

Siddiq, Fazilat & Scherer, Ronny

Is there a gender gap? A meta-analysis of the gender differences in students' ICT literacy

Educational Research Review, 27, 205-217.

DOI: [10.1016/j.edurev.2019.03.007](https://doi.org/10.1016/j.edurev.2019.03.007)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Accepted Manuscript

Is there a gender gap? A meta-analysis of the gender differences in students' ICT literacy

Fazilat Siddiq, Ronny Scherer

PII: S1747-938X(18)30290-2

DOI: <https://doi.org/10.1016/j.edurev.2019.03.007>

Reference: EDUREV 275

To appear in: *Educational Research Review*

Received Date: 11 June 2018

Accepted Date: 24 March 2019

Please cite this article as: Siddiq, F., Scherer, R., Is There a Gender Gap? A Meta-Analysis of the Gender Differences in Students' ICT Literacy, *Educational Research Review*, <https://doi.org/10.1016/j.edurev.2019.03.007>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Is There a Gender Gap?

A Meta-Analysis of the Gender Differences in Students' ICT Literacy

Fazilat Siddiq

Nordic Institute for Studies in Innovation, Research and Education (NIFU), Norway

Ronny Scherer

University of Oslo, Norway

Author Note

Fazilat Siddiq, Nordic Institute for Studies in Innovation, Research and Education (NIFU), Oslo, Norway; Ronny Scherer, Department of Teacher Education and School Research, Faculty of Educational Sciences, University of Oslo.

Correspondence concerning this article should be addressed to Fazilat Siddiq, Nordic Institute for Studies in Innovation, Research and Education (NIFU), Postboks 2815 Tøyen, 0608 Oslo, Norway. Tel: +47 92440588. E-Mail: fazilat.siddiq@nifu.no; fazilatu@gmail.com.

Abstract

The study of gender differences in academic achievement has been one of the core topics in education, especially because it may uncover possible gaps and inequalities in certain domains. Whereas these differences have largely been examined in traditional domains, such as mathematics, reading, and science, the existing body of empirical studies in the domain of ICT literacy is considerably smaller, yet abounds in diverse findings. One of the persistent findings however is that boys consider their ICT literacy to be higher than that of girls. This meta-analysis tests whether the same pattern holds for students' actual performance on ICT literacy tasks, as measured by performance-based assessments. In total, 69 effect sizes were extracted from 23 empirical studies using a random-effects model. Overall, the gender differences in ICT literacy were significant, positive, and favored girls ($g = +0.13$, 95 % CI = [0.06, 0.16]). This effect varied between studies, and moderation analyses indicated that the grade level students were taught at moderated its magnitude—effect sizes were larger in primary school as compared to secondary school. In conclusion, our findings contrast those obtained from previous meta-analyses that were based on self-reported ICT literacy and suggest that the ICT gender gap may not be as severe as it had been claimed to be.

Keywords. Gender differences; ICT literacy; Meta-analysis; Performance-based assessment; Twenty-first century skills

Is There a Gender Gap?

A Meta-Analysis of the Gender differences in Students' ICT Literacy

1 Introduction

Technological advancements and the increased availability of information and communication technology (ICT) resources have changed traditional learning environments and necessitated curricular reforms. Educational systems around the world have responded to these developments by including skills that are needed to solve problems in technology-rich environments and to become reflective and responsible ICT users in their national curricula (Balanskat, 2009; Pellegrino & Hilton, 2012). These skills are often labelled as “ICT literacy”, “digital competence”, or “digital skills” (Siddiq et al., 2016) and are often used interchangeably (Gellardo-Echenique, de Oliveira, Marques-Molias, & Esteve-Mon, 2015; Siddiq, Hatlevik, Olsen, Throndsen, & Scherer, 2016; Voogt & Roblin, 2012).

Over the last years, it has been a common understanding that male students hold more positive attitudes toward technology and technology use, use ICT more actively, have higher ICT self-efficacy and hence, perform better than their female peers (Jackson, Zhao, Koleniclll, Fitzgerald, Harold & Von Eye, 2008). A recently published meta-analysis on gender and attitudes toward technology partly confirms this view by revealing that there is a small but significant positive effect towards boys, suggesting that boys have higher ICT self-efficacy and hold more favorable attitudes toward technology than girls (Cai, Fan & Du, 2017). Although this finding concurs with current expectations on the direction of a possible gender gap in technology-related domains, it is based solely on self-report measures. Even though self-reports of one's competence might be predictive of performance to a certain extent, they do not represent actual performance (Honicke & Broadbent, 2016).

Research on gender differences in students' ICT literacy using performance-based assessments abounds with conflicting findings: While some studies identified differences in favor of girls (e.g., Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014), others have reported differences in favor of boys (e.g., Calvani, Fini, Ranieri, & Picci, 2012). Moreover, some studies could not identify any gender differences in students' ICT literacy scores (e.g., Claro et al., 2012). These inconsistent findings warrant in-depth examinations of the relations between students' ICT literacy achievement and gender in order to understand inequalities in ICT literacy. To our knowledge, this issue of inconsistent research findings concerning possible gender differences in students' ICT literacy performance has not yet been addressed. To contribute to filling this gap and to bringing forward up-to-date knowledge, this study presents a meta-analysis in which we investigate whether gender differences in ICT literacy exist, which direction and magnitude these differences have, and what might explain possible between-study variation.

