.2	143.4	142.8	142.3	141.9	142.2	144.5
305	179.6	175.4	171.3	167.2	163.4	159.8	156.6	153.8	151.8	150.0	148.4	147.1	146.0	145.1	144.3	143.6	143.1	142.7	143.1	145.4
306	180.5	176.3	172.1	168.1	164.3	160.7	157.4	154.6	152.6	150.8	149.3	147.9	146.8	145.9	145.1	144.5	144.0	143.6	144.0	146.3

307	181.4	177.1	173.0	169.0	165.1	161.5	158.3	155.5	153.4	151.6	150.1	148.8	147.7	146.7	146.0	145.3	144.8	144.4	144.9	147.3
308	182.3	178.0	173.9	169.9	166.0	162.4	159.1	156.3	154.3	152.5	150.9	149.6	148.5	147.6	146.8	146.2	145.6	145.2	145.7	148.2
309	183.2	178.9	174.8	170.7	166.9	163.2	160.0	157.2	155.1	153.3	151.8	150.5	149.3	148.4	147.6	147.0	146.5	146.1	146.6	149.1
310	184.1	179.8	175.7	171.6	167.7	164.1	160.8	158.0	155.9	154.2	152.6	151.3	150.2	149.2	148.5	147.8	147.3	146.9	147.5	150.0

IQ scores beyond certain points should not be considered to be reliable. This is because there are too few extreme scorers within a dataset to accurately capture the distribution of ability at these tails, and that these scorers tend to regress to the mean (Lohman & Korb, 2006). It's also unknown whether these norms apply to a representative sample of western countries, as these norms were calculated from an internet sample.

8. Appendix

Table A1. Associated factor by question (CFA model).

Questions	Associated factor
Poets	Literary Knowledge
Musicals	Literary Knowledge
Holidays	Cultural Knowledge
Makeup	Aesthetic Knowledge
Painkillers	Cultural Knowledge
STDs	Cultural Knowledge
Cigarette brands	Cultural Knowledge
Weed slang	Cultural Knowledge
Colonies of france	International Knowledge
Monarchies	International Knowledge
Oil producers	International Knowledge
Nuclear powers	International Knowledge
Video file types	Computational Knowledge
Web browsers	Computational Knowledge
Linux OSs	Computational Knowledge
HTTP status codes	Computational Knowledge
Garments	Aesthetic Knowledge
Craftsman's tools	Technical Knowledge

Red wines	Aesthetic Knowledge
Card games	Cultural Knowledge
Electronic components	Technical Knowledge
Cryptocurrencies	Computational Knowledge
Countries with pyramids	International Knowledge
Famous criminals	Cultural Knowledge
1000 page books	Literary Knowledge
Units of distance	Technical Knowledge
Exercise programs	Aesthetic Knowledge
Synonyms of fancy	Technical Knowledge
Computer cables	Computational Knowledge
Cancers	Cultural Knowledge
Fabric patterns	Aesthetic Knowledge

Table A2. General Knowledge by country (no bias adjustment).

Country	General Knowledge	Sample Size
Austria	95.8	55
Australia	97.2	810
Belgium	93.7	81
Brazil	88.8	176
Canada	100.3	1333
Switzerland	96.5	83
Czechia	94.6	68
Germany	93.9	383
Denmark	92.8	71
Spain	91.8	84
Finland	93.9	123
France	93.5	194
United Kingdom	96.9	1657
Greece	87.6	98
Croatia	92.3	78
Indonesia	84.6	143
Ireland	100.4	137

India	85.7	232
Italy	92.8	137
Japan	97.0	55
Mexico	90.9	83
Malaysia	79.4	139
Netherlands	94.1	221
Norway	94.8	95
New Zealand	98.6	202
Philippines	75.0	236
Poland	91.2	193
Portugal	89.8	66
Romania	86.0	112
Serbia	87.9	78
Russia	90.7	90
Sweden	94.9	203
Singapore	90.0	141
Turkey	85.6	65
United States	100.0	9494
South Africa	93.2	117

Fragment A1. Countries listed by regional category:

- Anglo: US, UK, Canada, New Zealand, Australia, Ireland, South Africa
- Latin American: Mexico, Nicaragua, Panama, Peru, Philippines, Puerto Rico, Paraguay, El Salvador, Uruguay, Argentina, Bolivia, Brazil, Belize, Chile, Columbia, Costa Rica, Cuba, Ecuador, Guatemala, Honduras, Guyana
- German: Germany, Switzerland, Austria
- Northern European: Norway, Sweden, Finland, Belgium, Denmark, Netherlands, Iceland, Luxembourg
- Southern European: Portugal, Spain, France, Andorra, Italy, Greece, Malta
- Eastern European: Estonia, Latvia, Lithuania, Russia, Belarus, Ukraine, Poland, Czechia, Slovakia, Moldova, Hungary, Romania, Slovenia,
- Balkan: Serbia, Macedonia, Albania, Micronesia, Bosnia, Montenegro, Croatia
- Caucasus: Turkey, Georgia, Azerbaijan, Armenia, Kazakhstan, Cyprus

- MENA: Afghanistan, Algeria, Iran, Israel, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Pakistan, Qatar, Saudi Arabia, Tunisia, Egypt
- South Asian: India, Bangladesh, Maldives, Nepal, Bahrain
- East Asian: Hong Kong, Singapore, Japan, China, South Korea, Taiwan, North Korea, Mongolia
- South East Asian: Laos, Malaysia, Thailand, Vietnam, Philippines, Cambodia
- African: Kenya, Sri Lanka, Madagascar, Mauritius, Malawi, Maldives, Nigeria, Mozambique, Seychelles, Sudan, Somalia, South Sudan, Tanizia, Ugandan, Zambia, Zimbabwe, Ethiopia, Ghana, Rwanda
-

Figure A1. Bias in Germans vs Anglos in the distractors.

