Intelligence 53 (2015) 118–137

Contents lists available at ScienceDirect

Intelligence

Intelligence and school grades: A meta-analysis

Bettina Roth *,[1], Nicolas Becker [1], Sara Romeyke, Sarah Schäfer, Florian Domnick, Frank M. Spinath

Saarland University, Campus A1 3, 66123 Saarbrücken, Germany

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form 7 September 2015
Accepted 7 September 2015
Available online xxxx

Keywords:
Intelligence
School grades
School achievement
Meta-analysis

ABSTRACT

Intelligence is considered as the strongest predictor of scholastic achievement. Research as well as educational policy and the society as a whole are deeply interested in its role as a prerequisite for scholastic success. The present study investigated the population correlation between standardized intelligence tests and school grades employing psychometric meta-analysis (Hunter & Schmidt, 2004). The analyses involved 240 independent samples with 105,185 participants overall. After correcting for sampling error, error of measurement, and range restriction in the independent variable, we found a population correlation of $\rho = .54$. Moderator analyses pointed to a variation of the relationship between *g* and school grades depending on different school subject domains, grade levels, the type of intelligence test used in the primary study, as well as the year of publication, whereas gender had no effect on the magnitude of the relationship.

1. Introduction

Intelligence is the strongest predictor of academic achievement with correlations ranging from .30 to .70 (e.g., Chamorro-Premuzic & Furnham, 2005; Colom & Flores-Mendoza, 2007; Deary, Strand, Smith, & Fernandes, 2007; Gottfredson, 2002; Gustafsson & Undheim, 1996; Jensen, 1998; Kuncel, Hezlett, & Ones, 2004; Kyttälä & Lehto, 2008; Laidra, Pillmann, & Allik, 2007; Lemos, Abad, Almeida, & Colom, 2014; Neisser et al., 1996; Primi, Ferrão, & Almeida, 2010; Rosander, Bäckström, & Stenberg, 2011; Taub, Keith, Floyd, & Mcgrew, 2008). Well known and much-quoted reviews (e.g., Gottfredson, 2002; Gustafsson & Undheim, 1996; Neisser et al., 1996; Sternberg, Grigorenko, & Bundy, 2001) refer to a mean correlation of .5, but none of them cites a study in which this was investigated. This is not surprising, since there was (and is) no current comprehensive meta-analytic examination of the association between *g* and scholastic achievement. Previous meta-analyses (see the following section) present data assessed before 1983, focus only on natural sciences and specific countries and do not correct for artifacts which might lower the correlation (i.e., unreliability, range restriction). Moreover, scholastic achievement is measured by achievement tests instead of school grades as a direct measure of scholastic success. However, school grades are crucial for accessing further scholastic and occupational qualification, and therefore, have an enormous influence on an individual's life (Sauer, 2006; Tent, 2006). With this study we try to close this research gap by integrating the extensive body of knowledge concerning the correlation between *g* and scholastic achievement measured by school grades. The main goals of the study were the following: (1) Consideration of all available studies presented in the international literature, (2) presentation of the mean correlation only weighted by sample size as well as the true score correlation corrected for unreliability and range restriction, (3) consideration of moderator variables which might influence the correlation.

1.1. Results of previous meta-analyses

1.1.1. Boulanger (1981)

This study dealt with the correlation of cognitive ability assessed by different standardized intelligence tests and school achievement in natural sciences in grade levels 6 to 12. Included were 34 studies between 1963 and 1978 yielding 62 correlations (total *N* not reported). The correlations were integrated by computing the mean correlation and corresponding standard deviations. For the complete sample, a mean correlation of $M(r) = .48$ with a standard deviation of $SD(\mathrm{r}) = .15$ was found. Furthermore, the mean correlations on different levels of a set of potential moderator variables were computed and compared using a *t*-test. Among several tested moderator variables, only the reliability of the outcome measure [$r < .80$: $M(r) = .42$ vs. $r \geq .80$: $M(r) = .55$; $p = .01$] had a significant influence on the strength of the relationship between cognitive ability and school achievement.

1.1.2. Fleming and Malone (1983)

The meta-analysis of Fleming and Malone (1983) analyzed correlations of different student variables (among others general ability, verbal and mathematical ability) and scholastic achievement in natural sciences. It was based on 42 correlation coefficients (number of studies

* Corresponding author at: Department of Psychology, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany.
E-mail address: bettina.roth@uni-saarland.de (B. Roth).

[1] The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

http://dx.doi.org/10.1016/j.intell.2015.09.002
0160-2896/

and total N not reported) between 1960 and 1981. Grade levels ranged from kindergarten to grade level 12. Intelligence was assessed by verbal and mathematical scholastic aptitude tests (SAT), scholastic achievement by standardized tests. The meta-analysis was based on the strategy of Glass, McGaw, and Smith (1981). For all studies, the mean true effect was $\rho = .43$ with a standard deviation of $\sigma_\rho = .22$. Analyses of moderator variables revealed a partially moderating effect of different grade levels (Elementary School: $\rho = .25$; $\sigma_\rho = .20$; Middle School: $\rho = .59$; $\sigma_\rho = .12$; High School: $\rho = .47$; $\sigma_\rho = .36$).

1.1.3. Steinkamp and Maehr (1983)

This meta-analysis integrated correlations between affect, cognitive ability and scholastic achievement in natural sciences. Since a central goal of this study was to analyze gender effects, only studies reporting gender-specific correlations were considered. For cognitive ability, 60 coefficients between 1965 and 1983 were found (number of studies and total N not reported), which were based exclusively on anglophone individuals. Grade levels ranged from elementary school to high school. Cognitive ability was assessed by standardized intelligence tests, scholastic ability by standardized and unstandardized tests. The authors employed the meta-analytic strategy of Glass (1977). For all studies, the mean true effect was $\rho = .34$ (σ_ρ not reported), with no significant effect for gender.

1.1.4. Summary of previous results

Ranging between $\rho = .34$ and .48, the mean correlation between cognitive abilities and scholastic achievement investigated in previous meta-analyses was slightly lower than generally assumed in the literature (e.g., Neisser et al., 1996). A wide range of possible moderator variables was analyzed, with significant effects only for grade level (Fleming & Malone, 1983) and the reliability of the outcome measure (Boulanger, 1981). As there was a strong focus on achievement in scientific school subjects, the bulk of primary studies addressing the impact of cognitive abilities for school achievement in other subject domains was not considered, nor were potential differences in the mean correlation between g and school grades across these subject domains analyzed.

1.2. The present study

Our goal was to identify the empirical estimate of the population correlation between g and scholastic success. We argue that school grades have a much stronger effect on an individual's subsequent school and occupational career than alternative measures of school achievement (e.g., teacher ratings, school achievement tests). Therefore we focus on scholastic success in a strict sense which means that we use school grades as a criterion exclusively.

The current study aims at investigating the population correlation between g and school grades in general and without restrictions on a specific subject domain or grade level as well as the country where the data were collected and the year the study was published. Moreover, we illustrate the moderating effect of third variables on the relationship between g and school grades.

1.3. Moderator hypotheses

To analyze moderating effects, we formulated hypotheses about potential moderating variables. We derived our assumptions from the previous meta-analyses by Boulanger (1981), Fleming and Malone (1983), and Steinkamp and Maehr (1983) as well as from the general literature on the topic. Thus, we identified five potential moderators which are presented below (details on the coding process for moderator variables will be presented in the method section).

1.3.1. Type of intelligence test

According to Gaedike (1974) and Sauer (2006), the performance in verbal intelligence tests is related more strongly to scholastic success than is the achievement in nonverbal ones. To test the moderating effect of the verbal or nonverbal character of intelligence tests, we built subgroups for either completely verbal or nonverbal intelligence tests as well as for such measurement instruments consisting of both verbal and nonverbal scales.

1.3.2. Subject domains

Previous meta-analyses (Boulanger, 1981; Fleming & Malone, 1983; Steinkamp & Maehr, 1983) concentrated on the mean correlation between g and school achievement in scientific school subjects. There are other school subjects beyond mathematics and science, which have not been considered in meta-analyses before. Hence we aimed at covering these and estimating the population correlation between g and school grades in a range of different school subjects. To reduce complexity and to reach a clear overview, we clustered the school subjects considered in the included primary studies into the following subgroups: Mathematics and Science (including e.g., mathematics, biology, and physics), Languages (including e.g., English, German, reading, and literature), Social Sciences (including e.g., social studies, history, and geography), Fine Art and Music, as well as Sports.

1.3.3. Grade level

The moderating effect of grade level on the correlation between general mental ability and school grades was analyzed by Boulanger (1981), Fleming and Malone (1983), and Steinkamp and Maehr (1983). Apart from that, Brody (1992) and Jensen (1998) point out the variation of the predictive value of g for scholastic success against the background of different grade levels. Jensen (1998) refers to correlations between g and grades which decrease from elementary school (.60 to .70) throughout high school (.50 to .60), college (.40 to .50) and graduate school (.30 to .40). As a consequence of the increasing drop-out of individuals with lower abilities during secondary school they expect a reduction in variance in g and hence a lower correlation between g and scholastic achievement in higher grade levels. In order to investigate the influence of grade level and to make the results of our analysis comparable to the previous meta-analytic findings, we clustered grade levels into the subgroups Elementary School, Middle School, and High School.

1.3.4. Gender

The meta-analysis by Steinkamp and Maehr (1983) did not reveal a significant difference between boys and girls in the correlation between g and scholastic achievement. In the current study, we tested gender as a variable moderating the relationship between g and school grades. We based the analysis on those samples that consisted of either male or female participants.

1.3.5. Year of publication

In order to investigate a potential change in the population correlation between g and scholastic achievement since the previous meta-analyses by Boulanger (1981), Fleming and Malone (1983) and Steinkamp and Maehr (1983), we separated the primary studies into two subgroups including primary studies published before 1983 and those published afterwards.

2. Method

2.1. Inclusion criteria

In this meta-analysis we considered primary studies that fulfilled the following inclusion criteria: (1) The independent variable general mental ability was measured either by standardized intelligence tests or highly comparable tests [e.g., Differential Aptitude Tests (Bennett, Seashore, & Wesman, 1947), Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy, & Kirk, 1974)]. We included primary studies with author-created measures, if it was possible to clearly classify them as intelligence tests. Primary studies that used achievement tests were not included because of the insufficient specificity of these measures. (2) The dependent variable school achievement was measured either by grade point average (GPA) or by other criteria of an individual's achievement in the narrow sense of school grades, for example, achievement expressed as a percentage in the school report. Articles using alternative measures of scholastic achievement such as teacher ratings or school achievement tests were not included. (3) School grades were derived from primary or secondary education. (4) The primary study reported zero-order correlations between the independent and the dependent variable or a coefficient that allowed us to calculate a zero-order correlation. (5) The sample size of the primary study was reported. (6) The primary study was available in German or English language. (7) The study was free from methodological flaws.

2.2. Literature search

We used two strategies to identify studies for the present meta-analysis: (1) We conducted a broad literature search using the databases PsycARTICLES, PsycINFO, PSYNDEXplus, ERIC, Science Direct, and Google Scholar. Search terms for intelligence were *intelligence*, *cognitive ability*, *mental ability*, *ability*, *g*, *g factor*, *general mental ability*, and *GMA*. Search terms for school grades were *school grades*, *school marks*, *scholastic achievement*, and *school achievement*. The search terms were combined in all 32 (8×4) possible ways. Entering these combinations in the six databases resulted in 192 (32×6) queries. The literature search covered all articles published before April 2014. We found 320 different studies that appeared to be relevant according to the title and abstract. (2) To reduce publication bias we gathered "gray literature" by contacting the mailing list of the Deutsche Gesellschaft für Psychologie, a German association of psychologists working in science and education. We asked the members to send us information about unpublished studies yielding correlations between intelligence and school grades. We received 19 additional appropriate datasets, including 12 unpublished studies. In total, the different search strategies yielded 335 primary studies.

Several primary studies did not fulfill the inclusion criteria: The pool of studies generated by the different search strategies included 135 studies that did not obtain school grades but alternative measures of scholastic achievement, did not report sample sizes, or did not provide a zero-order correlation coefficient or other information that allowed its calculation. These studies were omitted from our database. Some studies were based on identical datasets, so we had to exclude another 23 primary studies to ensure the independence of the included study coefficients. Finally, we were not able to retrieve 11 articles of those identified by the literature search and hence, to consider them for the current study. After the exclusion of these primary studies, 166 published and unpublished datasets remained for the current study.

Based on a sensitivity analysis regarding the sample sizes of the included primary studies, we had to exclude another four studies yielding a final dataset based on 162 primary studies (see Section 2.4.4).

2.3. Coding of studies

The final sample of primary studies was coded by three of the authors, including one of the first authors. For the coding, we employed a standardized coding scheme based in the information outlined in the following sections. All coders (i.e., for the initial and the double coding, see Section 2.3.6) received instructions on the coding process and the usage of the coding manual by one of the first authors.

2.3.1. Bibliographic information

We first documented bibliographic information about the included primary studies, consisting of (1) a consecutive primary study ID number, (2) the name(s) of the author(s), (3) the title of the article, (4) the year of publication, and (5) the name of the journal. In addition, we documented (6) the consecutive dataset ID number and (7) the name of the dataset which we either derived from the article or defined using the main characteristics of the sample (see also Appendix 1 where the study coefficients are given).

2.3.2. Sample characteristics

The documented sample characteristics comprised (1) information about the intelligence test or the specific scale of an intelligence test applied within the primary study, (2) characteristics of the grades reported in the primary studies, (3) sample size, (4) restriction of range, (5) school subjects the grades were obtained for, (6) grade level, (7) age, and (8) gender of the sample, as well as (9) the country the study was conducted in.

The name of each intelligence test applied in the primary studies was coded together with its reliability: if available, the exact reliability coefficient for the primary study sample was documented. Otherwise, we derived the reliability coefficient corresponding to the intelligence test from test manuals, compendia (Brähler, Holling, Leutner, & Petermann, 2002; Brickenkamp, 1997), or computerized databases containing basic information about commonly applied intelligence tests, such as Mental Measurements Yearbooks, Health and Psychosocial Instruments, and Hogrefe Testzentrale. In several instances we were not able to detect the reliability coefficient of a specific edition of an intelligence test applied within a primary study. In these cases, we obtained the reliability of a subsequent edition of the test as we assumed an improvement of the quality criteria of later editions — which would result in a conservative correction for attenuation (as described further below). Finally, if alternative sources did not provide particular information about the reliability of an intelligence test for a specific age group, gender or school type present in a primary study, we chose coefficients from manuals or computerized databases that matched the study sample characteristics best.

Regarding the school grades obtained within the included primary studies, we found an inverse polarity for several samples (meaning that lower numbers were associated with better grades). For these primary studies, the algebraic sign of the correlation coefficient (as described further below) was reversed.

The sample size was documented for each sample. We calculated an average sample size for those primary study samples which reported several correlation coefficients between general mental ability and school grades (e.g., for different school subjects) for the same overall sample but with slightly varying sample sizes for each of the correlation coefficients (e.g., slightly varying sample sizes for each school subject).

Indications of any range restriction in the independent variable (i.e., g) were documented and if reported, the standard deviation of the independent variable for the sample was coded.

We documented the school subjects for which school grades were obtained as far as they were reported within the articles. For each sample the corresponding grade level was documented.

2.3.3. Effect size

As for the effect measure, we coded the correlation between general mental ability and school grades for each study. In almost all cases, this was provided in the form of a correlation between the participants' results in the applied intelligence test(s) and their school grade(s).

One primary study reported t-values reflecting the significance of the correlations rather than the correlation coefficients themselves. We calculated the correlation coefficients for the corresponding samples.

2.3.4. Multiple effect sizes within primary studies

Various primary studies reported multiple correlation coefficients for a single sample (e.g., separate correlation coefficients for diverse intelligence tests, school subjects or grade levels). To avoid violating the independence assumption for study coefficients (see Hunter & Schmidt, 2004) without losing information, we decided to calculate an n-weighted average using Fisher's Z-values and included this single coefficient per sample. When primary studies reported separate coefficients for multiple samples (e.g., pupils from different schools or age groups) the independence assumption of study coefficients was not violated and the correlation coefficient for each sample was included.

2.3.5. Coding of moderator variables

2.3.5.1. Type of intelligence test. Samples for which intelligence was measured exclusively by either verbal or nonverbal intelligence tests were assigned to the subgroups Verbal and Nonverbal, respectively. A third subgroup Mixed contained all those samples for which *g* was measured by intelligence tests consisting of verbal as well as nonverbal scales. As only a small number of primary studies reported detailed information about the intelligence tests used, the sources described above (test manuals, compendia, and computerized databases) were consulted to obtain the relevant data for assigning the sample coefficients to the subgroups.

2.3.5.2. Subject domains. Depending on the school subject the school grade was derived for, samples were assigned to one of the following moderator subgroups: Mathematics and Science, Languages, Social Sciences, Fine Art and Music, and Sports.

2.3.5.3. Grade level. Each sample was assigned to one of the following subgroups, depending on the grade level for which school grades were assessed: Elementary School (including grade levels 1 to 4), Middle School (including grade levels 5 to 9), and High School (including grade levels 10 to 13).

2.3.5.4. Gender. We assigned samples consisting of either male or female participants to the corresponding subgroup Male and Female. Samples consisting of both boys and girls were not considered for this moderator analysis.

2.3.5.5. Year of publication. All primary studies published previously to the meta-analyses by Boulanger (1981), Fleming and Malone (1983) and Steinkamp and Maehr (1983), namely, before 1983, were assigned to the subgroup Before 1983. All primary studies published afterwards were assigned to the subgroup after 1983.