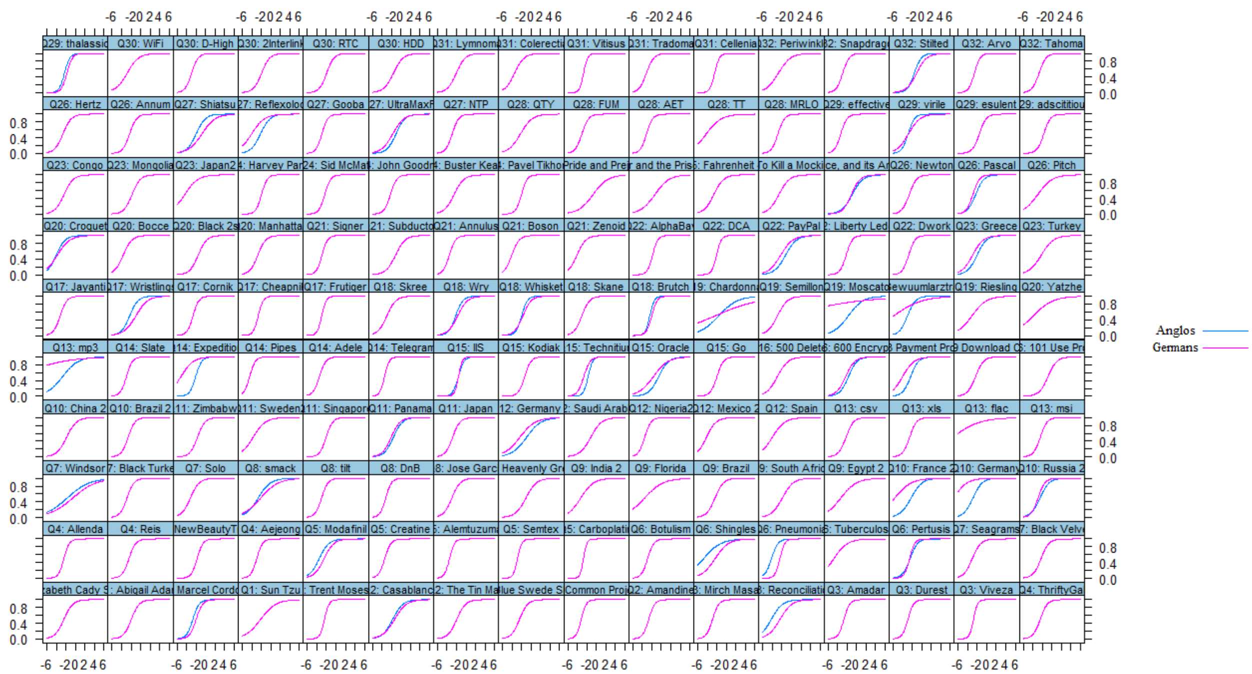


Figure A2. Bias in Germans vs Anglos in the answers.

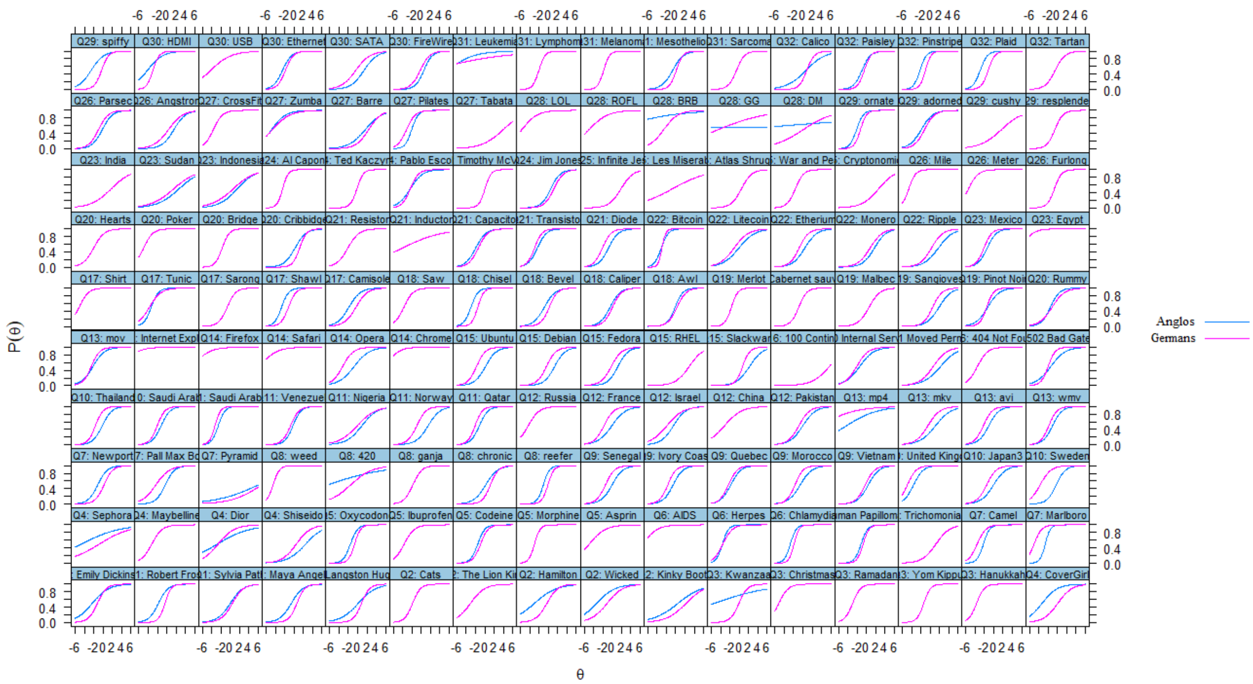


Table A3. g-loadings, pass rates, bias in favor of anglos, and bias in favor of women by item. Bias is reported as an odds-ratio. The reference group in the comparison with Anglo-Saxon countries (US, UK, Australia, New Zealand, South Africa, Canada, Ireland) is countries with similar national IQs to anglo-saxon nations but did not speak English as a first language (Germany, Switzerland, Austria, China, Japan, Hong Kong, Korea, Macao, Liechtenstein, Finland, Estonia, Netherlands, Belarus, Slovenia, Hungary, Belgium, Iceland, Norway, Denmark, Luxembourg, Sweden, France, Russia, Poland, Slovakia, and the Czech Republic).

Question	g-loading	distractor (1) or answer (0)	bias favoring anglos	bias favoring women	pass rate
Q1: Emily Dickinson	0.44	0	1.86	1.71	0.73
Q1: Robert Frost	0.60	0	3.63	1.39	0.72
Q1: Sylvia Path	0.44	0	1.56	2.23	0.47
Q1: Maya Angelou	0.54	0	3.71	2.00	0.52
Q1: Langston Hughes	0.43	0	3.56	1.32	0.40
Q2: Cats	0.64	0	1.24	1.81	0.80

Q2: The Lion King	0.48	0	1.91	1.66	0.81
Q2: Hamilton	0.39	0	2.76	1.49	0.70
Q2: Wicked	0.48	0	2.81	2.31	0.74
Q2: Kinky Boots	0.34	0	2.81	2.39	0.39
Q3: Kwanzaa	0.34	0	2.79	1.07	0.62
Q3: Christmas	0.56	0	0.46	0.77	0.99
Q3: Ramadan	0.53	0	1.20	0.86	0.94
Q3: Yom Kippur	0.70	0	1.90	0.96	0.73
Q3: Hanukkah	0.74	0	1.84	1.35	0.95
Q4: CoverGirl	0.53	0	6.07	3.24	0.76
Q4: Sephora	0.25	0	1.84	3.80	0.73
Q4: Maybelline	0.51	0	1.36	8.79	0.91
Q4: Dior	0.17	0	0.60	2.16	0.70
Q4: Shiseido	0.22	0	0.69	3.65	0.26
Q5: Oxycodone	0.71	0	3.03	0.93	0.74
Q5: Ibuprofen	0.58	0	1.18	1.44	0.92
Q5: Codeine	0.66	0	1.77	0.88	0.76
Q5: Morphine	0.69	0	0.97	0.53	0.97
Q5: Asprin	0.36	0	1.30	1.26	0.92
Q6: AIDS	0.32	0	0.76	1.02	0.99
Q6: Herpes	0.73	0	2.09	0.83	0.97
Q6: Chlamydia	0.75	0	1.52	1.24	0.93
Q6: Human Papillomavirus	0.64	0	2.29	1.39	0.72
Q6: Trichomoniasis	0.39	0	1.84	1.76	0.29
Q7: Camel	0.66	0	0.64	0.87	0.91
Q7: Marlboro	0.67	0	0.61	0.85	0.97
Q7: Newport	0.63	0	4.04	0.86	0.61

Q7: Pall Max Box	0.47	0	0.73	0.83	0.56
Q7: Pyramid	0.22	0	2.47	1.16	0.18
Q8: weed	0.72	0	1.40	0.67	0.99
Q8: 420	0.20	0	1.48	0.93	0.70
Q8: ganja	0.52	0	1.24	0.47	0.81
Q8: chronic	0.53	0	3.21	0.53	0.38
Q8: reefer	0.74	0	3.75	0.62	0.64
Q9: Senegal	0.34	0	0.56	0.74	0.52
Q9: Ivory Coast	0.40	0	0.60	0.71	0.62
Q9: Quebec	0.48	0	0.85	0.70	0.79
Q9: Morocco	0.34	0	0.64	0.81	0.61
Q9: Vietnam	0.40	0	0.90	0.44	0.41
Q10: United Kingdom	0.48	0	0.51	0.52	0.95
Q10: Japan3	0.27	0	0.56	0.54	0.54
Q10: Sweden	0.33	0	0.19	0.57	0.59
Q10: Thailand	0.34	0	0.45	0.57	0.57
Q10: Saudi Arabia	0.39	0	0.66	0.41	0.75
Q11: Saudi Arabia2	0.54	0	0.54	0.27	0.95
Q11: Venezuela	0.46	0	1.08	0.36	0.54
Q11: Nigeria	0.28	0	0.91	0.64	0.37
Q11: Norway	0.22	0	0.23	0.37	0.31
Q11: Qatar	0.40	0	0.56	0.47	0.75
Q12: Russia	0.40	0	0.68	0.23	0.96
Q12: France	0.24	0	0.47	0.23	0.53
Q12: Israel	0.29	0	0.74	0.45	0.59
Q12: China	0.25	0	0.89	0.36	0.84
Q12: Pakistan	0.31	0	0.95	0.46	0.50