2.3.6. Interrater agreement

Half of the studies were coded by two coders in a double-blind setting. The interrater agreement reached 95.0% for the coding of the correlation coefficients, 100.0% for the coding of the reliability of the independent variable, and 98.8% for the coding of the sample sizes. Furthermore, both coders agreed at 100.0% for the coding of the subject domain, at 97.5% for the coding of the grade level, at 97.3% for the coding of the type of the intelligence test, and finally at 100.0% for the year of publication. The occurring differences of codings were discussed by the coders and eventually resolved by the first authors. Codings were then adjusted accordingly.

2.4. Meta-analytic procedure

2.4.1. Main meta-analysis

In our analysis we followed the procedures described by Hunter and Schmidt (2004). According to this approach, primary study results are attenuated by various artifacts beyond sampling error, such as unreliability of measurement scales and restriction of range in the independent or the dependent variable. “Psychometric meta-analysis” (Hunter & Schmidt, 2004) allows to correct these artifacts and hence, to estimate population correlations ρ. Employing this method for a meta-analysis of correlation coefficients first consists of estimating a simple mean correlation based on the included observed correlation coefficients weighted by sample size. This results in the “bare-bones” mean correlation, which is corrected for sampling error only and hence is comparable to the results from methods in the tradition of Hedges and colleagues (Hedges & Olkin, 1985; Hedges & Vevea, 1998). In a second step, measurement error and range restriction are corrected yielding the population correlation ρ. Psychometric meta-analysis provides two options for the correction of artifacts — individual correction of each sample coefficient and artifact correction using artifact distributions. Whereas the individual correction requires information on the reliability of measurement scales and the variance of the observed variables for each sample included in the analysis, the correction using artifact distributions can be employed if information on study artifacts is available for only a part of the primary studies (see Hunter & Schmidt, 2004).

As individual artifact correction was not possible for our analysis due to lack of information, we used artifact distribution meta-analysis instead. We corrected for sampling error, error of measurement in the predictor variable (i.e., *g*), and indirect range restriction. Error of measurement in the criterion variable (i.e., school grades) was not corrected for since our aim was to estimate the population correlation between g and school grades under realistic conditions (i.e., results with imperfect reliability). Information for the generation of the artifact distribution for predictor reliability was available for 20.0 to 100.0% of the coefficients, depending on the analysis. Information on range restriction in the criterion variable was available for 2.0 to 73.0% of the coefficients (see Appendix 1). For the meta-analytic calculations we used the software provided by Schmidt and Le (2005).

2.4.2. Moderator analyses

After the correction of all artifacts, the remaining heterogeneity in the population correlation ρ is assessed on the basis of the “75% rule” (Schmidt & Hunter, 1977): If 75% or more of the variance of the observed correlations can be attributed to the corrected artifacts, it is assumed that the remaining variance is due to further study artifacts which were not corrected. However, if the variance of the observed correlations lies below 75% after the correction of study artifacts, moderator variables might contribute to the total amount of variance and a search for potential moderator variables is necessary (see Hunter & Schmidt, 2004).

To test the significance of the difference of ρ between different moderator levels we used 95% confidence intervals (95%-CIs) as recommended by Hunter and Schmidt (2002), Schmidt and Hunter (1999), and Whitener (1990). A distinct mean difference and especially nonoverlapping confidence intervals were considered to be a good indicator for moderating effects. As the software employed in our study (i.e., Schmidt & Le, 2005) does not provide this option, we computed the intervals manually using the formula suggested by Hunter and Schmidt (2004). All intervals are reported in Table 1.

2.4.3. Analysis of availability bias

Meta-analytic findings may be biased either by the selection of the primary studies available for analysis or by an accumulation of significant results of the primary studies reported in scientific journals (Hunter & Schmidt, 2004). In order to identify the robustness of our findings against file-drawer effects, we computed fail-safe *N*s as recommended by Hunter and Schmidt (2004). In doing so, we applied the formulae derived by Pearlman (1982) and Orwin (1983), and according to McNatt (2000) we regarded correlations of $r = .05$ and below as trivial. For additional analyses of availability bias we used the package metafor (Viechtbauer, 2010) in R (R Core Team, 2015). We generated funnel-plots of the correlations employed in the main meta-analysis, following the guidelines of Light (1984). This graphical test allows to assess whether the correlations are symmetrically distributed around their mean. If the correlations above the mean correlation are overrepresented, a file-drawer bias can be concluded, since correlations below the mean have not been published because they are too small or not significant. The funnel-plots were additionally adjusted for missing studies using the trim and fill method (c.f. Duval & Tweedie, 2000). The idea of this method is to complement correlations in order to

Table 1
Meta-analytic results and moderator analyses.

			Artifact distribution				Meta-analysis corrected for sampling error (bare-bones)			Meta-analysis with full artifact correction						
	k[a]	N[b]	Mean r_{xx}[c]	SD_{rxx}	Mean u[d]	SD_u	Mean r	VAR_r	SD_r	ρ	σ^2_ρ	σ_ρ	95%-CI LB[e]	95%-CI UB[f]	Variance reduction (%)[g]	N_{FS}[h]
Main meta-analysis	240	105,185	0.86	0.08	0.85	0.21	0.44	0.03	0.18	0.54	0.03	0.17	0.51	0.57	31.7	2,358
Moderator: Type of intelligence test[i]	197	114,114	0.84	0.08	0.88	0.27	0.41	0.02	0.15	0.50	0.01	0.09	0.48	0.53	71.8	1,787
Verbal	59	45,672	0.83	0.09	0.86	0.35	0.42	0.03	0.16	0.53	0.01	0.09	0.48	0.58	74.7	561
Nonverbal	89	49,538	0.84	0.07	0.94	0.28	0.37	0.02	0.13	0.44	0.01	0.07	0.40	0.47	74.2	686
Mixed	49	18,904	0.87	0.07	0.81	0.18	0.47	0.01	0.12	0.60	0.01	0.08	0.56	0.64	64.9	537
Moderator: Subject domains[i]	262	143,052	0.86	0.09	0.88	0.21	0.37	0.02	0.14	0.45	0.01	0.11	0.43	0.47	49.2	2,105
Mathematics and science	100	60,533	0.86	0.09	0.90	0.21	0.42	0.01	0.11	0.49	0.00	0.06	0.47	0.52	73.4	887
Languages	96	61,865	0.85	0.10	0.90	0.21	0.36	0.01	0.12	0.44	0.01	0.09	0.41	0.47	60.5	748
Social sciences	41	12,649	0.90	0.04	0.84	0.19	0.35	0.02	0.12	0.43	0.01	0.19	0.39	0.48	53.7	314
Fine art and music	14	2,269	0.83	0.09	0.73	0.15	0.21	0.02	0.15	0.31	0.03	0.17	0.19	0.43	34.9	73
Sports	11	5,736	0.88	0.04	0.90	0.24	0.08	0.01	0.07	0.09	0.00	0.06	0.05	0.14	53.3	10
Moderator: Grade level[i]	217	83,782	0.86	0.08	0.86	0.21	0.45	0.03	0.18	0.54	0.03	0.16	0.51	0.57	32.1	2,120
Elementary school	71	18,584	0.86	0.11	0.96	0.12	0.40	0.01	0.12	0.45	0.01	0.10	0.42	0.48	39.8	568
Middle school	75	49,771	0.86	0.07	0.90	0.20	0.46	0.04	0.19	0.54	0.03	0.18	0.49	0.59	22.5	729
High school	71	15,427	0.86	0.07	1.05	0.12	0.46	0.05	0.21	0.58	0.04	0.20	0.51	0.64	29.4	747
Moderator: Gender	68	15,273	0.88	0.08	0.83	0.21	0.46	0.02	0.13	0.58	0.01	0.07	0.54	0.62	73.0	722
Male	37	7,780	0.84	0.10	0.83	0.14	0.46	0.02	0.13	0.58	0.01	0.10	0.53	0.64	48.3	395
Female	31	7,493	0.86	0.06	0.84	0.27	0.46	0.01	0.12	0.58	0.00	0.03	0.53	0.63	95.8	327
Moderator: Year of publication[i]	240	105,185	0.86	0.08	0.85	0.21	0.44	0.03	0.18	0.54	0.03	0.17	0.51	0.57	31.6	2,358
Before 1983	100	35,046	0.86	0.09	0.80	0.12	0.56	0.04	0.20	0.68	0.03	0.19	0.63	0.73	15.1	1,260
After 1983	140	70,139	0.86	0.07	0.86	0.23	0.38	0.02	0.15	0.47	0.01	0.12	0.45	0.50	50.9	1,188

[a] Number of coefficients used for analysis.
[b] Total N.
[c] Mean reliability of the predictor computed as arithmetic mean.
[d] Mean range restriction.
[e] Lower bound of 95% confidence interval (CI).
[f] Upper bound of 95% confidence interval (CI).
[g] Percent variance in observed correlations attributable to all artifacts.
[h] Fail-safe N.
[i] As the number of coefficients from studies with sufficient information for the moderator analysis varied, we present both the results for the full group in a specific moderator analysis as well as the results for each moderator subgroup separately.

achieve symmetry of the funnel-plot. The difference between the mean correlation of the uncorrected and the complemented distribution of correlations is an estimate of the file-drawer bias. These analyses were carried out for the bare-bones-analysis as well as for the meta-analysis with full artifact correction. As metafor doesn't provide options for the correction of reliability and range restriction, we corrected the correlations manually with the mean levels of the artifact distributions, following the guidelines of Arthur, Bennett, and Huffcutt (2001).

2.4.4. Sensitivity analyses

We investigated the robustness of our meta-analytic findings towards extreme sample sizes, outliers among the included correlation coefficients, and the values used for the artifact distributions.

2.4.4.1. Sample size. The dataset contained several primary studies with samples consisting of more than 10,000 individuals (Almquist, 2011; Brunner, 2006; Calvin, Fernandes, Smith, Visscher, & Deary, 2010; Hauser & Palloni, 2010). To explore a potential effect of the sample size on the meta-analytic results, we conducted a moderator analysis comparing the full dataset to a reduced dataset in which samples consisting of more than 10,000 participants were excluded.

2.4.4.2. Correlation coefficients. Outliers among the included correlation coefficients may affect the mean as well as the variance of the observed correlations. We examined the potential influence of outliers by conducting a moderator analysis contrasting the dataset containing all correlation coefficients to a dataset in which the highest 5% and the lowest 5% of the correlation coefficients were deleted.

2.4.4.3. Artifact distributions. The values used for artifact correction differed between the main meta-analysis and the moderator analyses depending on the information available on study artifacts for the particular group of primary studies included in the analysis. To provide evidence that our findings were robust to varying values included in the artifact distributions, we conducted a moderator analysis comparing a dataset with all available information on predictor reliability to a dataset in which the highest 5% and the lowest 5% of the reliability coefficients were deleted. The procedure was repeated for range restriction coefficients.

3. Results

3.1. Descriptives

The final data set in our meta-analysis included 162 primary studies published between 1922 and 2014 which provided $k = 240$ independent samples. The overall sample size was $N = 105{,}185$ with sample sizes varying from 15 to 9776. The participants' age at the time their general mental ability was measured was reported for 117 samples. The average age was 13.9 years ($SD = 4.0$). For a total of 160 samples information on the participant's gender was provided. Thirty-seven samples consisted exclusively of male participants, 31 samples were exclusively female. Overall, the samples consisted on average of 50.8% female participants. The studies we included had been conducted in 33 countries: Australia ($k = 6$), Austria ($k = 7$), Brazil ($k = 4$), Canada ($k = 10$), Central Philippines ($k = 1$), China ($k = 3$), Croatia ($k = 2$), Czech Republic ($k = 1$), Dubai ($k = 2$), Estonia ($k = 1$), Finland ($k = 2$), France ($k = 1$), Germany ($k = 49$), Great Britain ($k = 20$), Guatemala ($k = 2$), India ($k = 18$), Iran ($k = 1$), Iraq ($k = 1$), Italy

($k = 5$), Kenya ($k = 2$), Lebanon ($k = 4$), Luxembourg ($k = 2$), Netherlands ($k = 5$), Poland ($k = 1$), Portugal ($k = 2$), Russia ($k = 1$), Slovenia ($k = 2$), South Africa ($k = 1$), Spain ($k = 2$), Sweden ($k = 9$), Switzerland ($k = 12$), USA ($k = 48$), and Yemen ($k = 3$).

3.2. Population correlation

The mean observed correlation weighted by sample size (bare-bones meta-analysis) was $r = .44$ (95%-CIs for all reported coefficients are given in Table 1). The correction for error of measurement and indirect range restriction in the predictor variable resulted in a corrected population correlation of $\rho = .54$. However, only 31.7% of the variance of observed correlations was attributable to the three artifacts, indicating that a substantial amount of variance across the studies was due to factors not corrected for. Since the 75%-rule was not met, a generalization of the population correlation of this analysis is not possible. Instead, a search for moderator variables was necessary (see Hunter & Schmidt, 2004, p. 401).

3.3. Moderator results

3.3.1. Type of intelligence test

Mixed intelligence tests yielded the highest population correlation ($\rho = .60$), followed by verbal intelligence tests ($\rho = .53$). Since the confidence intervals of both moderator subgroups overlapped (Verbal: $.48 \leq \rho \leq .58$; Mixed: $.56 \leq \rho \leq .64$), they did not differ significantly in their validities. Nonverbal intelligence tests also predicted school grades very well ($\rho = .44$), but with a significantly lower population correlation than mixed and verbal measures as the 95%-CI of this group ($.40 \leq \rho \leq .47$) shows no overlap with the confidence intervals of the other groups. For all three subgroups, the amount of variance reduction (64.9% to 74.7%) showed that the corrected artifacts had been responsible to a substantial degree for the heterogeneity of the study coefficients included.

3.3.2. Subject domains

The examination of subject domains indicated a partially moderating effect. The subgroup Mathematics and Science yielded the highest population correlation ($\rho = .49$), followed by Languages ($\rho = .44$), Social Sciences ($\rho = .43$), Fine Art and Music ($\rho = .31$), and Sports ($\rho = .09$). Confidence intervals overlapped especially for the subgroups Mathematics and Science ($.47 \leq \rho \leq .52$), Languages ($.41 \leq \rho \leq .47$), Social Sciences ($.39 \leq \rho \leq .48$). The confidence interval of Fine Art and Music ($.19 \leq \rho \leq .43$) showed no overlap with the Mathematics and Science subgroup but with the Languages and the Social Sciences subgroup. The subgroup Sports ($.05 \leq \rho \leq .14$) differed significantly in its population correlation from all other subgroups. The correction of study artifacts resulted in a pronounced variance reduction for the subgroups Mathematics and Science (73.4%) and Languages (60.5%), and considerably less variance reduction for the remaining subgroups (34.9% to 53.7%).

3.3.3. Grade level

For grade level we again found a partially moderating effect. The subgroup High School yielded the highest population correlation ($\rho = .58$), followed by the subgroups Middle School ($\rho = .54$) and Elementary School ($\rho = .45$). According to the corresponding confidence intervals, the population correlation between *g* and school grades was significantly higher for school grades in high school and middle school than in elementary school (Elementary School: $.42 \leq \rho \leq .48$; Middle School: $.49 \leq \rho \leq .59$; High School: $.51 \leq \rho \leq .64$). The extent of variance reduction after correcting for study artifacts was low to moderate for all three subgroups (22.5% to 39.8%).

3.3.4. Gender

The male samples yielded the same population correlation ($\rho = .58$) as the female samples ($\rho = .58$) and a similar confidence interval (Male: $.53 \leq \rho \leq .64$; Female: $.53 \leq \rho \leq .63$). The correction of study artifacts resulted in a moderate variance reduction for male samples (48.3%), while for female samples study artifacts accounted for most of the heterogeneity of the study coefficients (95.8%).

3.3.5. Year of publication

There was a significant moderating effect of the year in which the primary studies were published. The population correlation was higher for the period before the previous meta-analyses were conducted ($\rho = .68$), and lower for the period afterwards ($\rho = .47$), with confidence intervals clearly indicating a significant difference between both subgroups (Before 1983: $.63 \leq \rho \leq .73$; After 1983: $.45 \leq \rho \leq .50$). The extent of the variance reduction after correcting for study artifacts was small for the subgroup Before 1983 (15.1%), and larger for the subgroup after 1983 (50.9%).

3.4. Availability bias

According to the fail-safe *N* it would take 2358 null findings to reduce the mean effect to a trivial size. The fail-safe *N*s on the moderator levels also clearly exceed the amount of coefficients included in the analyses. We therefore consider the validity of this result as rather robust against file-drawer bias.

The funnel-plots for the correlations employed in the bare-bones and full artifact correction main meta-analysis can be found in parts (a) and (b) in Fig. 1. In both, the correlation coefficients are not distributed symmetrically around their mean. Instead, the correlations below the mean are overrepresented, indicating that studies with correlations above the mean are missing. As this is actually the opposite of a publication bias, one could only conclude that the mean correlation is lowered by missing studies. This assumption is supported by the results of the trim and fill analyses presented in parts (c) and (d) of Fig. 1. In both cases the trim and fill corrected correlations are higher than the uncorrected correlations. For the bare-bones meta-analysis 49 studies have to be complemented on the right side, which results in a corrected mean r of .49. For the full meta-analysis 53 studies have to be complemented on the right side, which results in a corrected ρ of .61.

3.5. Sensitivity analyses

We conducted sensitivity analyses to investigate the robustness of our findings towards extreme sample sizes and outliers among the included correlation coefficients and the values used for the artifact distributions (for detailed results see Table 2).

3.5.1. Sample size

We compared the full dataset including 244 samples obtained from our literature search to a reduced dataset in which samples consisting of more than 10,000 participants (Almquist, 2011; Brunner, 2006; Calvin et al., 2010; Hauser & Palloni, 2010) were excluded. The estimated population correlations substantially differed between both datasets, showing an upward bias for the full dataset compared to the reduced dataset. In order to estimate robust population level effects and to avoid bias due to extreme sample sizes, the corresponding samples were excluded from the analysis. As described in Section 3.1 the final dataset consisted of 162 primary studies.