Q13: mp4	0.04	0	0.48	0.41	0.81
Q13: mkv	0.14	0	0.52	0.40	0.40
Q13: avi	0.30	0	0.46	0.40	0.73
Q13: wmv	0.39	0	0.80	0.43	0.70
Q13: mov	0.37	0	0.67	0.51	0.83
Q14: Internet Explorer	0.57	0	0.47	0.59	0.99
Q14: Firefox	0.41	0	0.43	0.42	0.99
Q14: Safari	0.48	0	0.74	0.48	0.97
Q14: Opera	0.10	0	0.28	0.32	0.68
Q14: Chrome	0.40	0	0.68	0.87	0.99
Q15: Ubuntu	0.26	0	0.45	0.28	0.52
Q15: Debian	0.28	0	0.56	0.27	0.29
Q15: Fedora	0.30	0	0.66	0.31	0.29
Q15: RHEL	0.33	0	1.19	0.67	0.13
Q15: Slackware	0.34	0	1.08	0.70	0.19
Q16: 100 Continue	0.23	0	0.88	0.62	0.05
Q16: 500 Internal Server Error	0.32	0	0.83	0.75	0.71
Q16: 301 Moved Permanently	0.26	0	0.86	0.72	0.24
Q16: 404 Not Found	0.46	0	0.85	0.49	0.94
Q16: 502 Bad Gateway	0.27	0	0.56	0.44	0.72
Q17: Shirt	0.61	0	1.46	1.29	0.98
Q17: Tunic	0.69	0	1.37	1.15	0.95
Q17: Sarong	0.56	0	1.53	2.29	0.61
Q17: Shawl	0.66	0	4.52	2.04	0.86
Q17: Camisole	0.51	0	2.90	5.52	0.74
Q18: Saw	0.57	0	1.56	0.53	0.94

Q18: Chisel	0.69	0	3.62	0.64	0.91
Q18: Bevel	0.53	0	2.98	0.93	0.58
Q18: Caliper	0.47	0	2.14	0.47	0.53
Q18: Awl	0.60	0	2.55	0.85	0.45
Q19: Merlot	0.69	0	1.25	1.35	0.83
Q19: Cabernet sauvignon	0.56	0	1.20	1.10	0.67
Q19: Malbec	0.47	0	1.47	1.05	0.36
Q19: Sangiovese	0.36	0	0.86	1.05	0.21
Q19: Pinot Noir	0.47	0	0.90	1.12	0.69
Q20: Rummy	0.61	0	2.92	1.05	0.68
Q20: Hearts	0.44	0	1.72	0.91	0.80
Q20: Poker	0.55	0	1.22	0.61	0.99
Q20: Bridge	0.63	0	1.03	0.83	0.83
Q20: Cribbidge	0.55	0	3.90	0.88	0.37
Q21: Resistor	0.35	0	1.03	0.32	0.82
Q21: Inductor	0.14	0	0.93	1.14	0.72
Q21: Capacitor	0.41	0	2.07	0.24	0.80
Q21: Transistor	0.42	0	0.81	0.26	0.89
Q21: Diode	0.34	0	0.56	0.23	0.66
Q22: Bitcoin	0.61	0	0.55	0.18	0.96
Q22: Litecoin	0.22	0	0.76	0.38	0.53
Q22: Etherium	0.28	0	0.68	0.21	0.44
Q22: Monero	0.33	0	0.80	0.50	0.24
Q22: Ripple	0.27	0	0.82	0.36	0.26
Q23: Mexico	0.47	0	0.80	0.46	0.75
Q23: Egypt	0.47	0	0.97	0.63	1.00
Q23: India	0.26	0	0.89	0.81	0.31
Q23: Sudan	0.20	0	0.80	0.81	0.26

Q23: Indonesia	0.26	0	0.88	0.91	0.33
Q24: Al Capone	0.72	0	1.07	0.40	0.93
Q24: Ted Kaczynski	0.64	0	1.37	0.71	0.50
Q24: Pablo Escobar	0.36	0	0.81	0.35	0.83
Q24: Timothy McVeigh	0.69	0	2.14	0.67	0.42
Q24: Jim Jones	0.53	0	2.22	0.84	0.35
Q25: Infinite Jest	0.41	0	1.39	0.96	0.19
Q25: Les Miserables	0.19	0	1.20	1.81	0.54
Q25: Atlas Shrugged	0.47	0	1.47	0.87	0.38
Q25: War and Peace	0.57	0	1.38	1.16	0.78
Q25: Cryptonomicon	0.30	0	1.02	0.86	0.19
Q26: Mile	0.68	0	0.60	0.42	0.99
Q26: Meter	0.54	0	0.44	0.40	0.99
Q26: Furlong	0.58	0	1.58	0.55	0.60
Q26: Parsec	0.41	0	0.76	0.24	0.53
Q26: Angstrom	0.28	0	0.72	0.40	0.27
Q27: CrossFit	0.61	0	1.48	1.01	0.94
Q27: Zumba	0.47	0	2.47	3.08	0.93
Q27: Barre	0.44	0	2.54	3.92	0.25
Q27: Pilates	0.66	0	0.97	1.76	0.88
Q27: Tabata	0.23	0	0.75	1.16	0.16
Q28: LOL	0.44	0	1.06	0.51	0.96
Q28: ROFL	0.39	0	1.12	0.69	0.89
Q28: BRB	0.28	0	1.62	0.94	0.87
Q28: GG	-0.02	0	0.66	0.34	0.56
Q28: DM	0.10	0	1.69	1.10	0.61
Q29: ornate	0.69	0	2.74	0.94	0.85