3.5.2. Correlation coefficients

To examine the potential influence of outliers we contrasted the dataset containing all correlation coefficients to a dataset in which the highest 5% and the lowest 5% of the correlation coefficients were omitted. Both datasets produced comparable results indicating that the results of our study are robust to outliers among the included effect sizes.

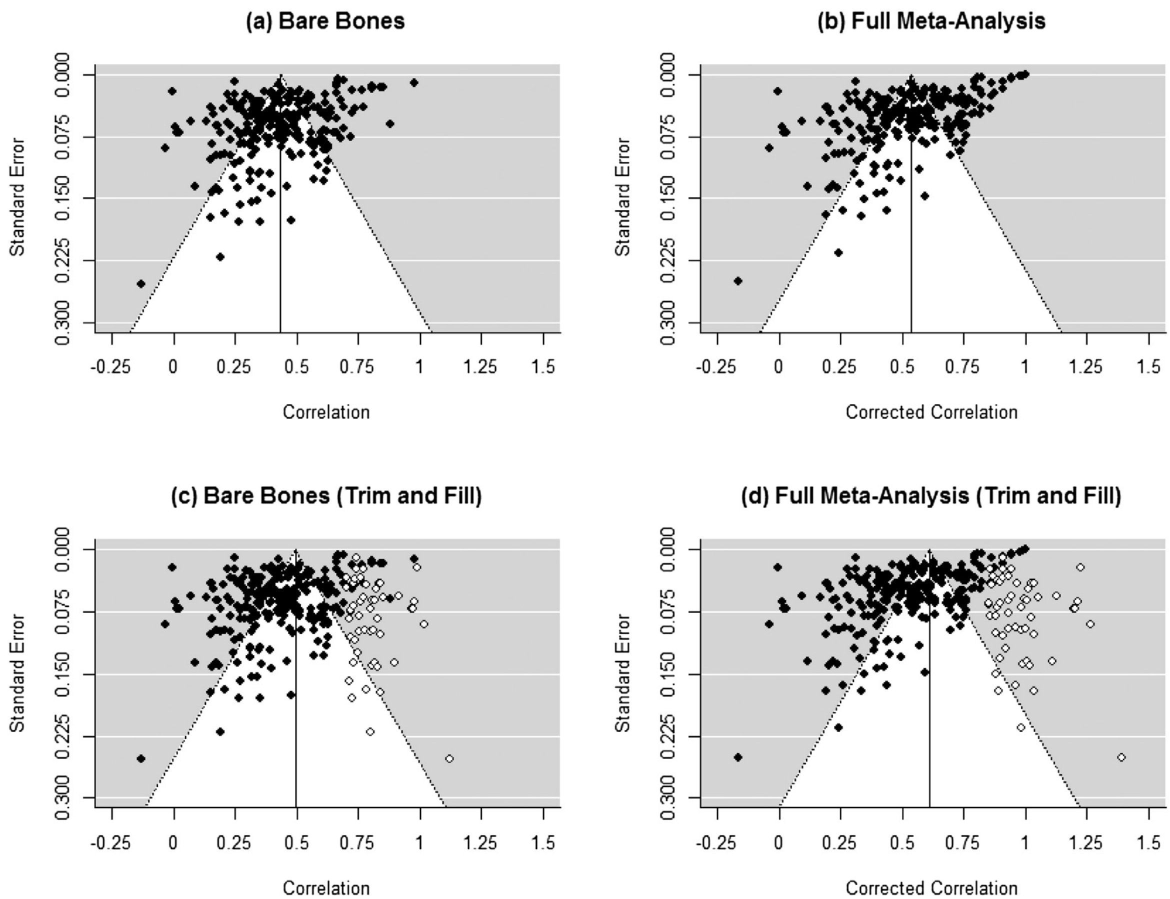

Fig. 1. Funnel-plots for the main meta-analysis. Dotted lines represent the 95%-standard error of the mean correlation. Black circles represent correlations employed in the analysis. White circles represent correlations complemented by the trim and fill method.

3.5.3. Artifact distributions

We analyzed the robustness of our findings to varying values included in the artifact distributions to correct for predictor reliability and range restriction comparing a dataset with all available information on the corresponding artifact to a dataset in which the highest 5% and the lowest 5% of the artifact coefficients were deleted. For both artifacts, we found

Table 2
Sensitivity analyses for sample size, correlation coefficients, and artifact distributions.

			Artifact distribution				Meta-analysis corrected for sampling error (bare-bones)			Meta-analysis with full artifact correction						
	k[a]	N[b]	Mean r_{xx}[c]	SD_{rxx}	Mean u[d]	SD_u	Mean r	VAR_r	SD_r	ρ	σ^2_ρ	σ_ρ	95%-CI LB[e]	95%-CI UB[f]	Variance reduction (%)[g]	N_{FS}[h]
Sample size																
Full dataset	244	336,386	0.86	0.08	0.85	0.21	0.58	0.04	0.20	0.69	0.03	0.17	0.66	0.72	26.4	3,111
Reduced dataset	240	105,185	0.86	0.08	0.85	0.21	0.44	0.03	0.18	0.54	0.03	0.17	0.51	0.57	31.6	2,358
Correlation coefficients																
All correlation coefficients	240	105,185	0.86	0.08	0.85	0.21	0.44	0.03	0.18	0.54	0.03	0.17	0.51	0.57	31.6	2,358
Outliers deleted	216	97,274	0.86	0.08	0.85	0.21	0.45	0.03	0.16	0.55	0.02	0.13	0.53	0.58	44.1	2,167
Artifact distribution: Predictor reliability																
All available information	240	105,185	0.86	0.08	0.85	0.21	0.44	0.03	0.18	0.54	0.03	0.17	0.51	0.57	31.6	2,358
Outliers deleted	240	105,185	0.86	0.06	0.85	0.21	0.44	0.03	0.18	0.54	0.03	0.17	0.51	0.57	30.9	2,349
Artifact distribution: Range restriction																
All available information	240	105,185	0.86	0.08	0.85	0.21	0.44	0.03	0.18	0.54	0.03	0.17	0.51	0.57	31.6	2,358
Outliers deleted	240	105,185	0.86	0.08	0.84	0.17	0.44	0.03	0.18	0.54	0.03	0.18	0.52	0.57	24.1	2,373

[a] Number of coefficients used for analysis.
[b] Total N.
[c] Mean reliability of the predictor computed as arithmetic mean.
[d] Mean range restriction.
[e] Lower bound of 95% confidence interval (CI).
[f] Upper bound of 95% confidence interval (CI).
[g] Percent variance in observed correlations attributable to all artifacts.
[h] Fail-safe N.

highly comparable results for the full and the reduced datasets and hence, no substantial effect of varying artifact distributions on the meta-analytical findings. This indicates that the results of our study are robust to the choice of values included in the artifact distributions.

4. Discussion

The goal of the present study was to provide a comprehensive and up-to-date meta-analytic examination of the correlation between intelligence and school grades. In doing so, we were additionally interested in evaluating the influence of a set of moderators. The meta-analytic strategy of Hunter and Schmidt (2004) was used, as a correction of statistical artifacts (i.e., sample size, reliability, range restriction) is possible when using this approach. In the following paragraphs the results of the main meta-analysis and the meta-analyses on the moderator levels will be interpreted and possible reasons for differences between moderator levels will be discussed.

The central finding of our meta-analysis for the full sample is a substantial mean correlation of $\rho = .54$ between intelligence and school grades which can be regarded as significant since the respective confidence interval does not include zero. The large size of this correlation clearly demonstrates the importance intelligence has for this measure of academic achievement. This result corresponds well with the assertion of narrative reviews which refer to a mean correlation of .5 between the two variables (e.g., Gottfredson, 2002; Gustafsson & Undheim, 1996; Neisser et al., 1996; Sternberg et al., 2001). Nevertheless, it would be an oversimplification to generalize this finding because of substantial residual variance, which cannot be explained by the methodological artifacts for which we corrected (i.e., sample size, predictor reliability, range restriction). Therefore, it is interesting to additionally regard the results of our moderator analyses.

All of the three test types (verbal, nonverbal, mixed) we considered in our analyses possessed significant and substantial population correlation between g and school grades. The comparatively higher population correlation for verbal and mixed tests may be explained by the fact that verbal skills (e.g., speech comprehension, linguistic expression) are of particular importance for the successful participation in class as well as in written exams, which are in turn the basis of school grades. As verbal and mixed tests cover these skills to a greater degree, they allow a better prediction of the grades. Although there is an overlap between the 95%-CIs for verbal ($.48 \le \rho \le .58$) and mixed tests ($.56 \le \rho \le .65$) it can be regarded as rather small. Therefore, we view a cautious interpretation of the difference of these two moderator groups as feasible. Mixed tests provide a broader measure of intelligence in the sense of g, as they consist of both, verbal and nonverbal material (c.f. Jensen, 1998). Thus, mixed tests are able to cover variance of nonverbal tests, which are not shared with verbal tests, but which are relevant for the prediction of the criterion. This is in line with the position of several authors (e.g., Gottfredson, 2002; Gustafsson & Undheim, 1996; Jensen, 1998; Mayer, 2000; Neisser et al., 1996; Sternberg et al., 2001) who view g as a better predictor of scholastic achievement than specific aspects of intelligence such as verbal ability. A glance at the variance which is reduced by the correction of the artifacts reveals that the population correlation for the verbal (74.7% variance reduction) and nonverbal subgroup (74.2% variance reduction) can be generalized. The mixed subgroup shows substantial residual variance (64.9% variance reduction). Therefore, the population correlation cannot be generalized. This may correspond with the fact that the tests used in this group are less homogenous (i.e., they can contain more or less nonverbal or verbal items) than in the other two groups in which solely verbal or nonverbal items are employed.

Concerning the moderating effect of the subject domains, our analyses showed that the mean corrected correlation between scores of intelligence tests and school grades was highest in the Mathematics and Science subgroup ($\rho = .49$). Although there is an overlap between the 95%-CI ($.47 \le \rho \le .52$) of this group and the confidence intervals of the Languages ($.41 \le \rho \le .47$) and the Social Sciences subgroups ($.39 \le \rho \le .48$) it can be regarded as rather small. Therefore a moderating effect can be assumed carefully. This finding appears to be rather straightforward since mathematics and science are subjects which deal with content that relies heavily on logic. As logical thinking is the most dominant competence assessed by intelligence tests, persons with higher tests scores should understand the content of these subject domains better and thus have better grades. A second explanation involves the reliability of school grades in these subjects. As answers in written exams in mathematics and science can easily be evaluated as right or wrong there is no margin of judgment for the teachers when giving the grades. Thus, the reliability of grades in these school subjects should be higher than in the other school subjects where there is clearly a wider margin of judgment, which in turn influences the height of the correlation which can be achieved maximally. The 95%-CI of the Languages subgroup ($.41 \le \rho \le .47$) is completely included in the confidence interval of the Social Sciences ($.39 \le \rho \le .48$) subgroup. Thus, no moderating effect can be assumed between the two subgroups. This and the finding that the mean corrected validities are lower than in the Mathematics and Science subgroup can be explained by the fact that the successful participation in these subjects requires learning content (e.g. vocabulary, historical data) by heart. Thus, motivational aspects play a more important role than in the Mathematics and Science subgroup where it is predominantly important to understand content. Furthermore, there is a wider margin of judgment for teachers when grading answers in written exams, especially in higher grades where pupils have to interpret e.g. texts or historical constellation. This results in a lower reliability of the grades which in turn lowers the correlation that can be maximally achieved between intelligence and school grades. Nevertheless, the mean corrected correlation between intelligence and school grades in Language ($\rho = .44$) and Social Sciences ($\rho = .43$) still is rather high, which indicates that cognitive ability is a substantial prerequisite for scholastic success in these subjects. The 95%-CI of the Fine Art and Music subgroup ($.19 \le \rho \le .43$) shows negligible overlap with the confidence interval of the Languages subgroup ($.41 \le \rho \le .47$). Although there is a greater overlap with the 95%-CI of the Social Sciences subgroup ($.39 \le \rho \le .48$) we tend to carefully regard the difference as substantial since the confidence interval of the Fine Art and Music subgroup is rather wide. This difference can be explained by the fact that the margin of judgment when giving grades in these subjects can be regarded as the widest, as there are no truly right or wrong “answers”. Thus, the reliability of the grades can be regarded as the lowest, which leads to an attenuation of the correlation between intelligence test scores and the grades in these subjects. Nevertheless there is still a substantial corrected mean correlation ($\rho = .31$) between the two variables which is in line with the literature suggesting a relationship between musicality and intelligence (e.g., Schellenberg, 2005) as well as creativity and intelligence (Batey & Furnham, 2006). A clear moderating effect was found for the Sports subgroup indicating that the population correlation for this subgroup ($\rho = .09$) is significantly the lowest. This can be explained by the fact that the grades in this subject reflect aspects of bodily functioning rather than academic achievement. As intelligence tests assess cognitive competencies which are relevant for academic achievement they are less relevant for the grades in this subject. Nevertheless, the population correlation still is significant, since the 95%-CI does not include zero, reflecting that intelligence correlates with variables that are in turn associated with athletic success, e.g., body height (Gale, 2005) or absence of obesity (Smith, Hay, Campbell, & Trollor, 2011).

The moderator analyses concerning grade levels revealed that population correlation was lowest in elementary school ($\rho = .45$), increased throughout middle school ($\rho = .54$), and was highest in high school ($\rho = .58$). The 95%-CI of the Elementary School subgroup ($.42 \le \rho \le .48$) showed no overlap with the 95%-CIs of the other subgroups. Thus, a clear moderating effect can be assumed. The 95%-CIs of the Middle School ($.49 \le \rho \le .59$) and the High School subgroup

($.51 \leq \rho \leq .64$) showed considerable overlap. Therefore, the differences between the mean corrected validities of these subgroups can at the most be interpreted as a tendency. The ranking of the correlations appears to be independent from range restriction since the results of the bare-bones analyses in which range restriction was not corrected show a relatively similar pattern [Elementary School: $M(r) = .40$; Middle School: $M(r) = .46$; High School: $M(r) = .46$]. Thus, the results run counter to the position brought forward by Brody (1992) and Jensen (1998) who expected a decrease of the population correlation throughout grade levels because of range restriction. What might explain our findings to some degree is that it may be easier to compensate deficits in intelligence by practice in lower grades because the content taught is easier to understand. As the content becomes more demanding throughout grade levels it should be increasingly difficult to compensate for intelligence deficits through practice alone. This effect may be strong enough to superimpose the decreasing effect of range restriction and lead to higher validities in higher grade levels.

Results concerning the year of publication suggested a substantially higher mean correlation in the studies published before 1983 ($\rho = .68$) compared to the studies published afterwards ($\rho = .47$). A possible explanation for this finding might involve grade inflation, which describes the observation that throughout the last decades progressively better grades are awarded for work that would have received worse grades in the past. This effect appears to be rather robust for Germany (Kersting, 2015) and the USA (Rojstaczer & Healy, 2010). Since most of the studies in our analyses come from Germany and the USA, this may have influenced the observed results. As the upper limit of the grade scale is fixed, better grades lead to range restriction, which in turn leads to lower correlations between intelligence and school grades. A comparison between our findings and the findings of previous meta-analyses is feasible only on a very careful basis since different meta-analytic strategies were applied. Furthermore, only the results of our bare-bones analyses can be used since reliability and range restriction were not corrected for in the previous meta-analyses. The comparison shows that the mean correlation of the studies published before 1983 was higher in our analysis [$M(r) = .56$] than the mean correlations reported by Boulanger [1981; $M(r) = .48$] Fleming and Malone (1983; $\rho = .43$) and Steinkamp and Maehr (1983; $\rho = .34$). It has to be noted that the previous meta-analyses used school achievement tests as the criterion whereas we used school grades. School achievement tests represent a narrower measure of scholastic performance since they are based on a selective assessment, which is only based on written sources and might be influenced by the current mental state of the testees (e.g., vigilance, mood). School grades are based on a broader information basis since they incorporate scholastic performance over a longer period of time and stem from different sources (written exams, participation in class). As school grades are less prone to error resulting from temporary mental states and individual strengths of the testees (written vs. lingual performance), they provide a more reliable (and arguably more relevant) measure of scholastic achievement, which interestingly shows higher correlations to intelligence.

Finally, in the analyses concerning the participant's gender no sex differences emerged. The mean corrected validities as well as the 95%-CIs were nearly identical for female ($\rho = .58$; $.53 \leq \rho \leq .63$) and male participants ($\rho = .58$; $.53 \leq \rho \leq .64$). Nevertheless, only the population correlation for the female participants can be generalized since there was a substantial variance reduction due to the correction of artifacts (95.8%). The distribution of validities for male participants was inhomogeneous even after correcting for artifacts (48.3%). Sex differences in non-cognitive variables, which have relevance for school grades in addition to intelligence, may serve as an explanation for this finding. For example, it was reported that school-related intrinsic motivation, school anxiety, and performance-avoidance goals explain additional variance in school grades only for male pupils (Freudenthaler, Spinath, & Neubauer, 2008). As the distribution of these variables may vary between studies but would only have impact on the validities in the male subgroup, one would expect a higher variability there but not in the subgroup of female pupils.

The analyses concerning file-drawer bias suggest that the correlations we found are not boosted by the publication practice, which prefers high and significant correlations. The fail-safe *Ns* indicate that it would take an enormous amount of insignificant studies to reduce the mean effect of the main meta-analysis as well as of the meta-analyses on moderator levels to a trivial size. The inspection of the funnel-plots shows that if anything, a negative file-drawer bias can be concluded, which artificially reduced the mean correlation of our analysis. The results of the trim and fill analyses suggest that the potential file-drawer bias is not negligible, since the mean correlation is boosted from .44 to .49 for the bare-bones main meta-analysis and from .54 to .61 for the full artifact correction meta-analysis.

We examined the robustness of our findings towards a potential effect of outliers among sample sizes, correlation coefficients and the values used for the artifact distributions. Regarding extreme sample sizes, we had to exclude four primary studies from the dataset used for the current meta-analysis as they produced an upward bias in our findings. However, with regard to outliers among correlation coefficients, we found no effect on the meta-analytical findings. Likewise, the investigation of extreme values included in the artifact distributions to correct for predictor reliability and range restriction revealed unbiased results. In summary, after the exclusion of primary studies with extreme sample sizes, the sensitivity analyses did not reveal any bias due to outliers among correlation coefficients and the values included in the artifact distributions and thus confirm the robustness of the findings reported in the current study towards extreme values.

5. Conclusion

The results of our study clearly show that intelligence has substantial influence on school grades and thus can be regarded as one of the most (if not *the* most) influential variables in this context. Although intelligence turned out to be a significant predictor on all moderator levels, we were able to identify some scenarios in which even higher validities can be obtained. First of all, the population correlation was highest for tests relying on both verbal and nonverbal materials, indicating that a broad measure of intelligence or *g* respectively is the best predictor of school grades. Furthermore, the importance of intelligence increases throughout grade levels. This leads us to the conclusion that intelligence has special importance in educational contexts which deal with content that is more complex and thus can be mastered fully only with an appropriate cognitive ability level. Regarding the learning content it can be concluded that intelligence has its highest population correlation in subject domains, which focus on content that has a clearer logical structure, because this is a key component of what intelligence tests assess. Our results also show that the population correlation between *g* and school grades was higher in studies published before 1983 compared to studies published afterwards. Although the relevance of intelligence seems to be lower nowadays, it is still substantial. Furthermore, this finding suggests that selection procedures for tertiary education (colleges, universities) and employment should incorporate intelligence tests in addition to school grades, since intelligence most certainly provides incremental validity, as the amount of shared variance between intelligence and school grades is lower at this level. In our analysis the population correlation between *g* and school grades was independent from the testees' gender. Nevertheless, only the results for female groups were generalizable, suggesting that there are additional variables which influence the population correlation in the male group (e.g., motivation, school anxiety, performance-avoidance-goals).

Acknowledgments

We thank two anonymous reviewers for their valuable comments on earlier versions of this article.

Appendix 1. Overview of the characteristics of the included primary studies

Author (year)	Description of the sample	Country	n[a]	Intelligence test	Type of intelligence test	Subject domain	Grade level	Gender[b]	Age	r[c]	r_{xx}[d]	u_x[e]
Adkins (1937)	High school sample	USA	77	Point Subject Ratio	n.a.	n.a.	8	n.a.	n.a.	0.56[f]	0.89	n.a.
Agnoli et al. (2012)	Bolognia sample	Italy	352	Raven's Progressive Matrices	Nonverbal	n.a.	4	43.4	9.4	0.38[f]	0.80	n.a.
Al-Ghamri (2012)	Arab sample I	Yemen	682	Otis–Lennon Mental Ability Test	n.a.	n.a.	10	48.0	16.9	0.81	0.83	n.a.
Al-Ghamri (2012)	Arab sample II	Yemen	466	Otis–Lennon Mental Ability Test	n.a.	n.a.	11	48.1	16.9	0.81	0.83	n.a.
Al-Ghamri (2012)	Arab sample III	Yemen	413	Otis–Lennon Mental Ability Test	n.a.	n.a.	12	50.6	16.9	0.80	0.83	n.a.
Allana (2010)	Private School sample I	Dubai	15	Raven's Standard Progessive Matrices	Nonverbal	n.a.	5	n.a.	9.5	0.98	0.87	n.a.
Allana (2010)	Private School sample II	Dubai	15	Raven's Standard Progressive Matrices	Nonverbal	n.a.	6	n.a.	9.5	0.88	0.87	n.a.
Ames (1943)	1939 sample	USA	256	Otis S–A Intelligence Test	Mixed	n.a.	9	n.a.	n.a.	0.59	n.a.	n.a.
Ames (1943)	1940 sample	USA	277	Otis S–A Intelligence Test	Mixed	n.a.	9	n.a.	n.a.	0.59	n.a.	n.a.
Ames (1943)	1941 sample	USA	281	Otis S–A Intelligence Test	Mixed	n.a.	9	n.a.	n.a.	0.54	n.a.	n.a.
Anonymous[g]	Anonymous sample	n.a.	93	Performance Test System 1–2 (German version); Intelligenz-Struktur-Test 2000 R	n.a.	n.a.	13	n.a.	n.a.	0.23[f]	0.97	n.a.
Anonymous[g]	Anonymous sample	n.a.	58	Performance Test System 1–2 (German version); Intelligenz-Struktur-Test 2000 R	n.a.	n.a.	13	n.a.	n.a.	0.24[f]	0.97	n.a.
Anonymous[g]	Anonymous sample	n.a.	1,508	German Cognitive Ability Test — 4–12 — Revision	Nonverbal	n.a.	5	49.0	10.9	0.25[f]	0.87	1.13
Anonymous[g]	Anonymous sample	n.a.	529	Culture Fair Intelligence Test-20 (German adaptation)	Nonverbal	n.a.	3	49.2	8.9	0.33[f]	0.90	1.13
Anonymous[g]	Anonymous sample	n.a.	718	Culture Fair Intelligence Test-20 (German adaptation)	Nonverbal	n.a.	3	49.5	9.5	0.35[f]	0.90	1.12
Anonymous[g]	Anonymous sample	n.a.	700	Culture Fair Intelligence Test-20 (German adaptation)	Nonverbal	n.a.	4	48.9	10.4	0.36[f]	0.90	1.07
Anonymous[g]	Anonymous sample	n.a.	710	Culture Fair Intelligence Test-20 (German adaptation)	Nonverbal	n.a.	5	49.5	11.4	0.33[f]	0.90	1.08
Anonymous[g]	Anonymous sample	n.a.	715	n.a.	Verbal	n.a.	n.a.	n.a.	n.a.	0.25	n.a.	n.a.
Anonymous[g]	Anonymous sample	n.a.	1,389	Berlin Structure of Intelligence Test for Youth: Assessment of Talent and Giftedness	Mixed	n.a.	9	59.0	14.5	0.29[f]	0.74	0.69
Anonymous[g]	Anonymous sample	n.a.	85	Berlin Structure of Intelligence Test	Mixed	n.a.	11	46.5	16.7	0.44	0.83	0.55
Anonymous[g]	Anonymous sample	n.a.	277	Analyse des Schlussfolgernden und Kreativen Denkens	Verbal	n.a.	12	61.1	17.6	0.46	0.89	0.66
Anonymous[g]	Anonymous sample	n.a.	312	Wonderlic Personal Test (German version)	Verbal	n.a.	12	59.9	17.4	0.07	0.90	0.66
Aswal (2001)	P.I.C. Tehri College students sample	India	65	General Mental Ability Group Test by Jolata and Singh	n.a.	n.a.	11	n.a.	n.a.	0.58	n.a.	n.a.
Aswal (2001)	G.G.I.C. Tehri College students sample	India	25	General Mental Ability Group Test by Jolata and Singh	n.a.	Mathematics and Science	11	n.a.	n.a.	0.61	n.a.	n.a.
Aswal (2001)	V.H.S. Koti College students sample	India	50	General Mental Ability Group Test by Jolata and Singh	n.a.	Mathematics and Science	11	n.a.	n.a.	0.39	n.a.	n.a.
Aswal (2001)	G.I.C. Pokhal College students sample	India	35	General Mental Ability Group Test by Jolata and Singh	n.a.	Mathematics and Science	11	n.a.	n.a.	0.46	n.a.	n.a.
Aswal (2001)	G.I.C. Sirai College students sample	India	25	General Mental Ability Group Test by Jolata and Singh	n.a.	Mathematics and Science	11	n.a.	n.a.	0.35	n.a.	n.a.
Axelsson (2009)	Sweden sample	Sweden	124	Raven's Progressive matrices; Deltatest — submodule Number Series	n.a.	Mathematics and Science	n.a.	100.0	29.9	0.50	n.a.	n.a.
Barton, Dielman, and Cattell (1971)	Sixth grade sample	USA	169	Culture Fair Intelligence Test	Nonverbal	n.a.	6	50.0	n.a.	0.43[f]	n.a.	n.a.
Barton et al. (1971)	Seventh grade sample	USA	142	Culture Fair Intelligence Test	Nonverbal	n.a.	7	50.0	n.a.	0.49[f]	n.a.	n.a.
Beckett, Castle, Rutter, and Sonuga-Barke (2010)	GCSE-sample	Great Britain	161	Wechsler Intelligence Scale for Children	n.a.	n.a.	n.a.	n.a.	15.0	0.64	0.97	1.09
Bipp, Steinmayr, and Spinath (2012)	Goal orientation sample	Germany	164	Intelligenz Struktur Test 2000 R	Mixed	n.a.	11	53.1	16.5	0.35	0.89	0.84
Blöschl (1966)	Blöschl's sample	Germany	125	Performance Test System (German version); Hamburger-Wechsler-Intelligence Test for Children	n.a.	n.a.	n.a.	41.6	10.9	−0.03[f]	0.92	n.a.
Blue (2009)	Blue's sample	USA	452	Cognitive Abilities Test	n.a.	n.a.	7	50.7	n.a.	0.32[f]	0.95	n.a.
Bose (1982)	Calcutta sample	India	250	Culture Fair Intelligence Test	Nonverbal	n.a.	n.a.	n.a.	n.a.	0.18	0.90	n.a.
Bouffard, Vezeau, Roy, and Lengelé (2011)	Quebec sample	Canada	462	Mental Ability Test (French version)	Mixed	n.a.	n.a.	43.3	8.6	0.77[f]	0.85	n.a.
Boulon-Díaz (1992)	Puerto Rican sample	USA	65	Wechsler Intelligence Scale for Children — Revised, Puerto Rican version	Mixed	n.a.	5	n.a.	10.0	0.66	0.89	n.a.
Bowers (1966)	Bowers' sample	USA	278	Otis Quick Scoring Mental Ability Test	Mixed	n.a.	9	n.a.	n.a.	0.57	0.86	0.77
Bratko, Chamorro-Premuzic, and	Bratko et al.'s sample	Croatia	255	Multifactor Test Battery; Serial Numbers; Surface Development	n.a.	n.a.	n.a.	n.a.	n.a.	0.49	0.91	n.a.

(continued on next page)

Appendix 1. *(continued)*

Author (year)	Description of the sample	Country	n[a]	Intelligence test	Type of intelligence test	Subject domain	Grade level	Gender[b]	Age	r[c]	r_{xx}[d]	u_x[e]
Saks (2006)				Test–Vz 3; Adaptation of Thurstone First Letter Test								
Brucks (1978)	Hollfeld sample	Germany	185	Frankfurter Analogietest 4–6	Mixed	n.a.	5	54.6	11.0	0.61[f]	0.93	n.a.
Burgert (1937)	San Diego sample	USA	191	Terman Group Test	n.a.	n.a.	6	n.a.	n.a.	0.48	0.93	0.86
Carter (1959)	Grade 11 small California High School sample	USA	116	Hemnon–Nelson Intelligence Test	Verbal	n.a.	11	n.a.	n.a.	0.66	n.a.	n.a.
Carter (1959)	Grade 11 large California High School sample	USA	239	ACE Intelligence Test	n.a.	n.a.	10	n.a.	n.a.	0.53	n.a.	n.a.
Carter (1959)	Grade 11 California High School sample	USA	211	ACE Intelligence Test	n.a.	n.a.	10	n.a.	n.a.	0.50	n.a.	n.a.
Chamorro-Premuzic, Quiroga, and Colom (2009)	Madrid sample	Spain	248	Raven's Advanced Progressive Matrices; Solid figures; Primary Mental Ability Battery; Differential Aptitude Test Battery	n.a.	n.a.	n.a.	81.0	20.1	0.01[f]	n.a.	n.a.
Cocking and Holy (1927)	Iowa sample	USA	266	Thorndike Intelligence Examination for High School Graduates	n.a.	n.a.	n.a.	40.2	n.a.	0.42	0.92	n.a.
Cooper (1974)	Ottawa sample	Canada	527	Hemnon–Nelson Tests; Differential Aptitude Test	n.a.	n.a.	10	n.a.	n.a.	0.20[f]	n.a.	n.a.
Craig (1988)	Craig's sample	USA	105	Test of Cognitive Skills	n.a.	n.a.	n.a.	n.a.	n.a.	0.50[f]	0.81	n.a.
Daley et al. (2005)	Female sample	Kenya	243	Adaptation of the Peabody Picture Vocabulary Test; Adaptation of the Wechsler Intelligence Scale for Children — Revised	n.a.	n.a.	1	n.a.	n.a.	0.34[f]	n.a.	n.a.
Daley et al. (2005)	Male sample	Kenya	234	Adaptation of the Peabody Picture Vocabulary Test; Adaptation of the Wechsler Intelligence Scale for Children — Revised	n.a.	n.a.	1	n.a.	7.5	0.40[f]	n.a.	n.a.
Dash, Mohanty, and Kar (1989)	Orissa sample	India	60	Compound-Stimulus Visual Information Test; Serial-Stimulus Visual Information Test; Forward Digit Span Test; Backward Digit Span Test; Adaptation of the Wechsler Intelligence Scale for Children: Similarities subtest & Vocabulary subtest; Raven's Colored Progressive Matrices; Wechsler Intelligence Test: Block Design subtest	n.a.	n.a.	5	n.a.	10.0	0.31[f]	0.82	n.a.
Day, Hanson, Maltby, Proctor, and Wood (2010)	A-level sample	GB	129	Raven's Progressive Matrices	Nonverbal	n.a.	n.a.	59.7	18.6	0.43	0.92	n.a.
Di Fabio and Palazzeschi (2009)	Toscana sample	Italy	124	Advanced Progressive Matrices	Nonverbal	n.a.	n.a.	72.6	17.5	0.32	n.a.	1.09
Dodonova and Dodonov (2012)	Moscow sample	Russia	184	Raven's Advanced Progressive Matrices; Intelligence-Structure-Test	n.a.	n.a.	n.a.	62.0	16.0	0.19[f]	n.a.	0.74
Downey, Lomas, Billings, Hansen, and Stough (2013)	Victoria sample	Australia	243	Raven's Standard Progressive Matrices	Nonverbal	n.a.	9	100.0	14.6	0.46	n.a.	1.44
Dresel, Fasching, Steuer, and Berner (2010)	Dresel et al.'s sample	Germany	796	Culture Fair Test-20x	Nonverbal	Languages	4	54.0	9.8	0.23	0.69	0.74
Duckworth, Quinn, and Tsukayama (2012)	Texas sample I	USA	706	Wechsler Abbreviated Scale of Intelligence	Mixed	n.a.	8	48.0	n.a.	0.48[f]	n.a.	0.96
Duckworth et al. (2012)	Texas sample II	USA	510	Raven's Progressive Matrices	Nonverbal	n.a.	n.a.	n.a.	11.7	0.41[f]	0.80	n.a.
Edds and McCall (1933)	Edd and McCall's sample	USA	85	Otis Group Intelligence Scale	Mixed	n.a.	n.a.	n.a.	n.a.	0.32[f]	n.a.	n.a.
Edminston and Rhoades (1959)	Edminston and Rhoades's sample	USA	94	California Test of Mental Maturity	Mixed	n.a.	n.a.	n.a.	n.a.	0.56	n.a.	n.a.
Falch and Sandgren Massih (2011)	Malmö sample	Sweden	637	Raven's Matrices	Nonverbal	n.a.	3	0.0	15.0	0.62[f]	0.80	n.a.
Fischbach, Baudson, Preckel, Martin, and Brunner (2013)	MAGRIP sample	Luxembourg	1601	Performance Test System (German version)	Mixed	n.a.	6	49.7	11.9	0.51[f]	0.83	n.a.
Fischer, Schult, and Hell (2013)	Female sample	Germany	232	Intelligenz-Struktur-Test 2000 R	Mixed	n.a.	13	100.0	20.3	0.37	0.90	0.62
Fischer et al. (2013)	Male sample	Germany	220	Intelligenz-Struktur-Test 2000 R	Mixed	n.a.	13	0.0	n.a.	0.43	0.90	0.64
Flere, Krajnc, Klanjšek, Musil, and Kirbiš (2010)	Flere et al.'s sample	Slovenia	1,308	POGACIK Test	Nonverbal	n.a.	n.a.	48.9	17.0	0.48	0.84	n.a.
Frandsen (1950)	Frandsen's sample	USA	83	Wechsler-Bellevue; Henmon-Nelson	n.a.	n.a.	12	n.a.	17.0	0.60[f]	0.87	0.80
Freudenthaler et al. (2008)	Female sample	Austria	779	Intelligenz-Struktur-Analyse	Mixed	n.a.	8	100.0	14.5	0.55	0.84	n.a.
Freudenthaler et al. (2008)	Male sample	Austria	526	Intelligenz-Struktur-Analyse	Mixed	n.a.	8	0.0	14.5	0.53	0.84	n.a.
Freund and Holling (2011)	Matrices sample	Germany	646	MatrixDeveloper	Nonverbal	n.a.	10	34.0	15.7	0.36	0.74	n.a.
Freund, Holling, and Preckel (2007)	Freund et al.'s sample	Germany	1,135	Berlin Structure of Intelligence Test for Youth: Assessment of Talent and Giftedness	Mixed	n.a.	n.a.	46.2	14.5	0.48[f]	0.89	n.a.
Friedhoff (1955)	Female Boone sample	USA	48	Otis Quick Scoring Mental Ability Test; Science Research Associates Test of Primary Mental Abilities	n.a.	n.a.	7	100.0	n.a.	0.61[f]	0.88	0.86