Q29: adorned	0.47	0	1.99	1.27	0.76
Q29: cushy	0.27	0	1.39	1.02	0.31
Q29: resplendent	0.53	0	1.74	0.82	0.58
Q29: spiffy	0.59	0	3.96	0.91	0.73
Q30: HDMI	0.34	0	0.71	0.45	0.91
Q30: USB	0.25	0	0.98	0.62	0.86
Q30: Ethernet	0.52	0	1.87	0.28	0.86
Q30: SATA	0.23	0	0.59	0.18	0.31
Q30: FireWire	0.44	0	0.86	0.32	0.36
Q31: Leukemia	0.34	0	2.61	1.27	0.91
Q31: Lymphoma	0.71	0	2.31	1.19	0.88
Q31: Melanoma	0.71	0	2.12	1.17	0.85
Q31: Mesothelioma	0.54	0	3.14	0.99	0.54
Q31: Sarcoma	0.56	0	1.66	1.42	0.61
Q32: Calico	0.38	0	2.44	1.48	0.37
Q32: Paisley	0.66	0	2.60	2.31	0.61
Q32: Pinstripe	0.72	0	4.00	1.47	0.83
Q32: Plaid	0.71	0	4.02	2.20	0.88
Q32: Tartan	0.51	0	1.19	1.57	0.55
Q1: Elizabeth Cady Stanton	0.57	1	0.70	1.00	0.90
Q1: Abigail Adams	0.59	1	1.00	1.29	0.86
Q1: Marcel Cordoba	0.66	1	1.52	1.15	0.95
Q1: Sun Tzu	0.35	1	1.27	1.07	0.76
Q1: Trent Moseson	0.77	1	1.33	1.07	0.96
Q2: Casablanca	0.48	1	1.23	1.12	0.77
Q2: The Tin Man	0.53	1	1.01	1.29	0.84

Q2: Blue Swede Shoes	0.58	1	1.04	1.20	0.90
Q2: Common Projects	0.75	1	1.04	1.07	0.99
Q2: Amandine	0.70	1	1.14	1.49	0.95
Q3: Mirch Masala	0.68	1	1.76	0.76	0.96
Q3: Reconciliation	0.46	1	2.44	0.83	0.95
Q3: Amadar	0.74	1	0.72	0.74	0.96
Q3: Durest	0.72	1	0.61	0.64	0.98
Q3: Viveza	0.81	1	0.98	0.74	0.99
Q4: ThriftyGal	0.62	1	0.66	1.45	0.93
Q4: Allenda	0.71	1	1.07	1.68	0.96
Q4: Reis	0.62	1	0.71	1.64	0.94
Q4: NewBeautyTruth	0.69	1	0.72	0.97	0.89
Q4: Aejeong	0.69	1	1.14	1.23	0.97
Q5: Modafinil	0.55	1	1.85	1.19	0.94
Q5: Creatine	0.69	1	0.80	0.68	0.98
Q5: Alemtuzumab	0.73	1	1.41	1.14	0.99
Q5: Semtex	0.77	1	1.38	0.98	0.99
Q5: Carboplatin	0.85	1	0.96	0.59	0.99
Q6: Botulism	0.68	1	0.74	0.88	0.97
Q6: Shingles	0.34	1	1.77	1.20	0.90
Q6: Pneumonia	0.68	1	2.16	1.03	0.99
Q6: Tuberculosis	0.40	1	1.24	1.07	0.95
Q6: Pertusis	0.62	1	0.68	1.32	0.93
Q7: Seagrams	0.59	1	0.58	0.64	0.95
Q7: Black Velvet	0.58	1	1.26	0.86	0.89
Q7: Windsor	0.26	1	1.33	1.15	0.66

Q7: Black Turkey	0.64	1	0.75	0.70	0.95
Q7: Solo	0.56	1	0.88	0.79	0.95
Q8: smack	0.42	1	1.43	0.72	0.84
Q8: tilt	0.67	1	1.47	0.63	0.97
Q8: DnB	0.72	1	0.95	0.53	0.98
Q8: Jose Garcia	0.56	1	1.36	0.64	0.94
Q8: Heavenly Green	0.48	1	1.14	0.75	0.68
Q9: India 2	0.39	1	0.65	0.79	0.90
Q9: Florida	0.27	1	0.74	1.18	0.86
Q9: Brazil	0.54	1	0.47	0.58	0.91
Q9: South Africa	0.46	1	0.73	0.56	0.81
Q9: Egypt 2	0.32	1	1.13	1.20	0.86
Q10: France 2	0.44	1	0.36	0.61	0.86
Q10: Germany	0.51	1	0.34	0.52	0.94
Q10: Russia 2	0.52	1	0.27	0.46	0.91
Q10: China 2	0.47	1	0.57	0.52	0.83
Q10: Brazil 2	0.60	1	0.68	0.71	0.95
Q11: Zimbabwe	0.59	1	0.55	0.66	0.85
Q11: Sweden2	0.42	1	1.09	1.04	0.93
Q11: Singapore	0.63	1	0.81	0.63	0.91
Q11: Panama	0.52	1	0.72	0.63	0.80
Q11: Japan	0.59	1	0.59	0.48	0.95
Q12: Germany 2	0.35	1	0.64	0.85	0.65
Q12: Saudi Arabia 3	0.44	1	0.71	0.36	0.72
Q12: Nigeria2	0.66	1	0.87	0.28	0.99
Q12: Mexico 2	0.51	1	0.71	0.57	0.98
Q12: Spain	0.41	1	0.71	1.01	0.95