Friedhoff (1955)	Male Boone sample	USA	59	Otis Quick Scoring Mental Ability Test; Science Research Associates Test of Primary Mental Abilities	n.a.	n.a.	7	0.0	n.a.	0.49[f]	0.88	0.96
Friedhoff (1955)	Female Mason City sample	USA	103	Otis Quick Scoring Mental Ability Testx; Science Research Associates Test of Primary Mental Abilities	n.a.	n.a.	7	100.0	n.a.	0.50[f]	0.88	0.90
Friedhoff (1955)	Male Mason City sample	USA	99	Otis Quick Scoring Mental Ability Test; Science Research Associates Test of Primary Mental Abilities	n.a.	n.a.	7	0.0	n.a.	0.51[f]	0.88	0.96
Furnham and Monsen (2009)	South East of England sample	Great Britain	265	Wonderlic Personnel Test; Baddeley Reasoning Test	Verbal	n.a.	10	41.3	15.6	0.35[f]	0.86	n.a.
Furnham, Monsen, and Ahmetoglu (2009)	Furnham et al.'s Sample	Great Britain	212	Wonderlic Personnel Test	Mixed	n.a.	10	58.0	15.8	0.49	0.90	1.00
Gagné and St. Père (2002)	Greater Montreal sample	Canada	156	Raven's Progressive Matrices; Otis–Lennon Mental Ability Test	n.a.	n.a.	8	100.0	13.5	0.46[f]	0.87	n.a.
Ghosh (1960)	1956 sample	India	120	Viva	n.a.	n.a.	3	0.0	n.a.	0.29	0.54	n.a.
Ghosh (1960)	1957 sample	India	120	Viva	n.a.	n.a.	3	0.0		0.39	0.54	n.a.
Gilles and Bailleux (2001)	Gilles and Bailleux's sample	France	122	Culture Fair Test: subtest 3; Ability Factor Battery: Numerical Addition subtest, Mental Rotation subtest	n.a.	n.a.	6	49.2	12.0	0.27[f]	0.88	n.a.
Gralewski and Karwowski (2012)	Creativity sample	Poland	589	Raven's Progressive Matrices	Nonverbal	n.a.	3	51.8	17.6	0.15	0.85	n.a.
Hartlage and Steele (1977)	Hartlage and Steele's sample	USA	36	Wechsler Intelligence Scale; Wechsler Intelligence Scale for Children — Revised	Mixed	n.a.	n.a.	32.0	8.0	0.59[f]	0.89	n.a.
Hartson (1939)	Ohio sample	USA	2,121	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.00	n.a.	n.a.
Haugwitz, Nesbit, and Sandmann (2010)	Biology sample	Germany	248	German Cognitive Ability Test — 4–12 — Revision	n.a.	Mathematics and Science	n.a.	56.0	13.9	0.17[f]	0.73	0.68
Heaven and Ciarrochi (2012)	Sydney sample	Australia	786	Standardized Australian test	Mixed	n.a.	10	53.2	n.a.	0.46	n.a.	n.a.
Hinkelman (1955)	Hinkelman's sample	USA	29	Kuhlmann–Anderson intelligence rating	n.a.	n.a.	n.a.	50.0	n.a.	0.57[f]	n.a.	n.a.
Hintsanen et al. (2012)	Helsinki sample	Finnland	309	Three standardized tests		Mathematics and Science	8	46.6	15.0	0.62[f]	0.74	n.a.
Hofer, Kuhnle, Kilian, and Fries (2012)	Bielefeld sample	Germany	697	Berlin Structure of Intelligence Tes: Reasoning subtest	Mixed	n.a.	8	52.0	13.4	0.34[f]	0.68	n.a.
Irwin, Engle, Yarbrough, Klein, and Townsend (1978)	Female sample	Guatemala	125	Adaptation of the 22-test INCAP Preschool Battery	n.a.	n.a.	n.a.	100.0	7.0	0.34[f]	n.a.	n.a.
Irwin et al. (1978)	Male sample	Guatemala	130	Adaptation of the 22-test INCAP Preschool Battery	n.a.	n.a.	n.a.	0.0	7.0	0.42[f]	n.a.	n.a.
Johnson and McGowan (1984)	Houston Parent–Child Development Center sample	USA	51	Stanford-Binet; McCarthy Scales of Childrens Ability; Wechsler Intelligence Scale for Children — Revised	n.a.	n.a.	n.a.	n.a.	n.a.	0.17[f]	0.90	n.a.
Johnson (1967)	Minnesota sample	USA	126	Lorge-Thorndike Intelligence Test	Mixed	n.a.	n.a.	n.a.	n.a.	0.63	n.a.	n.a.
Johnson, Deary, and Iacono (2009)	Twin sample	Germany	1,648	Wechsler Intelligence Scale for Children — Revised	Mixed	n.a.	n.a.	53.8	11.0	0.40	0.97	0.92
Jordan (1922)	Arkansas sample	USA	67	Otis; Army Alpha; Miller; Terman	n.a.	n.a.	n.a.	n.a.	n.a.	0.47[f]	0.94	n.a.
Kanderian (1970)	Kanderian's sample	Iraq	304	Cattell Culture Fair Intelligence Test; Modification of Thurstone Primary Mental Abilities	n.a.	n.a.	6	n.a.	n.a.	0.30[f]	0.86	n.a.
Karbach, Gottschling, Spengler, Hegewald, and Spinath (2013)	Cosmos sample	Germany	334	German Cognitive Ability Test — 4–12 — Revision (German version); Wechsler Intelligence Scale-III	n.a.	n.a.	n.a.	50.6	12.4	0.38	n.a.	n.a.
Kaufman et al. (2010)	Implicit sample	Great Britain	109	Raven's Advanced Progressive Matrices; Differential Aptitudes Test (The Psychological Corporation, 1995); Mental Rotation Test	n.a.	n.a.	10	69.3	n.a.	0.23[f]	0.79	1.19
Keehn and Prothro (1955)	Section secondaire sample	Lebanon	54	Cattell's Culture-Free Test; Raven's Progressive Matrices; Dominoe's Test D 48 (French version); number series test	Nonverbal	n.a.	10	n.a.	n.a.	0.35[f]	0.70	n.a.
Keehn and Prothro (1955)	Preparatory section 2 sample	Lebanon	48	Cattell's Culture-Free Test; Raven's Progressive Matrices; Dominoe's Test D 48 (French version); number series test	Nonverbal	n.a.	10	n.a.	n.a.	0.19[f]	0.85	n.a.
Keehn and Prothro (1955)	Preparatory section 3 sample	Lebanon	48	Cattell's Culture-Free Test; Raven's Progressive Matrices; Dominoe's Test D 48 (French version); number series test	Nonverbal	n.a.	11	n.a.	n.a.	0.16[f]	0.85	n.a.
Keehn and Prothro (1955)	Preparatory section 4 sample	Lebanon	56	Cattell's Culture-Free Test; Raven's Progressive Matrices; Dominoe's Test D 48 (French version); number series test	Nonverbal	n.a.	12	n.a.	n.a.	0.35[f]	0.85	n.a.
Kessels and Steinmayr (2013)	Macho sample	Germany	180	Intelligence-Structure-Test 2000R	Mixed	n.a.	11	45.6	16.4	0.30[f]	0.89	0.80
Kleitman and Moscrop (2010)	Kleitman and Moscrop's sample	Australia	184	Raven's Progressive Matrices	Nonverbal	n.a.	5	n.a.	10.4	0.35	0.88	0.88
Krumm, Lipnevich, Schmidt-Atzert, and Bühner (2012)	Krumm et al.'s sample	Germany	161	Intelligence-Structure-Test 2000R	Mixed	n.a.	n.a.	67.0	21.0	0.27[f]	0.76	n.a.
Kundu (1962)	Calcutta sample	India	58	Self-developed test including items from M- and L-form from the Terman–Merrill Intelligence Scale (1955)	n.a.	n.a.	8	0.0	n.a.	0.31[f]	n.a.	n.a.
Kundu (1975)	Female Science stream	India	50	Self-developed test including items from M- and L-form from the	n.a.	n.a.	10	100.0	n.a.	0.68	n.a.	0.65

(continued on next page)

Appendix 1. *(continued)*

Author (year)	Description of the sample	Country	n[a]	Intelligence test	Type of intelligence test	Subject domain	Grade level	Gender[b]	Age	r[c]	r_{xx}[d]	u_x[e]
	sample			Terman–Merrill Intelligence Scale (1955)								
Kundu (1975)	Male Science stream sample	India	50	Self-developed test including items from M- and L-form from the Terman–Merrill Intelligence Scale (1955)	n.a.	n.a.	10	0.0	n.a.	0.63	n.a.	0.75
Kundu (1975)	Female Humanities stream sample	India	50	Self-developed test including items from M- and L-form from the Terman–Merrill Intelligence Scale (1955)	n.a.	n.a.	10	100.0	n.a.	0.31	n.a.	0.62
Kundu (1975)	Male Humanities stream sample	India	50	Self-developed test including items from M- and L-form from the Terman–Merrill Intelligence Scale (1955)	n.a.	n.a.	10	0.0	n.a.	0.57	n.a.	0.69
Kuusinen and Leskinen (1988)	Kuusinen and Leskinen's sample	Finland	234	Illinois Test of Psycholiguistic Abilities	Verbal	n.a.	9	50.0	15.0	0.30[f]	0.95	n.a.
Kwall and Lackner (1966)	Kwall and Lackner's sample	USA	100	Otis Quick-Scoring Mental Abilities Test	n.a.	n.a.	4	50.0	n.a.	0.77	0.88	n.a.
Lemos et al. (2014)	Old sample	Portugal	1,101	Reasoning Tests Battery — Elementry School	n.a.	n.a.	11	n.a.	16.8	0.32[f]	0.74	1.00
Lemos et al. (2014)	Young sample	Portugal	1,695	Reasoning Tests Battery — Secondary School	n.a.	n.a.	8	n.a.	13.5	0.37[f]	0.74	1.00
Levpušček, Zupančič, and Sočan (2013)	Slovenia sample	Slovenia	400	Raven's Standard Progressive Matrices	Nonverbal	Mathematics and Science	9	51.9	13.5	0.51	0.93	0.86
Lindgren and De Almeida Guedes (1963)	Elementary school sample	Brazil	88	Two 60-item multiple-choice, paper-pencile intelligence tests	Verbal	n.a.	4	n.a.	10.7	0.61	0.66	n.a.
Lindgren and De Almeida Guedes (1963)	Secondary school sample	Brazil	55	Two 60-item multiple-choice, paper-pencile intelligence tests	Verbal	n.a.	7	n.a.	14.9	0.61	0.64	n.a.
Line and Glen (1935)	Second grade sample	Canada	61	National Intelligence Test	n.a.	n.a.	n.a.	n.a.	n.a.	0.57	n.a.	n.a.
Line and Glen (1935)	Junior third grade sample	Canada	129	National Intelligence Test	n.a.	n.a.	n.a.	n.a.	n.a.	0.46	n.a.	n.a.
Line and Glen (1935)	Senior third grade sample	Canada	119	National Intelligence Test	n.a.	n.a.	n.a.	n.a.	n.a.	0.39	n.a.	n.a.
Line and Glen (1935)	Junior forth grade sample	Canada	91	National Intelligence Test	n.a.	n.a.	n.a.	n.a.	n.a.	0.15	n.a.	n.a.
Line and Glen (1935)	Senior forth grade sample	Canada	124	National Intelligence Test	n.a.	n.a.	n.a.	n.a.	n.a.	0.47	n.a.	n.a.
Lu, Weber, Spinath, and Shi (2011)	German-Chinese sample	China	179	Culture Fair Intelligence Test	Nonverbal	n.a.	4	n.a.	n.a.	0.45[f]	0.78	0.96
Lubbers, Van Der Werf, Kuyper, and Hendriks (2010)	Dutch sample	Netherlands	9,776	Entry Test	n.a.	n.a.	7	52.0	13.0	0.25[f]	0.79	n.a.
Marcus (2000)	Marcus's sample	n.a.	210	Wonderlic Personnel Test (German version)	Verbal	n.a.	13	41.8	23.9	0.36	0.90	n.a.
Marcus, Wagner, Poole, Powell, and Carswell (2009)	Butchers sample	n.a.	50	Wonderlic Personnel Test (German version)	Verbal	n.a.	n.a.	n.a.	n.a.	0.39	0.90	0.56
Marcus et al. (2009)	Butchery shop assistant sample	n.a.	115	Wonderlic Personnel Test (German version)	Verbal	n.a.	n.a.	n.a.	n.a.	0.45	0.90	0.58
Marcus et al. (2009)	Bakery shop assistant sample	n.a.	99	Wonderlic Personnel Test (German version)	Verbal	n.a.	n.a.	n.a.	n.a.	0.21	0.90	0.54
Matlin and Mendelsohn (1965)	Matlin and Mendelsohn's sample	USA	68	Otis Quick-Scoring Mental Abilities Test	Mixed	n.a.	5	n.a.	n.a.	0.60	0.86	n.a.
Matthews, Marulis, and Williford (2014)	NICHD SECCYD sample	USA	923	Woodcock–Johnson Picture Vocabulary	Verbal	n.a.	5	n.a.	4.5	0.42	0.76	n.a.
Mayer (1958)	Delaware sample	US	100	Wechsler Intelligence Scale for Children	Mixed	n.a.	7	n.a.	n.a.	0.73	0.89	n.a.
Meili, Aebi, Heizmann, and Schoefer (1977)	Lucerne 1970 sample	Switzerland	185	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.34[f]	n.a.	n.a.
Meili et al. (1977)	Lucerne 1971 sample	Switzerland	153	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.32[f]	n.a.	n.a.
Meili et al. (1977)	Zurich 1971 sample	Switzerland	128	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.38[f]	n.a.	n.a.
Meili et al. (1977)	Winterthur 1971 sample	Switzerland	104	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.39[f]	n.a.	n.a.
Meili et al. (1977)	Winterthur 1972 sample	Switzerland	87	Self-developed test battery	n.a.	n.a.	n.a.	n.a.	n.a.	0.51[f]	n.a.	n.a.
Meili et al. (1977)	Basel 1971 sample I	Switzerland	175	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.39[f]	n.a.	n.a.
Meili et al. (1977)	Basel 1971 sample II	Switzerland	209	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.44[f]	n.a.	n.a.
Meili et al. (1977)	Basel 1972 sample I	Switzerland	139	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.45[f]	n.a.	n.a.
Meili et al. (1977)	Basel 1972 sample II	Switzerland	281	Self-developed test battery	n.a.	n.a.	n.a.	100.0	n.a.	0.38[f]	n.a.	n.a.
Millikin (1976)	South Texas sample	USA	306	Differential Aptitude Test	Mixed	n.a.	11	54.6	n.a.	0.68	n.a.	n.a.
Moenikia and Zahed-Babelan (2010)	Iran sample	Iran	1670	Raven's Intelligence Test	Nonverbal	Mathematics and Science	n.a.	44.2	n.a.	0.45	0.86	n.a.
Moore (1939)	Moore's sample	USA	174	Otis Group Intelligence Scale	Mixed	n.a.	6	n.a.	n.a.	0.69	n.a.	n.a.
Mõttus, Guljajev, Allik, Laidra, and Pullmann (2012)	Mõttus et al.'s sample	Estonia	1,159	Raven's Standard Progessive Matrices	Nonverbal	n.a.	n.a.	0.0	14.6	0.43	0.80	n.a.

Myburgh, Grobler, and Niehaus (1999)	Roodepoort sample	South Africa	656	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.56	n.a.	n.a.
Neuenschwander (2013)	PSB sample	Switzerland	187	Prüfsystem für Schul- und Bildungsberatung 6–13	n.a.	n.a.	6	49.7	n.a.	0.63[f]	0.90	1.12
Neuenschwander, Röthlisberger, Cimeli, and Roebers (2012)	Mixed-class-sample	Switzerland	459	Test of Nonverbal Intelligence — Third Edition	Nonverbal	n.a.	1	49.0	7.4	0.35[f]	n.a.	n.a.
Oates (1929)	Form IIa sample	Great Britain	35	Self-developed test battery	n.a.	n.a.	2	0.0	n.a.	0.34	n.a.	n.a.
Oates (1929)	Form IIb sample	Great Britain	35	Self-developed test battery	n.a.	n.a.	2	0.0	n.a.	0.32	n.a.	n.a.
Oates (1929)	Form IIIa sample	Great Britain	34	Self-developed test battery	n.a.	n.a.	3	0.0	n.a.	0.63	n.a.	n.a.
Oates (1929)	Form IIIb sample	Great Britain	33	Self-developed test battery	n.a.	n.a.	3	0.0	n.a.	0.21	n.a.	n.a.
Oates (1929)	Form IVa sample	Great Britain	34	Self-developed test battery	n.a.	n.a.	4	0.0	n.a.	0.34	n.a.	n.a.
Oates (1929)	Form IVb sample	Great Britain	33	Self-developed test battery	n.a.	n.a.	4	0.0	n.a.	0.15	n.a.	n.a.
Oates (1929)	Form R sample	Great Britain	16	Self-developed test battery	n.a.	n.a.	n.a.	0.0	n.a.	−0.13	n.a.	n.a.
Oates (1929)	Form Va sample	Great Britain	28	Self-developed test battery	n.a.	n.a.	5	0.0	n.a.	0.26	n.a.	n.a.
Oates (1929)	Form Vb sample	Great Britain	29	Self-developed test battery	n.a.	n.a.	5	0.0	n.a.	0.62	n.a.	n.a.
Peterson, Pihl, Higgins, Séguin, and Tremblay (2003)	Peterson et al.'s sample	Canada	148	Wechsler Intelligence Scale for Children — Revised	Mixed	n.a.	n.a.	0.0	14.0	0.46[f]	0.91	n.a.
Petrides, Chamorro-Premuzic, Frederickson, and Furnham (2005)	Female sample	Great Britain	297	Verbal Reasoning Test (Department of Assessment and Measurement at the National Foundation for Educational Research)	Verbal	n.a.	11	100.0	n.a.	0.85	0.97	1.01
Petrides et al. (2005)	Male sample	Great Britain	321	Verbal Reasoning Test (Department of Assessment and Measurement at the National Foundation for Educational Research)	Verbal	n.a.	11	0.0	n.a.	0.84	0.97	1.00
Phillipson and Phillipson (2012)	Hong-Kong sample	China	780	Raven's Progressive Matrices	Nonverbal	n.a.	n.a.	52.5	n.a.	0.28[f]	0.80	n.a.
Powell and Nettelbeck (2014)	Curiosity sample	Australia	160	Raven's Advanced Progressive Matrices-Short Form	Nonverbal	n.a.	12	68.8	19.4	0.49	0.75	0.61
Preckel and Brüll (2010)	Trier sample	Germany	678	German Cognitive Ability Test — 4–12 — Revisionx	n.a.	Mathematics and Science	5	52.1	10.4	0.43	0.92	n.a.
Preckel, Götz, and Frenzel (2010)	Boredom sample	Germany	144	German Cognitive Ability Test — 4–12 — Revision	n.a.	n.a.	9	43.0	14.8	0.62	0.93	0.90
Preckel et al. (2013)	Morning sample	Germany	237	Culture Fair Intelligence Test-20 (German adaptation)	Nonverbal	n.a.	10	46.7	15.6	0.44	0.93	1.41
Qualter, Gardner, Pope, Hutchinson, and Whiteley (2012)	Female sample	Great Britain	214	Cognitive Ability Test	n.a.	n.a.	11	100.0	11.2	0.43[f]	n.a.	n.a.
Qualter et al. (2012)	Male sample	Great Britain	199	Cognitive Ability Test	n.a.	n.a.	11	0.0	11.2	0.64[f]	n.a.	n.a.
Rauh (1977)	Rauh's sample	n.a.	42	Wechsler Preschool and Primary Scale of Intelligence; Hamburg-Wechsler-Intelligence tests for Children-1; Hamburg-Wechsler-Intelligence tests for Chrildren-2; Picture Test for 2. and 3. grade; Aufgaben zum Nachdenken 4+; Prüfsystem für Schul- und Bildungsberatung	n.a.	n.a.	n.a.	n.a.	n.a.	0.72[f]	0.95	0.91
Rindermann and Neubauer (2000)	Rindermann and Neubauer's sample	Germany	157	German Cognitive Ability Test; Advanced Progressive Matrices	n.a.	n.a.	10	n.a.	n.a.	0.44[f]	0.83	n.a.
Rindermann and Neubauer (2004)	Rindermann and Neubauer's sample	Germany	271	Raven's Advanced Progressive Matrices; German Cognitive Ability Test	n.a.	n.a.	10	n.a.	15.5	0.52	0.83	n.a.
Rogers (1933)	Public school sample	Canada	80	Pintner Group Test	Verbal	n.a.	4	n.a.	n.a.	0.62	0.97	n.a.
Rosander and Bäckström (2012)	Female sample	Sweden	245	Wonderlic Personnel Test	Mixed	n.a.	n.a.	100.0	17.2	0.49	0.90	0.79
Rosander and Bäckström (2012)	Male sample	Sweden	197	Wonderlic Personnel Test	Mixed	n.a.	n.a.	0.0	17.2	0.49	0.90	0.79
Rosander et al. (2011)	Lund sample	Sweden	297	Wonderlic Personnel Test	Mixed	n.a.	n.a.	53.0	17.6	0.38[f]	0.90	0.90
Sauer and Gamsjäger (1996)	Sauer and Gamsjäger's sample	Austria	651	Prüfsystem für die Schul- und Bildungsberatung	Mixed	n.a.	4	47.8	n.a.	0.56[f]	0.98	n.a.
Sauer and Gattringer (1986)	Salzburg sample	Austria	599	Prüfsystem für die Schul- und Bildungsberatung	Mixed	n.a.	4	50.0	n.a.	0.66	0.98	n.a.
Scarr, Caparlo, Ferndman, Tower, and Caplan (1983)	British Midlands sample	Great Britain	639	Young Nonreaders' IQ Test; Nonverbal (National Foundation for Educational Research)	Nonverbal	n.a.	12	n.a.	n.a.	0.73[f]	n.a.	n.a.

(continued on next page)

Appendix 1. *(continued)*

Author (year)	Description of the sample	Country	n[a]	Intelligence test	Type of intelligence test	Subject domain	Grade level	Gender[b]	Age	r[c]	r_{xx}[d]	u_x[e]
Schaefer and McDermott (1999)	Supplementary national sample	USA	420	Differential Ability Scales	n.a.	n.a.	n.a.	50.0	n.a.	0.40[f]	0.82	n.a.
Schneider, Grabner, and Paetsch (2009)	MPI Sample	Germany	195	German Cognitive Ability Test — 4–12 — Revision	Mixed	Mathematics and Science	6	51.2	11.3	0.60[f]	0.95	n.a.
Schult (2013)	Female sample I	Germany	41	Berlin Structure of Intelligence	n.a.	n.a.	n.a.	100.0	n.a.	0.36	0.82	n.a.
Schult (2013)	Female sample II	Germany	435	Berlin Structure of Intelligence	n.a.	n.a.	n.a.	100.0	n.a.	0.33	0.82	n.a.
Schult (2013)	Female sample III	Germany	1,033	Berlin Structure of Intelligence	n.a.	n.a.	n.a.	100.0	n.a.	0.51	0.73	n.a.
Schult (2013)	Male sample I	Germany	46	Berlin Structure of Intelligence	n.a.	n.a.	n.a.	0	n.a.	0.63	0.82	n.a.
Schult (2013)	Male sample II	Germany	293	Berlin Structure of Intelligence	n.a.	n.a.	n.a.	0	n.a.	0.35	0.82	n.a.
Schult (2013)	Male sample III	Germany	768	Berlin Structure of Intelligence	n.a.	n.a.	n.a.	0	n.a.	0.43	0.73	n.a.
Schwinger, Steinmayr, and Spinath (2009)	Gießen Sample	Germany	231	Culture Fair Test-3	Nonverbal	n.a.	12	60.2	16.8	0.23	0.78	0.62
Seitz (1971)	Würzburg sample	Germany	99	Analytischer Intelligenztest; Culture Fair Intelligence Test; Frankfurter Schultest[g]	n.a.	n.a.	6	52.5	12.3	0.18[f]	0.78	n.a.
Sharma, Sharma, and Sharma (2011)	Female sample	India	100	General Mental Ability Test; Standard Progressiv Matrices	n.a.	n.a.	11	100.0	n.a.	0.45[f]	0.91	n.a.
Sharma et al. (2011)	Male sample	India	100	General Mental Ability Test; Standard Progressiv Matrices	n.a.	n.a.	11	0.0	n.a.	0.52[f]	0.91	n.a.
Singh and Varma (1995)	Rural sample	India	200	Raven's Standard Progressive Matrices (1960)	Nonverbal	n.a.	11	50.0	n.a.	0.02	0.93	n.a.
Singh and Varma (1995)	Urban sample	India	200	Raven's Standard Progressive Matrices (1960)	Nonverbal	n.a.	11	50.0	n.a.	0.02	0.93	n.a.
Sonnleitner, Keller, Martin, and Brunner (2013)	Problem solving sample	Luxembourg	563	Intelligence-Structure-Test 2000R	Mixed	n.a.	10	48.7	17.4	0.51	0.73	n.a.
Spinath, Harald Freudenthaler, and Neubauer (2010)	Female sample	Austria	801	Intelligenz Struktur Analyse	n.a.	n.a.	8	100.0	13.7	0.41[f]	0.90	n.a.
Spinath et al. (2010)	Male sample	Austria	522	Intelligenz Struktur Analyse	n.a.	n.a.	8	0.0	13.7	0.47[f]	0.90	n.a.
Spinath, Spinath, and Plomin (2008)	Twin sample	Great Britain	4,492	Wechsler Intelligence Scale for Children-III; Cognitive Ability Test — Third Version	n.a.	n.a.	4	n.a.	n.a.	0.43[f]	n.a.	n.a.
St. John (1930)	Female sample	USA	455	Standford–Binet-Test; Dearborn General; Otis Primary; Myers Mental Measure	n.a.	n.a.	n.a.	100.0	n.a.	0.55	n.a.	n.a.
St. John (1930)	Male sample	USA	503	Standford–Binet-Test; Dearborn General; Otis Primary; Myers Mental Measure	n.a.	n.a.	n.a.	0.0	n.a.	0.44	n.a.	n.a.
Steinmayr, Dinger, and Spinath (2010)	Personality sample	Germany	548	Intelligence-Structure-Test-2000-R	n.a.	n.a.	12	61.2	17.0	0.27	0.90	0.84
Steinmayr and Meißner (2013)	Realschule sample	Germany	187	Culture Fair Test-20-Revision	Nonverbal	Mathematics and Science	8	48.8	13.6	0.26	0.86	0.62
Steinmayr and Meißner (2013)	Gymnasium sample	Germany	276	Culture Fair Test-20-Revision	Nonverbal	Mathematics and Science	8	48.8	13.6	0.39	0.86	0.62
Steinmayr and Spinath (2009)	Steinmayr and Spinath's sample	Germany	328	Intelligence-Structure-Test	n.a.	n.a.	12	59.7	16.9	0.35	0.90	n.a.
Steinmayr, Wirthwein, and Schöne (2014)	Math sample	Germany	301	Intelligence-Structure-Test-2000-R	Mixed	Mathematics and Science	12	49.2	17.5	0.41	0.92	0.53
Stejskal (1935)	Prague sample	Czech Republic	3,214	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.66	n.a.	n.a.
Süß (2001)	Süß's sample	Germany	137	Berlin Structure of Intelligence-Test	Mixed	n.a.	12	n.a.	n.a.	0.16[f]	0.82	n.a.
Svensson (1971)	Elementary school sample	Sweden	1,500	n.a.	n.a.	n.a.	n.a.	n.a.	13.0	0.66	0.90	n.a.
Svensson (1971)	Comprehensive school sample	Sweden	6,144	n.a.	n.a.	n.a.	n.a.	n.a.	13.0	0.69	0.90	n.a.
Svensson (1971)	Elementary school sample	Sweden	8,905	n.a.	n.a.	n.a.	n.a.	n.a.	13.0	0.67	0.86	n.a.
Nijenhuis, Resing, Tolboom, and Bleichrodt (2004)	7.8 year old Dutch sample	Netherlands	196	Amsterdam Intelligence Test for Children — Revised	n.a.	n.a.	n.a.	n.a.	7.8	0.54	n.a.	n.a.
Nijenhuis et al. (2004)	9.8 year old Dutch sample	Netherlands	204	Amsterdam Intelligence Test for Children — Revised	n.a.	n.a.	n.a.	n.a.	9.8	0.43	n.a.	n.a.
Nijenhuis et al. (2004)	7.8 year old immigrants sample	Netherlands	192	Amsterdam Intelligence Test for Children — Revised	n.a.	n.a.	n.a.	n.a.	7.8	0.50	n.a.	n.a.
Nijenhuis et al. (2004)	9.8 year old immigrants sample	Netherlands	184	Amsterdam Intelligence Test for Children — Revised	n.a.	n.a.	n.a.	n.a.	9.8	0.40	n.a.	n.a.
Telles Da Silva, Borges Osorio, and	Monozygotic twins sample	Brazil	79	Dominoes Test; Differential Aptitude Test		n.a.	n.a.	65.0	20.0	0.27[f]	0.70	n.a.

Salzano (1975)												
Telles Da Silva et al. (1975)	Dizygotic twins sample	Brazi	75	Dominoes Test; Differential Aptitude Test	n.a.	n.a.	n.a.	77.3	19.0	0.25[f]	0.70	n.a.
Tiedemann and Faber (1992)	Tiedemann and Faber's sample	Germany	96	Columbia Mental Maturity Scale (German version)	Nonverbal	n.a.	n.a.	77.3	7.3	0.19[f]	0.88	n.a.
Todt (1966)	Language stream sample	Germany	95	Wilde Intelligence Test	Mixed	n.a.	12	0.0	n.a.	0.33	0.81	n.a.
Todt (1966)	Science stream sample	Germany	113	Wilde Intelligence Test	Mixed	n.a.	12	0.0	n.a.	0.43	0.81	n.a.
Trautwein, Lüdtke, Roberts, Schnyder, and Niggli (2009)	Trautwein et al.'s sample I	Germany	571	German Cognitive Ability Test — 4–12 — Revision	Nonverbal	n.a.	8	51.4	14.7	0.16[f]	0.87	n.a.
Trautwein et al. (2009)	Trautwein et al.'s sample II	Germany	415	German Cognitive Ability Test — 4–12 — Revision	Nonverbal	Mathematics and Science	8	58.5	13.5	0.29[f]	0.90	n.a.
Trautwein et al. (2009)	Trautwein et al.'s sample III	Switzerland	1,535	German Cognitive Ability Test 4–13	Verbal	Languages	7	53.0	23.8	0.22	0.89	n.a.
Trost and Bickel (1979)	Female sample	n.a.	113	Test der akademischen Befähigung	Verbal	n.a.	13	100.0	n.a.	0.31	0.89	n.a.
Trost and Bickel (1979)	Male sample	n.a.	526	Test der akademischen Befähigung	Verbal	n.a.	13	0.0	n.a.	0.26	0.89	n.a.
Turney (1930)	Freshmen sample	USA	68	Army Group Examination Alpha; Pressey Senior Classification Test; Haggerty Intelligence Examination; Terman Group Test of Mental Ability; Miller Mental Ability Test	n.a.	n.a.	9	n.a.	n.a.	0.65	0.92	n.a.
Turney (1930)	Sophomore sample	USA	70	Army Group Examination Alpha; Haggerty Intelligence Examination; Terman Group Test of Mental Ability; Miller Mental Ability Test	n.a.	n.a.	10	n.a.	n.a.	0.56	0.93	n.a.
Turney (1930)	Juniors sample	USA	65	Army Group Examination Alpha; Pressey Senior Classification Test; Haggerty Intelligence Examination; Terman Group Test of Mental Ability; Miller Mental Ability	n.a.	n.a.	11	n.a.	n.a.	0.61	0.92	n.a.
Turney (1930)	Seniors sample	USA	65	Terman Group Test of Mental Ability; Otis Self-Administering Tests of Mental Ability, Higher Examination; Army Group Examination Alpha; Stanford–Binet Individual Examination	n.a.	n.a.	12	n.a.	n.a.	0.69	0.94	n.a.
Valenzuela (1971)	Spanish-American sample	USA	20	n.a.	n.a.	n.a.	11	n.a.	17.0	0.19	n.a.	n.a.
Valenzuela (1971)	Anglo-American sample	USA	20	n.a.	n.a.	n.a.	11	n.a.	17.0	0.48	n.a.	n.a.
Vecchione, Alessandri, and Marsicano (2014)	Female sample	Italy	54	Culture Fair Test	Nonverbal	n.a.	12	100.0	n.a.	0.09	0.78	n.a.
Vecchione et al. (2014)	Male sample	Italy	47	Culture Fair Test	Nonverbal	n.a.	12	0.0	n.a.	0.26	0.78	n.a.
Vock, Preckel, and Holling (2011)	BIS sample	Germany	1,135	Berlin Structure of Intelligence Test for Youth: Assessment of Talent and Giftedness	Mixed	n.a.	n.a.	46.2	14.5	0.40[f]	0.88	1.03
Vrdoljak and Velki (2012)	Croatia sample	Croatia	161	Cognitive-nonverbal test	Nonverbal	n.a.	8	47.7	13.1	0.45	0.93	n.a.
Watkins and Astilla (1980)	Watkins and Astilla's sample	Central Philippines	187	Otis–Lennon Mental Abilities Test	Mixed	n.a.	n.a.	100.0	n.a.	0.63	0.86	0.65
Watterson, Schuerger, and Melnyk (1976)	Watterson et al.'s sample	USA	163	Culture Fair Intelligence Test	Nonverbal	n.a.	n.a.	0.0	n.a.	0.31[f]	n.a.	n.a.
Weber, Lu, Shi, and Spinath (2013)	German-Chinese sample	Germany	320	Culture Fair Test-20-Revised	Nonverbal	n.a.	4	54.0	9.7	0.32[f]	n.a.	n.a.
Wellman (1957)	Iowa sample	USA	120	Otis Quick Scoring Mental Ability Tests; SRA Primary Mental Abilities Test	n.a.	n.a.	12	50.0	n.a.	0.74[f]	0.89	n.a.
Wright and Parker (1978)	Aboriginal sample	Australia	35	ACER Intermediate Test D (Australian Council for Educational Research, 1958)	Verbal	n.a.	8	51.4	13.8	0.27[f]	0.93	n.a.
Wright and Parker (1978)	Non-aboriginal sample	Australia	58	ACER Intermediate Test D (Australian Council for Educational Research, 1958)	Verbal	n.a.	8	44.8	13.4	0.52[f]	0.93	n.a.
Xin and Zhang (2009)	Chinese math sample	China	119	Revised Raven's Standard Progressive Matrices	Nonverbal	Mathematics and Science	5	52.1	11.5	0.45	0.85	n.a.
Zaunbauer, Retelsdorf, and Möller (2009)	Realschule Zaunbauer	Germany	292	German Cognitive Ability Test — 4–12 — Revision	Nonverbal	Languages	5	50.3	10.9	0.14[f]	0.83	n.a.
Zaunbauer et al. (2009)	Gymnasium Zaunbauer	Germany	418	German Cognitive Ability Test — 4–12 — Revision	Nonverbal	Languages	5	51.4	10.9	0.12[f]	0.83	n.a.
Ziegler, Knogler, and Bühner (2009)	Psychology sample	Germany	271	Intelligence-Structure-Test-2000-Revised	Mixed	n.a.	n.a.	82.0	20.0	0.22	0.92	0.87
Zuffianò et al. (2013)	Rom sample	Italy	170	Culture Fair Test	Nonverbal	n.a.	7	51.2	13.5	0.30[f]	0.77	0.71

Note. n.a. = not available.