Q13: csv	0.59	1	0.62	0.53	0.91
Q13: xls	0.60	1	0.59	0.56	0.95
Q13: flac	0.20	1	1.07	1.34	0.92
Q13: msi	0.60	1	0.66	0.75	0.94
Q13: mp3	0.30	1	0.47	0.60	0.85
Q14: Slate	0.65	1	0.24	0.54	0.98
Q14: Expedition	0.64	1	0.23	0.39	0.94
Q14: Pipes	0.69	1	1.10	0.81	1.00
Q14: Adele	0.63	1	0.36	0.56	0.99
Q14: Telegram	0.78	1	0.93	0.44	0.99
Q15: IIS	0.70	1	0.55	0.67	0.88
Q15: Kodiak	0.66	1	0.68	0.74	0.82
Q15: Technitium	0.76	1	0.47	0.57	0.91
Q15: Oracle	0.48	1	0.64	0.70	0.62
Q15: Go	0.65	1	0.81	0.97	0.90
Q16: 500 Deleted	0.44	1	0.89	1.03	0.91
Q16: 600 Encrypted	0.53	1	0.66	0.73	0.83
Q16: 303 Payment Processing	0.57	1	0.65	0.71	0.91
Q16: 209 Download Complete	0.57	1	0.68	0.56	0.92
Q16: 101 Use Proxy	0.53	1	0.86	0.67	0.67
Q17: Jayanti	0.67	1	1.22	1.55	0.97
Q17: Wristlings	0.53	1	1.72	1.71	0.88
Q17: Cornik	0.69	1	1.15	1.91	0.96
Q17: Cheapnik	0.71	1	1.38	1.33	0.97
Q17: Frutiger	0.77	1	0.96	1.82	0.99
Q18: Skree	0.56	1	1.20	0.69	0.93

Q18: Wry	0.56	1	1.86	0.74	0.89
Q18: Whisket	0.66	1	1.47	0.68	0.89
Q18: Skane	0.62	1	1.20	0.69	0.95
Q18: Brutch	0.74	1	1.63	0.67	0.95
Q19: Chardonnay	0.28	1	1.15	1.14	0.66
Q19: Semillon	0.50	1	1.20	0.93	0.92
Q19: Moscato	0.35	1	0.58	0.88	0.75
Q19: Gewuumlarztraminer	0.54	1	1.36	0.81	0.94
Q19: Riesling	0.36	1	1.02	0.88	0.87
Q20: Yatzhe	0.27	1	0.98	0.90	0.86
Q20: Croquet	0.55	1	1.40	0.92	0.97
Q20: Bocce	0.53	1	1.29	0.88	0.97
Q20: Black 2s	0.57	1	0.94	0.70	0.79
Q20: Manhattan	0.56	1	1.85	1.24	0.93
Q21: Signer	0.68	1	1.13	0.44	0.98
Q21: Subductor	0.48	1	0.64	0.37	0.73
Q21: Annulus	0.67	1	0.44	0.54	0.99
Q21: Boson	0.58	1	0.91	0.48	0.99
Q21: Zenoid	0.51	1	0.98	0.79	0.98
Q22: AlphaBay	0.75	1	0.81	0.59	0.91
Q22: DCA	0.76	1	0.77	0.63	0.94
Q22: PayPal	0.43	1	0.47	0.30	0.73
Q22: Liberty Ledger	0.66	1	0.65	0.64	0.93
Q22: Dwork	0.66	1	0.84	0.78	0.96
Q23: Greece	0.49	1	0.60	0.59	0.85
Q23: Turkey	0.42	1	0.73	0.84	0.83
Q23: Congo	0.51	1	0.79	0.86	0.89