[a] Sample size.

[b] Gender in % females.

[c] Correlation between intelligence test score and school grades within the primary study.

[d] Reliability of the intelligence test.

[e] Information on range restriction.

[f] Mean correlation for the sample calculated for the main meta-analysis.

[g] Datasets provided by members of the Deutsche Gesellschaft für Psychologie (DGPs).

References*

*Adkins, D. C. (1937). The efficiency of certain intelligence tests in predicting scholarship scores. *Journal of Educational Psychology*, *28*(2), 129–134.

*Agnoli, S., Mancini, G., Pozzoli, T., Baldaro, B., Russo, P. M., & Surcinelli, P. (2012). The interaction between emotional intelligence and cognitive ability in predicting scholastic performance in school-aged children. *Personality and Individual Differences*, *53*(5), 660–665.

*Al-Ghamri, S. J. (2012). A Study in the relationship of mental ability with age, sex, and school achievement of secondary school students in the Republic of Yemen (Unpublished Doctoral Dissertation). University of Pune, Pune.

*Allana, G. (2010). Is academic excellence in school children based on IQ: A review of private schools in Dubai regarding the gifted and talented (Unpublished Doctoral Dissertation). British University, Dubai.

Almquist, Y. (2011). Social isolation in the classroom and adult health: A longitudinal study of a 1953 cohort. *Advances in Life Course Research*, *16*(1), 1–12.

*Ames, V. (1943). Factors related to high-school achievement. *Journal of Educational Psychology*, *34*(4), 229–236.

Arthur, W., Bennett, W., & Huffcutt, A. I. (2001). *Conducting meta-analysis using SAS*. Mahwah, NJ: Lawrence Earlbaum Associates.

*Aswal, G. S. (2001). Intelligence as a correlate of achievement in mathematics across different levels of SES. *Psycholingua*, *31*(2), 127–130.

*Axelsson, G. B. (2009). Mathematical identity in women: The concept, its components and relationship to educative ability, achievement and family support. *International Journal of Lifelong Education*, *28*(3), 383–406.

*Barton, K., Dielman, T. E., & Cattell, R. B. (1971). The prediction of school grades from personality and IQ measures. *Personality*, *2*(4), 325–333.

Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, *132*(4), 355–429.

*Beckett, C., Castle, J., Rutter, M., & Sonuga-Barke, E. J. (2010). Institutional deprivation, specific cognitive functions, and scholastic achievement: English and Romanian Adoptee (ERA) study findings. *Monographs of the Society for Research in Child Development*, *75*(1), 125–142.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1947). *Differential aptitude tests*. New York: Psychological Corporation.

*Bipp, T., Steinmayr, R., & Spinath, B. (2012). A functional look at goal orientations: Their role for self-estimates of intelligence and performance. *Learning and Individual Differences*, *22*(3), 280–289.

*Blöschl, L. (1966). BTS, HAWIK, und schulisches Arbeitsverhalten. *Diagnostica*, *12*, 47–52.

*Blue, L. (2009). An investigation into the relationship between cognitive ability, standardized achievement, and grades in middle school (Unpublished doctoral dissertation). Temple University, Philadelphia.

*Bose, U. (1982). Scholastic achievement and its correlates. *Psychological Research Journal*, *6*(1), 33–38.

*Bouffard, T., Vezeau, C., Roy, M., & Lengelé, A. (2011). Stability of biases in self-evaluation and relations to well-being among elementary school children. *International Journal of Educational Research*, *50*(4), 221–229.

Boulanger, F. D. (1981). Ability and science learning: A quantitative synthesis. *Journal of Research in Science Teaching*, *18*, 113–121.

*Boulon-Díaz, F. (1992, August). The effects of intelligence, social class, early development and pre-school experience on school achievement of Puerto Rican children. *Paper presented at the Annual Convention of the American Psychological Association, Washington, DC*.

*Bowers, J. E. (1966). A study of the relationships among measures of productive thinking intelligence, and ninth-grade achievement (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.

Brähler, E., Holling, H., Leutner, D., & Petermann, F. (Eds.). (2002). *Brickenkamp Handbuch psychologischer und pädagogischer Tests, Band 1* (3rd ed.). Göttingen: Hogrefe.

*Bratko, D., Chamorro-Premuzic, T., & Saks, Z. (2006). Personality and school performance: Incremental validity of self- and peer-ratings over intelligence. *Personality and Individual Differences*, *41*(1), 131–142.

Brickenkamp, R. (Ed.). (1997). *Handbuch psychologischer und pädagogischer Tests* (2nd ed.). Göttingen: Hogrefe.

Brody, N. (1992). *Intelligence* (2nd ed.). San Diego, CA: Academic Press.

*Brucks, C. A. (1978). Über die Beziehungen zwischen Anspruchsniveau, Intelligenz und Schulleistung (unter Berücksichtigung der Stabilität und der Generalität des Anspruchsniveaus; unpublished doctoral dissertation). Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen-Nürnberg.

Brunner, M. (2006). Mathematische Schülerleistung: Struktur, Schulform-unterschiede und Validität (unpublished doctoral dissertation). Humboldt-Universität, Berlin.

*Burgert, R. H. (1937). The relation of school marks to intelligence in secondary schools. *Journal of Applied Psychology*, *19*(5), 606–614.

Calvin, C. M., Fernandes, C., Smith, P., Visscher, P. M., & Deary, I. J. (2010). Sex, intelligence and educational achievement in a national cohort of over 175,000 11-year-old schoolchildren in England. *Intelligence*, *38*(4), 424–432.

*Carter, H. D. (1959). Improving the prediction of school achievement by the use of the California Study Methods Survey. *Education Administration & Supervision*, *45*, 255–260.

Chamorro-Premuzic, T., & Furnham, A. (2005). *Personality and intellectual competence*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

*Chamorro-Premuzic, T., Quiroga, M. A., & Colom, R. (2009). Intellectual competence and academic performance: A Spanish study. *Learning and Individual Differences*, *19*(4), 486–491.

*Cocking, W. D., & Holy, T. C. (1927). Relation of intelligence scores to highschool and university marks. *Educational Research Bulletin*, *6*, 383–384.

Colom, R., & Flores-Mendoza, C. E. (2007). Intelligence predicts scholastic achievement irrespective of SES factors: Evidence from Brasil. *Intelligence*, *35*, 243–251.

*Cooper, M. (1974). Factor analysis of measures of aptitude, intelligence, personality, and performance in high school subjects. *The Journal of Experimental Education*, *42*(3), 7–10.

*Craig, H. M. (1988). The relationship of motives, subjective perceptions, occupational aspirations and cognitive ability to career salience and scholastic achievement for eighth-grade and tenth-grade students (unpublished doctoral dissertation). State University of New York, Buffalo.

*Daley, T. C., Whaley, S. E., Sigman, M. D., Guthrie, D., Neumann, C. G., & Bwibo, N. (2005). Background and classroom correlates of child achievement, cognitive, and behavioural outcomes in rural Kenyan schoolchildren. *International Journal of Behavioral Development*, *29*(5), 399–408.

*Dash, U. N., Mohanty, M., & Kar, B. C. (1989). Mental capacity and cognitive functioning: Relationship with school achievement. *Indian Journal of Current Psychological Research*, *4*(2), 103–110.

*Day, L., Hanson, K., Maltby, J., Proctor, C., & Wood, A. (2010). Hope uniquely predicts objective academic achievement above intelligence, personality, and previous academic achievement. *Journal of Research in Personality*, *44*(4), 550–553.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13–21.

*Di Fabio, A., & Palazzeschi, L. (2009). An in-depth look at scholastic success: Fluid intelligence, personality traits or emotional intelligence? *Personality and Individual Differences*, *46*(5), 581–585.

*Dodonova, Y. A., & Dodonov, Y. S. (2012). Processing speed and intelligence as predictors of school achievement: Mediation or unique contribution? *Intelligence*, *40*(2), 163–171.

*Downey, L. A., Lomas, J., Billings, C., Hansen, K., & Stough, C. (2013). Scholastic success: Fluid intelligence, personality, and emotional intelligence. *Canadian Journal of School Psychology*, *29*(1), 40–53.

*Dresel, M., Fasching, M., Steuer, G., & Berner, V. D. (2010). The role and the interplay of classroom goal structures, individual motivation and general intelligence for (excellent) school achievement in elementary school classrooms. *Talent Development & Excellence*, *2*(1), 63–81.

*Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What "No child left behind" leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, *104*(2), 439–451.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463.

*Edds, J. H., & McCall, W. M. (1933). Predicting the scholastic success of college freshmen. *Journal of Educational Research*, *27*, 127–130.

*Edminston, R. W., & Rhoades, B. J. (1959). Predicting achievement. *Journal of Educational Research*, *52*, 177–180.

*Falch, T., & Sandgren Massih, S. (2011). The effect of education on cognitive ability. *Economic Inquiry*, *49*(3), 838–856.

*Fischbach, A., Baudson, T. G., Preckel, F., Martin, R., & Brunner, M. (2013). Do teacher judgments of student intelligence predict life outcomes? *Learning and Individual Differences*, *27*, 109–119.

*Fischer, F., Schult, J., & Hell, B. (2013). Sex differences in secondary school success: why female students perform better. *European Journal of Psychology of Education*, *28*(2), 529–543.

Fleming, M. L., & Malone, M. R. (1983). The relationship of student characteristics and student performance in science as viewed by metaanalysis research. *Journal of Research in Science Teaching*, *20*, 481–495.

*Flere, S., Krajnc, M. T., Klanjšek, R., Musil, B., & Kirbiš, A. (2010). Cultural capital and intellectual ability as predictors of scholastic achievement: A study of Slovenian secondary school students. *British Journal of Sociology of Education*, *31*(1), 47–58.

*Frandsen, A. N. (1950). The Wechsler–Bellevue intelligence scale and high school achievement. *Journal of Applied Psychology*, *34*(6), 406–411.

*Freudenthaler, H. H., Spinath, B., & Neubauer, A. C. (2008). Predicting school achievement in boys and girls. *European Journal of Personality*, *22*(3), 231–245.

*Freund, P. A., & Holling, H. (2011). Retest effects in matrix test performance: Differential impact of predictors at different hierarchy levels in an educational setting. *Learning and Individual Differences*, *21*(5), 597–601.

*Freund, P. A., Holling, H., & Preckel, F. (2007). A multivariate, multilevel analysis of the relationship between cognitive abilities and scholastic achievement. *Journal of Individual Differences*, *28*(4), 188–197.

*Friedhoff, W. H. (1955). Relationships among various measures of socio-economic status, social class identification, intelligence, and school achievement. *Dissertation Abstracts*, *15*.

*Furnham, A., & Monsen, J. (2009). Personality traits and intelligence predict academic school grades. *Learning and Individual Differences*, *19*(1), 28–33.

*Furnham, A., Monsen, J., & Ahmetoglu, G. (2009). Typical intellectual engagement, Big Five personality traits, approaches to learning and cognitive ability predictors of academic performance. *British Journal of Educational Psychology*, *79*(4), 769–782.

Gaedike, A. -K. (1974). Determinanten der Schulleistung. In K. Heller (Ed.), *Leistungsbeurteilung in der Schule* (pp. 46–93). Heidelberg: Quelle & Meyer.

*Gagné, F., & St. Père, F. (2002). When IQ is controlled, does motivation still predict achievement? *Intelligence*, *30*(1), 71–100.

Gale, C. (2005). Height and intelligence. *International Journal of Epidemiology*, *34*(3), 678–679.

* *References marked with an asterisk are included in the meta-analysis.*

*Ghosh, S. N. (1960). Prediction of scholastic achievement through the selection tests used in a Calcutta school. *Indian Journal of Psychology*, *35*, 55–62.

*Gilles, P. -Y., & Bailleux, C. (2001). Personality traits and abilities as predictors of academic achievement. *European Journal of Psychology of Education*, *16*(1), 3–15.

Glass, G. V. (1977). Integrating findings: the meta-analysis of research. *Review of Research in Education*, *5*, 351–379.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Gottfredson, L. S. (2002). g: Highly general and highly practical. In R. J. Sternberg, & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 331–380). Mahwah, New Jersey: Lawrence Erlbaum Associates.

*Gralewski, J., & Karwowski, M. (2012). Creativity and school grades: A case from Poland. *Thinking Skills and Creativity*, *7*(3), 198–208.

Gustafsson, J. E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner, & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186–242). New York: Prentice Hall International.

*Hartlage, L. C., & Steele, C. T. (1977). WISC and WISC-R correlates of academic achievement. *Psychology in the Schools*, *14*(1), 15–18.

*Hartson, L. D. (1939). Relative value of school marks and intelligence tests as bases for rating secondary schools. *School & Society*, *49*, 354–356.

*Haugwitz, M., Nesbit, J. C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and Individual Differences*, *20*(5), 536–543.

Hauser, R. M., & Palloni, A. (2010). *Why do intelligent people live longer*. University of Wisconsin-Madison: Center for Demography and Ecology.

Health and Psychosocial Instruments. Retrieved from https://www.ebscohost.com/academic/health-and-psychosocial-instruments-hapi. Access on 7/24/2015.

*Heaven, P. C., & Ciarrochi, J. (2012). When IQ is not everything: Intelligence, personality and academic performance at school. *Personality and Individual Differences*, *53*(4), 518–522.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486.

*Hinkelman, E. A. (1955). Relationship of intelligence to elementary school achievement. *Education Administration & Supervision*, *41*, 176–179.

*Hintsanen, M., Alatupa, S., Jokela, M., Lipsanen, J., Hintsa, T., & Leino, M. (2012). Associations of temperament traits and mathematics grades in adolescents are dependent on the rater but independent of motivation and cognitive ability. *Learning and Individual Differences*, *22*(4), 490–497.

*Hofer, M., Kuhnle, C., Kilian, B., & Fries, S. (2012). Cognitive ability and personality variables as predictors of school grades and test scores in adolescents. *Learning and Instruction*, *22*(5), 368–375.

Hogrefe Testzentrale. Retrieved from http://www.testzentrale.de. Access on 7/24/2015.

Hunter, J. E., & Schmidt, F. L. (2002). Fixed effects vs. random effects meta-analysis models: implications for cumulative research knowledge. *International Journal of Selection and Assessment*, *8*(4), 275–292.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis. correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

*Irwin, M., Engle, P. L., Yarbrough, C., Klein, R. E., & Townsend, J. (1978). The relationship of prior ability and family characteristics to school attendance and school achievement in rural Guatemala. *Child Development*, *49*, 415–427.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

*Johnson, D. L., & McGowan, R. J. (1984). Comparison of three intelligence tests as predictors of academic achievement and classroom behaviors of Mexican-American children. *Journal of Psychoeducational Assessment*, *2*(4), 345–352.

*Johnson, J. O. (1967). The relationship between science achievement and selected student characteristics (unpublished doctoral dissertation). University of Minnesota, Minneapolis.

*Johnson, W., Deary, I. J., & Iacono, W. G. (2009). Genetic and environmental transactions underlying educational attainment. *Intelligence*, *37*(5), 466–478.

*Jordan, A. M. (1922). Correlations of four intelligence tests with grades. *Journal of Educational Psychology*, *13*(7), 419–429.

*Kanderian, S. S. (1970). Study of the relationship between school achievement and measures of intelligence and creativity for students in Iraq (unpublished doctoral dissertation). University of Southern California, Los Angeles.

*Karbach, J., Gottschling, J., Spengler, M., Hegewald, K., & Spinath, F. M. (2013). Parental involvement and general cognitive ability as predictors of domain-specific academic achievement in early adolescence. *Learning and Instruction*, *23*, 43–51.

*Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321–340.

*Keehn, J. D., & Prothro, E. T. (1955). Nonverbal tests as predictors of academic success in Lebanon. *Educational and Psychological Measurement*, *15*, 495–498.

Kersting, M. (2015, March). Abitur ohne Hochschulreife. *Forschung & Lehre*, *3*, 200–202.

*Kessels, U., & Steinmayr, R. (2013). Macho-man in school: Toward the role of gender role self-concepts and help seeking in school performance. *Learning and Individual Differences*, *23*, 234–240.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1974). *Illinois Test of psycholinguistic abilities*. Urbana, IL: University of Illinois Press.

*Kleitman, S., & Moscrop, T. (2010). Self-confidence and academic achievements in primary-school children: Their relationships and links to parental bonds, intelligence, age, and gender. In A. Efklides, & M. Plousia (Eds.), *Trends and prospects in metacognition research* (pp. 293–326). New York: Springer US.

*Krumm, S., Lipnevich, A. A., Schmidt-Atzert, L., & Bühner, M. (2012). Relational integration as a predictor of academic achievement. *Learning and Individual Differences*, *22*(6), 759–769.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, *86*, 148–161.

*Kundu, R. (1962). An investigation in the prediction of school success. *Indian Journal of Psychology*, *37*(2), 75–83.

*Kundu, R. (1975). An investigation into the relationship between abstract intelligence and attainments in different subjects of higher secondary course. *Journal of Psychological Researches*, *19*(2), 71–75.

*Kuusinen, J., & Leskinen, E. (1988). Latent structure analysis of longitudinal data on relations between intellectual abilities and school achievement. *Multivariate Behavioral Research*, *23*(1), 103–118.

*Kwall, D. S., & Lackner, F. M. (1966, September). *Ability, sociometric and parent–child relationship variables in the prediction of elementary school achievement. Paper presented at the Annual Convention of the American Psychological Association, New York City*.

Kyttälä, M., & Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *European Journal of Psychology of Education*, *23*, 77–94.

Laidra, K., Pillmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, *42*, 441–451.

*Lemos, G. C., Abad, F. J., Almeida, L. S., & Colom, R. (2014). Past and future academic experiences are related with present scholastic achievement when intelligence is controlled. *Learning and Individual Differences*, *32*, 148–155.

*Levpušček, M. P., Zupančič, M., & Sočan, G. (2013). Predicting achievement in mathematics in adolescent students. The role of individual and social factors. *The Journal of Early Adolescence*, *33*(4), 523–551.

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Cambridge, Massachusetts: Harvard University Press.

*Lindgren, H. C., & De Almeida Guedes, H. (1963). Social status, intelligence, and educational achievement among elementary and secondary students in Sao Paulo, Brazil. *The Journal of Social Psychology*, *60*(1), 9–14.

*Line, W., & Glen, J. S. (1935). Some relationships between intelligence and achievement in the public school. *Journal of Educational Research*, *28*, 582–588.

*Lu, L., Weber, H. S., Spinath, F. M., & Shi, J. (2011). Predicting school achievement from cognitive and non-cognitive variables in a Chinese sample of elementary school children. *Intelligence*, *39*(2), 130–140.

*Lubbers, M. J., Van Der Werf, M. P., Kuyper, H., & Hendriks, A. A. (2010). Does homework behavior mediate the relation between personality and academic performance? *Learning and Individual Differences*, *20*(3), 203–208.

*Marcus, B. (2000). *Kontraproduktives Verhalten im Betrieb*. Göttingen: Verlag für Angewandte Psychologie.

*Marcus, B., Wagner, U., Poole, A., Powell, D. M., & Carswell, J. (2009). The relationship of GMA to counterproductive work behavior revisited. *European Journal of Personality*, *23*(6), 489–507.

*Matlin, A. H., & Mendelsohn, F. A. (1965). The relationship between personality and achievement variables in the elementary school. *Journal of Educational Research*, *58*(10), 457–459.

*Matthews, J. S., Marulis, L. M., & Williford, A. P. (2014). Gender processes in school functioning and the mediating role of cognitive self-regulation. *Journal of Applied Developmental Psychology*, *35*(3), 128–137.

Mayer, R. E. (2000). Intelligence and education. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 519–533). New York: Cambridge University Press.

*Mayer, R. W. (1958). A study of the STEP Reading, SCAT and WISC tests, and school grades. *Reading Teacher*, *12*, 117–142.

McNatt, D. B. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, *85*, 314–322.

*Meili, R., Aebi, H. J., Heizmann, M. -L., & Schoefer, E. (1977). Intelligence and scholastic achievements in high schools of German speaking Switzerland. *Schweizerische Zeitschrift für Psychologie und ihre Anwendungen*, *36*(2), 77–99.

Mental Measurements Yearbooks. Retrieved from http://www.ebscohost.com/ACADEMIC/mental-measurements-yearbook. Access on 7/24/2015.

*Millikin, J. L. (1976). Some correlates of test-wiseness among high school students. *Dissertation Abstracts International*, *36*(8-A), 5155.

*Moenikia, M., & Zahed-Babelan, A. (2010). A study of simple and multiple relations between mathematics attitude, academic motivation and intelligence quotient with mathematics achievement. *Procedia-Social and Behavioral Sciences*, *2*(2), 1537–1542.

*Moore, C. C. (1939). The elementary school mark. *The Pedagogical Seminary and Journal of Genetic Psychology*, *54*, 285–294.

*Mõttus, R., Guljajev, J., Allik, J., Laidra, K., & Pullmann, H. (2012). Longitudinal associations of cognitive ability, personality traits and school grades with antisocial behaviour. *European Journal of Personality*, *26*(1), 56–62.

*Myburgh, C. P. H., Grobler, R. C., & Niehaus, L. (1999). Predictors of scholastic achievement: IQ, self-concept, time concept, and background characteristics. *South African Journal of Education*, *19*(3), 165–178.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., ... Urbina, S. (1996). Intelligence, knowns, and unknowns. *American Psychologist*, *51*, 77–101.

*Neuenschwander, M. (2013). Intelligenzverteilung in den Sekundarstufenniveaus des Kantons Basel-Land (BL). *Psychologie und Erziehung*, *39*(1), 58–66.

*Neuenschwander, R., Röthlisberger, M., Cimeli, P., & Roebers, C. M. (2012). How do different aspects of self-regulation predict successful adaptation to school? *Journal of Experimental Child Psychology*, *113*(3), 353–371.

*Nijenhuis, J. t., Resing, W., Tolboom, E., & Bleichrodt, N. (2004). Short-term memory as an additional predictor of school achievement for immigrant children? *Intelligence*, *32*(2), 203–213.

*Oates, D. W. (1929). The relation of temperament and intelligence to scholastic ability. *Forum of Education*, *7*, 171–185.

Orwin, R. G. (1983). A fail-safe N for effect size. *Journal of Educational Statistics*, *8*, 147–159.

Pearlman, K. (1982). The Bayesian approach to validity generalization: A systematic examination of the robustness of procedures and conclusions (unpublished doctoral dissertation). George Washington University, Washington.

*Peterson, J. B., Pihl, R. O., Higgins, D. M., Séguin, J. R., & Tremblay, R. E. (2003). Neuropsychological performance, IQ, personality, and grades in a longitudinal grade-school male sample. *Individual Differences Research*, *1*(3), 159–172.

*Petrides, K. V., Chamorro-Premuzic, T., Frederickson, N., & Furnham, A. (2005). Explaining individual differences in scholastic behaviour and achievement. *British Journal of Educational Psychology*, *75*(2), 239–255.

*Phillipson, S., & Phillipson, S. N. (2012). Children's cognitive ability and their academic achievement: the mediation effects of parental expectations. *Asia Pacific Education Review*, *13*(3), 495–508.

*Powell, C., & Nettelbeck, T. (2014). Intellectual curiosity may not incrementally predict academic success. *Personality and Individual Differences*, *64*, 7–11.

*Preckel, F., & Brüll, M. (2010). The benefit of being a big fish in a big pond: Contrast and assimilation effects on academic self-concept. *Learning and Individual Differences*, *20*(5), 522–531.

*Preckel, F., Götz, T., & Frenzel, A. (2010). Ability grouping of gifted students: Effects on academic self-concept and boredom. *British Journal of Educational Psychology*, *80*(3), 451–472.

*Preckel, F., Lipnevich, A. A., Boehme, K., Brandner, L., Georgi, K., Könen, T., ... Roberts, R. D. (2013). Morningness–eveningness and educational outcomes: the lark has an advantage over the owl at high school. *British Journal of Educational Psychology*, *83*(1), 114–134.

Primi, R., Ferrão, M. E., & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences*, *20*, 446–451.

*Qualter, P., Gardner, K. J., Pope, D. J., Hutchinson, J. M., & Whiteley, H. E. (2012). Ability emotional intelligence, trait emotional intelligence, and academic success in British secondary schools: A 5 year longitudinal study. *Learning and Individual Differences*, *22*(1), 83–91.

R Core Team (2015). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing (URL http://www.R-project.org).

*Rauh, H. (1977). Schulleistung und Übertrittsempfehlung am Ende des 4. Schuljahres. Längsschnittliche empirische Analyse eines komplexen Bega- bungsurteils. In K. Ingenkamp (Ed.), *Schüler- und Lehrerbeurteilung* (pp. 15–64). Weinheim: Beltz.

*Rindermann, H., & Neubauer, A. C. (2000). Informationsverarbeitungsge- schwindigkeit und Schulerfolg: Weisen basale Maße der Intelligenz prädiktive Validität auf? *Diagnostica*, *46*(1), 8–17.

*Rindermann, H., & Neubauer, A. C. (2004). Processing speed, intelligence, creativity, and school performance: Testing of causal hypotheses using structural equation models. *Intelligence*, *32*(6), 573–589.

*Rogers, K. H. (1933). 'Intelligence' and 'perseveration' related to school achievement. *Journal of Experimental Education*, *2*, 35–43.

Rojstaczer, S., & Healy, C. (2010, March). Grading in American colleges and universities. *Teachers College Record* (Retrieved from http://www.tcrecord.org).

*Rosander, P., & Bäckström, M. (2012). The unique contribution of learning approaches to academic performance, after controlling for IQ and personality: Are there gender differences? *Learning and Individual Differences*, *22*(6), 820–826.

*Rosander, P., Bäckström, M., & Stenberg, G. (2011). Personality traits and general intelligence as predictors of academic performance: A structural equation modelling approach. *Learning and Individual Differences*, *21*(5), 590–596.

Sauer, J. (2006). Prognose von Schulerfolg. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (pp. 584–595). Weinheim: Beltz PVU.

*Sauer, J., & Gamsjäger, E. (1996). *Ist Schulerfolg vorhersagbar? Die Determinanten der Grundschulleistung und ihr prognostischer Wert für den Sekundarschulerfolg.* Göttingen: Hogrefe.

*Sauer, J., & Gattringer, H. (1986). Zur Aufklärung der Schulleistung durch spezifische und gemeinsame Varianzanteile von Intelligenz und Motivation. In K. Daumenlang, & J. Sauer (Eds.), *Aspekte psychologischer Forschung. Festschrift zum 60. Geburtstag von Erwin Roth* (pp. 237–257). Göttingen: Hogrefe.

*Scarr, S., Caparlo, B. K., Ferndman, B. M., Tower, R. B., & Caplan, J. (1983). Developmental status and school achievements of minority and non-minority children from birth to 18 years in a British Midlands town. *British Journal of Developmental Psychology*, *1*(1), 31–48.

*Schaefer, B. A., & McDermott, P. A. (1999). Learning behavior and intelligence as explanations for children's scholastic achievement. *Journal of School Psychology*, *37*(3), 299–313.

Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science*, *14*(6), 317–320.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, *62*, 529–540.

Schmidt, F. L., & Hunter, J. E. (1999). Comparison of three meta-analysis methods revisited: An analysis of Johnson, Mullen, and Salas (1995). *Journal of Applied Psychology*, *84*, 144–148.

Schmidt, F. L., & Le, H. (2005). *Hunter and Schmidt metaanalysis programs (Version 1.1) [Computer software].* Iowa City, IA: Department of Management and Organizations, University of Iowa.

*Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in grades 5 and 6. *Journal of Educational Psychology*, *101*(2), 359–372.

*Schult, J. (2013). Once upon a time in the test: Sex differences in the prediction of academic achievement and job performance (Unpublished doctoral dissertation). Universität Konstanz, Konstanz.

*Schwinger, M., Steinmayr, R., & Spinath, B. (2009). How do motivational regulation strategies affect achievement: Mediated by effort management and moderated by intelligence. *Learning and Individual Differences*, *19*(4), 621–627.

*Seitz, W. (1971). The relationship between personality characteristics of public school pupils, their scholastic achievements, and intelligence test performance. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *18*(2), 307–336.

*Sharma, A., Sharma, P., & Sharma, D. (2011). Cognitive correlates of different academic subjects in school setting. *International Journal of Psychological Studies*, *3*(2), 142–148.

*Singh, R., & Varma, S. K. (1995). The effect of academic aspiration and intelligence on scholastic success of XI graders. *Indian Journal of Psychometry & Education*, *26*(1), 43–48.

Smith, E., Hay, P., Campbell, L., & Trollor, J. N. (2011). A review of the association between obesity and cognitive function across the lifespan: implications for novel approaches to prevention and treatment. *Obesity Reviews*, *12*(9), 740–755.

*Sonnleitner, P., Keller, U., Martin, R., & Brunner, M. (2013). Students' complex problem-solving abilities: Their structure and relations to reasoning ability and educational success. *Intelligence*, *41*(5), 289–305.

*Spinath, B., Harald Freudenthaler, H., & Neubauer, A. C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Personality and Individual Differences*, *48*(4), 481–486.

*Spinath, F. M., Spinath, B., & Plomin, R. (2008). The nature and nurture of intelligence and motivation in the origins of sex differences in elementary school achievement. *European Journal of Personality*, *22*(3), 211–229.

*St. John, C. W. (1930). *Educational achievement in relation to intelligence.* Oxford: Harvard University Press.

Steinkamp, M. W., & Maehr, M. L. (1983). Affect, ability, and science achievement: A quantitative synthesis of correlational research. *Review of Educational Research*, *53*, 369–396.

*Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80–90.

*Steinmayr, R., & Meißner, A. (2013). Zur Bedeutung der Intelligenz und des Fähigkeitsselbstkonzeptes bei der Vorhersage von Leistungstests und Noten in Mathematik. *Zeitschrift für Pädagogische Psychologie*, *27*(4), 273–282.

*Steinmayr, R., Dinger, F. C., & Spinath, B. (2010). Parents' education and children's achievement: the role of personality. *European Journal of Personality*, *24*(6), 535–550.

*Steinmayr, R., Wirthwein, L., & Schöne, C. (2014). Gender and numerical intelligence: Does motivation matter? *Learning and Individual Differences*, *32*, 140–147.

*Stejskal, C. (1935). Selecting pupils for secondary schools. *Paper presented at the Conférence Internationale de Psychotechnique, Prague*.

Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, *47*, 1–41.

*Süß, H. -M. (2001). Prädiktive Validität der Intelligenz im schulischen und außerschulischen Bereich. In E. Stern, & J. Guthke (Eds.), *Perspektiven der Intelligenzforschung* (pp. 109–135). Lengerich: Pabst.

*Svensson, A. (1971). *Relative achievement: School performance in relation to intelligence, sex and home Environment.* Oxford: Almqvist & Wiksell.

Taub, G. E., Keith, T. Z., Floyd, R. G., & Mcgrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly*, *23*, 187–198.

*Telles Da Silva, B. T., Borges Osorio, M. R., & Salzano, F. M. (1975). School achievement, intelligence, and personality in twins. *Acta Geneticae Medicae et Gemellologiae*, *24*, 213–219.

Tent, L. (2006). Zensuren. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (pp. 873–880). Weinheim: Beltz PVU.

*Tiedemann, J., & Faber, G. (1992). Preschoolers' maternal support and cognitive competencies as predictors of elementary achievement. *Journal of Educational Research*, *85*(6), 348–354.

*Todt, E. (1966). Untersuchungen zur Vorhersage von Schulnoten. *Psychologische Forschung*, *29*, 32–51.

*Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology*, *97*(6), 1115–1128.

*Trost, G., & Bickel, H. (1979). *Studierfähigkeit und Studienerfolg.* München: Minerva Publikation.

*Turney, A. H. (1930). *Factors other than intelligence that affect success in high school.* Oxford: University of Minnesota Press.

*Valenzuela, A. M. (1971). The relationships between self-concept, intelligence, socioeconomic status and school Achievement among Spanish-American children in OMAHA (unpublished diploma thesis). University of Nebraska, Omaha.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48 (URL http://www.jstatsoft.org/v36/i03/).

*Vecchione, M., Alessandri, G., & Marsicano, G. (2014). Academic motivation predicts educational attainment: Does gender make a difference? *Learning and Individual Differences*, *32*, 124–131.

*Vock, M., Preckel, F., & Holling, H. (2011). Mental abilities and school achievement: A test of a mediation hypothesis. *Intelligence*, *39*(5), 357–369.

*Vrdoljak, G., & Velki, T. (2012). Metacognition and intelligence as predictors of academic success. *Hrvatski Časopis za Odgoj i Obrazovanje*, *14*(4), 799–815.

*Watkins, D., & Astilla, E. (1980). Relationship between field independence, intelligence, and school achievement for Filipino girls. *Perceptual and Motor Skills*, *51*(2), 593–594.

*Watterson, D. G., Schuerger, J. M., & Melnyk, G. I. (1976). The addition of personality measures to ability measures for predicting and understanding school achievement. *Multivariate Experimental Clinical Research*, *2*(3), 113–122.

*Weber, H. S., Lu, L., Shi, J., & Spinath, F. M. (2013). The roles of cognitive and motivational predictors in explaining school achievement in elementary school. *Learning and Individual Differences*, *25*, 85–92.

*Wellman, F. E. (1957). Differential prediction of high school achievement using single score and multiple factor tests of mental maturity. *Personnel & Guidance Journal*, *35*, 512–517.

Whitener, E. M. (1990). Confusion on confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, *75*, 315–321.

*Wright, M. M., & Parker, J. L. (1978). The relationship of intelligence, self-concept and locus of control to school achievement for aboriginal and non-aboriginal children. *Exceptional Child*, *25*(3), 167–179.

*Xin, Z., & Zhang, L. (2009). Cognitive holding power, fluid intelligence, and mathematical achievement as predictors of children's realistic problem solving. *Learning and Individual Differences*, *19*(1), 124–129.

*Zaunbauer, A. M., Retelsdorf, J., & Möller, J. (2009). Die Vorhersage von Englischleistungen am Anfang der Sekundarstufe. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *41*(3), 153–164.

*Ziegler, M., Knogler, M., & Bühner, M. (2009). Conscientiousness, achievement striving, and intelligence as performance predictors in a sample of German psychology students: Always a linear relationship? *Learning and Individual Differences*, *19*(2), 288–292.

*Zuffianò, A., Alessandri, G., Gerbino, M., Luengo Kanacri, B. P., Di Giunta, L., Milioni, M., & Caprara, G. V. (2013). Academic achievement: The unique contribution of self-efficacy beliefs in self-regulated learning beyond intelligence, personality traits, and self-esteem. *Learning and Individual Differences*, *23*, 158–162.