Q23: Mongolia	0.55	1	0.70	0.78	0.84
Q23: Japan2	0.40	1	0.77	0.97	0.95
Q24: Harvey Parnell	0.69	1	0.76	0.66	0.91
Q24: Sid McMath	0.69	1	0.84	0.54	0.94
Q24: John Goodman	0.66	1	0.96	0.67	0.92
Q24: Buster Keaton	0.52	1	0.93	0.82	0.89
Q24: Pavel Tikhonov	0.70	1	0.89	0.80	0.92
Q25: Pride and Prejudice	0.34	1	0.80	1.13	0.61
Q25: Harry Potter and the Prisoner of Azkaban	0.36	1	0.70	0.65	0.70
Q25: Fahrenheit 451	0.50	1	1.28	0.83	0.89
Q25: To Kill a Mockingbird	0.47	1	1.01	0.87	0.88
Q25: Science, and its Antecedents	0.47	1	0.76	0.85	0.67
Q26: Newton	0.58	1	0.45	0.39	0.95
Q26: Pascal	0.55	1	0.45	0.67	0.92
Q26: Pitch	0.36	1	1.40	0.88	0.84
Q26: Hertz	0.53	1	0.58	0.33	0.93
Q26: Annum	0.61	1	1.07	0.82	0.92
Q27: Shiatsu	0.48	1	1.72	0.77	0.83
Q27: Reflexology	0.50	1	0.44	1.01	0.87
Q27: Gooba	0.71	1	1.17	1.52	0.97
Q27: UltraMaxFit	0.52	1	0.67	1.05	0.71
Q27: NTP	0.65	1	1.25	1.59	0.96
Q28: QTY	0.42	1	0.86	0.62	0.84
Q28: FUM	0.64	1	0.83	0.71	0.97
Q28: AET	0.68	1	0.67	0.55	0.98

Q28: TT	0.36	1	0.99	0.66	0.92
Q28: MRLO	0.67	1	0.77	0.74	0.97
Q29: effective	0.60	1	1.50	0.67	0.98
Q29: virile	0.62	1	1.86	1.14	0.96
Q29: esulent	0.61	1	0.81	0.93	0.83
Q29: adscititious	0.64	1	0.98	0.81	0.85
Q29: thalassic	0.69	1	1.58	0.92	0.94
Q30: WiFi	0.45	1	0.53	0.30	0.90
Q30: D-High	0.65	1	0.94	0.72	0.96
Q30: 2Interlink	0.56	1	0.76	0.61	0.86
Q30: RTC	0.60	1	0.79	0.74	0.88
Q30: HDD	0.48	1	0.98	0.74	0.78
Q31: Lymnoma	0.53	1	1.15	1.05	0.85
Q31: Colerectia	0.53	1	1.25	1.03	0.95
Q31: Vitisus	0.80	1	1.13	0.84	0.98
Q31: Tradoma	0.63	1	1.12	0.82	0.96
Q31: Cellenia	0.75	1	1.23	0.91	0.97
Q32: Periwinkle	0.45	1	0.98	1.65	0.87
Q32: Snapdragon	0.66	1	1.01	0.95	0.96
Q32: Stilted	0.59	1	1.37	1.20	0.80
Q32: Arvo	0.70	1	0.89	0.95	0.96
Q32: Tahoma	0.61	1	1.38	1.22	0.94

Table A4. Observed differences by specific ability by region. Reference group is anglos. COK - Computational knowledge, TK - Technical knowledge, IK - International Knowledge, AK - Aesthetic knowledge, CK - Cultural Knowledge, LK - Literary knowledge, GK - General knowledge.

Cultural category	COK	TK	IK	AK	LK	CK	GK
German (n = 521)	0.40***	-0.57***	0.62***	-0.75***	-0.72***	-0.87***	-0.35***
Latin American (n = 628)	-0.03	-0.84***	-0.2***	-1.04***	-0.83***	-1.32***	-1.05***
Northern European (n = 814)	0.60***	-0.56***	0.52***	-0.59***	-0.8***	-0.81***	-0.36***
Southern European (n = 587)	0.06	-0.68***	0.44***	-0.64***	-0.66***	-0.97***	-0.54***
Eastern European (n = 742)	0.50***	-0.76***	0.3***	-0.89***	-0.96***	-1.04***	-0.64***
Balkans (n = 199)	0.34	-0.97***	0.2**	-0.87***	-0.98***	-0.88***	-0.73***
Caucasus (n = 84)	-0.27	-0.98***	0.19	-1.37***	-1.02***	-1.57***	-1.05***
MENA (n = 138)	-0.22	-0.92***	0.17*	-1.25***	-1.07***	-1.4***	-1.1***
South Asian (n = 256)	0.14	-0.36***	-0.12	-1.24***	-0.98***	-1.63***	-0.96***
East Asian (n = 354)	0.2**	-0.33***	0.28***	-0.38***	-0.59***	-1.18***	-0.47***
South East Asian (n = 413)	0.22	-0.97***	-0.57***	-1.15***	-1.19***	-1.9***	-1.52***
African (n = 66)	-0.2*	-0.37**	0.37**	-0.81***	-1.1***	-1.09***	-0.71***

References

Alexopoulos, D. S. (1996). Sex differences and I.Q. *Personality and Individual Differences*, *20*(4), 445–450. https://doi.org/10.1016/0191-8869(95)00187-5

Breuer, S., Scherndl, T., & Ortner, T. M. (2020). Effects of response format on psychometric properties and fairness of a matrices test: Multiple choice vs. free response. *Frontiers in Education*, *5*. https://doi.org/10.3389/feduc.2020.00015

Brown, M. I., Grossenbacher, M. A., & Warman, Z. (2023). Self-selection as an explanation for general mental ability test score differences between mobile and nonmobile devices in observational studies. *Journal of Applied Psychology*, *108*(7), 1190–1206. https://doi.org/10.1037/apl0001067

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for theREnvironment. *Journal of Statistical Software*, *48*(6). https://doi.org/10.18637/jss.v048.i06

Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, *12*(4), 381–398. https://doi.org/10.1037/1082-989x.12.4.381

Cox, S. R., Ritchie, S. J., Fawns-Ritchie, C., Tucker-Drob, E. M., & Deary, I. J. (2019). Structural brain imaging correlates of general intelligence in UK Biobank. *Intelligence*, *76*, 101376. https://doi.org/10.1016/j.intell.2019.101376

Hsu, H.-Y., Troncoso Skidmore, S., Li, Y., & Thompson, B. (2014). Forced zero cross-loading misspecifications in measurement component of structural equation models. *Methodology*, *10*(4), 138–152. https://doi.org/10.1027/1614-2241/a000084

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

Hunt, E. (2010). *Human intelligence*. Cambridge University Press.

Jensen, A. R. (1980). *Bias in mental testing*. Free Press.

Johnson, W., Bouchard, T. J., Jr., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, *32*(1), 95–107. https://doi.org/10.1016/s0160-2896(03)00062-x

Logos, K., Brewer, N., & Young, R. L. (2021). *Convergent validity of a quick online self-administered measure of verbal IQ for psychology researchers*. Center for Open Science. http://dx.doi.org/10.31234/osf.io/7csvm

Lohman, D. F., & Korb, K. A. (2006). Gifted today but not tomorrow? Longitudinal changes in ability and achievement during elementary school. *Journal for the Education of the Gifted*, *29*(4), 451–484. https://doi.org/10.4219/jeg-2006-245

Lynn, R., & Becker, D. (2019). *Intelligence of nations.* Ulster Institute.

Lynn, R., & Kanazawa, S. (2011). A longitudinal study of sex differences in intelligence at ages 7, 11 and 16 years. *Personality and Individual Differences*, *51*(3), 321–324. https://doi.org/10.1016/j.paid.2011.02.028

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology*, *38*(1), 115–142.

https://doi.org/10.1037/0012-1649.38.1.115

Nyborg, H. (2005). Sex-related differences in general intelligence g, brain size, and social status. *Personality and Individual Differences*, *39*(3), 497–509. https://doi.org/10.1016/j.paid.2004.12.011

Palmer, P., Haywood, C. S., Fairbank, B. A., & Earles, J. A. (1990). *Comparison of the armed services vocational aptitude battery to the general aptitude test battery*. Defense Technical Information Center. http://dx.doi.org/10.21236/ada221551

Polat, M. (2020). Analysis of Multiple-choice versus Open-ended Questions in Language Tests According to Different Cognitive Domain Levels. *Novitas-Royal (Research on Youth and Language)*, *14*(2), 76–96.

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., Harris, M. A., Alderson, H. L., Hunter, S., Neilson, E., Liewald, D. C. M., Auyeung, B., Whalley, H. C., Lawrie, S. M., Gale, C. R., Bastin, M. E., McIntosh, A. M., & Deary, I. J. (2018). Sex differences in the adult human brain: Evidence from 5216 UK biobank participants. *Cerebral Cortex*, *28*(8), 2959–2975. https://doi.org/10.1093/cercor/bhy109

Rohwedder, S., & Willis, R. J. (2010). Mental retirement. *Journal of Economic Perspectives*, *24*(1), 119–138. https://doi.org/10.1257/jep.24.1.119

Roid, G., & Pomplun, M. (2011). Stanford-Binet Intelligence scales, fifth edition. In *Encyclopedia of Child Behavior and Development* (pp. 1439–1439). Springer US. http://dx.doi.org/10.1007/978-0-387-79061-9_6176

Sirota, M., & Juanchich, M. (2018). Effect of response format on cognitive reflection: Validating a two- and four-option multiple choice question version of the Cognitive Reflection Test. *Behavior Research Methods*, *50*(6), 2511–2522. https://doi.org/10.3758/s13428-018-1029-4

Sivo, S. A., Fan, X., Witta, E. L., & Willse, J. T. (2006). The search for "optimal" cutoff properties: Fit index criteria in structural equation modeling. *The Journal of Experimental Education*, *74*(3), 267–288. https://doi.org/10.3200/jexe.74.3.267-288

Tran, U. S., Hofer, A. A., & Voracek, M. (2014). Sex differences in general knowledge: Meta-Analysis and new data on the contribution of school-related moderators among high-school students. *PLoS ONE*, *9*(10), e110391. https://doi.org/10.1371/journal.pone.0110391

Voronin, I., Te Nijenhuis, J., & Malykh, S. B. (2015). THE CORRELATION BETWEEN g LOADINGS AND HERITABILITY IN RUSSIA. *Journal of Biosocial Science*, *48*(6), 833–843. https://doi.org/10.1017/s0021932015000395

Weiss, L. G., Saklofske, D. H., Coalson, D., & Raiford, S. E. (2010). *WAIS-IV clinical use* *and interpretation: Scientist-Practitioner perspectives*. Academic Press.

Wilmer, H. H., Sherman, L. E., & Chein, J. M. (2017). Smartphones and cognition: A review of research exploring the links between mobile technology habits and cognitive functioning. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00605

Ximénez, C., Revuelta, J., & Castañeda, R. (2022). What are the consequences of ignoring cross-loadings in bifactor models? A simulation study assessing parameter recovery and sensitivity of goodness-of-fit indices. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.923877

Young, S. R., & Keith, T. Z. (2020). An examination of the convergent validity of the ICAR16 and WAIS-IV. *Journal of Psychoeducational Assessment*, *38*(8), 1052–1059. https://doi.org/10.1177/0734282920943455