2 Theoretical Framework

ICT competences are considered vital for social interaction, civic participation, information retrieval and processing, academic performance, and professional success (Pagani, Argentin, Gui, & Stanca, 2016; Zhong, 2011). Their inclusion in education has therefore been emphasized as the growing number of assessment of ICT literacy across several countries and school systems testify (Siddiq et al., 2016). Next to the term "ICT literacy", a vast number of concepts (e.g., digital competence, computer literacy, ICT fluency, technological literacy, Internet skills, information literacy, media literacy) have been used in the literature to describe the knowledge, skills, and attitudes related to ICT (Ala-Mutka, 2011; Law, Lee, & Yuen, 2009). Given this diversity, efforts were made to define these concepts and identify their similarities and differences (Lankshear & Knobel, 2008). Many researchers concluded

that most of the terms are interchangeable and largely reflect the same content (Law et al., 2009; Sjøby, 2013). In this meta-analysis, we will use the term ICT literacy.

2.1 ICT Literacy

ICT literacy can be defined as “the interest, attitude, and ability of individuals to appropriately use digital technology and communication tools to access, manage, integrate, and evaluate information; construct new knowledge; and communicate with others in order to participate effectively in society” (Lennon et al., 2003, p. 8). This definition is in line with several other definitions (Educational Testing Service [ETS], 2007; Markauskaite, 2006; MCEETYA, 2007) and resonates with that of “digital competence” as an educational goal to foster the confident and critical use of ICT for fully participating in the knowledge society (Ferrari, 2013). These definitions are quite general and do not specify what is meant by being ICT literate or which specific skills, attitudes, and competences students should attain. The theoretical frameworks accompanying such concepts, however, are more specific as they outline both the content domains and the skills that underlie ICT literacy.

2.2 Frameworks of ICT Literacy

The existing frameworks of ICT literacy outline the knowledge and skills students need in order to become digitally literate. Although these frameworks are diverse and structured differently, they converge to a great extent to a set of core knowledge domains and skills of ICT literacy (Voogt & Roblin, 2012). The so-called DIGCOMP (Developing and Understanding Digital Competence in Europe) framework, which was initially developed by the European Commission (Ferrari, 2013) and revised in a recently published systematic review (Siddiq et al., 2016), describes these domains and skills: The revised DIGCOMP framework postulates six areas: *Information*, *Communication*, *Content Creation*, *Safety*, *Problem Solving*, and *Technical Operational Skills*, and each area further consists of several competences. For instance, *Information* includes competences such as search for- and use of

information, storing and retrieving information, including evaluation of information. *Content Creation* includes the knowledge and skills to develop new ideas or content and redefine existing ones. Moreover, programming and acknowledging copyright and licenses are also included in this competence areas. *Communication* is related to interaction with others, both synchronously and asynchronously, and includes collaboration, a competence needed to participate and contribute to teams using ICT, and to participate in and create social networks. *Safety* includes competences related to protecting personal data, health, devices and the environment. Moreover, netiquette is an important aspect of this area. *Problem solving* includes innovating and creatively using technology and solving problems individually or collaboratively. At last, *Technical Operational Skills* include generic ICT skills, such as operating devices and applications and identifying and solving technical problems.

Furthermore, to facilitate the present meta-analysis, the test content was recorded by further dividing these competence areas along with the task descriptions into two main groups: (1) *applied skills*, which consist of the skills that require the test-taker to apply knowledge to solve a problem or come to a solution by taking several actions (i.e., problem solving, communication and technological skills), and (2) *other skills*, which consist of competence or knowledge the students have and was shown by providing answers to tasks without taking further actions (i.e., knowledge-lean and theoretical skills). Thus, in the present meta-analysis, only studies will be included that conceptualized ICT literacy by at least one of these two knowledge domains and skills.

2.3 Assessment of ICT Literacy

Over the last two decades, assessment of ICT literacy and related concepts have been developed and conducted in several educational systems. While some assessment projects were initiated by national educational authorities (e.g., Australia, South Korea), others were initiated and conducted by researchers (e.g., Aesaert, van Nijlen, Vanderlinde, & van Braak,

2014; Siddiq, Gochyyev, & Wilson, 2017). Acknowledging the different initiatives and the resultant diversity in the assessments and their foci on certain knowledge domains and skills, Siddiq et al. (2016) reviewed ICT literacy assessments in K-12 education in order to identify their differences and commonalities. This systematic review summarized the contextual information about the assessments (e.g., country, age group, sample size, level of authenticity), information about the knowledge domains and skills the tests measured, and the reported test quality. Siddiq et al. (2016) conclude that most tests measure the knowledge and skills in the area of *Information*, whereas few tests measure knowledge and skills in the areas of *Communication*, *Collaboration*, *Safety*, and *Problem solving*. Moreover, their review revealed that the reporting of the reliability and validity of the assessments in the primary studies was insufficient. Even though the review provides an overview of the past research on ICT literacy assessment, it neither addresses the factors determining students' ICT literacy performance nor the equity and diversity issues, such as gender differences in performance.

2.4 Gender and ICT literacy Performance

Previous studies revealed gender differences in a variety of ICT-related constructs, such as the attitudes toward ICT (Cai et al., 2017; Pamuk & Peker, 2009), interest and self-efficacy in using ICT (Sáinz & Eccles, 2012), and ICT use in general (Vekiri & Chronaki, 2008; Volman, van Eck, Heemskerck, & Kuiper, 2005). A meta-analysis based on studies using self-reports of ICT literacy and attitudes found that there was a small but significant positive effect favoring boys. In other words, boys held more favorable attitudes toward technology and considered themselves more competent than girls did (Cai et al., 2017). This meta-analysis, however, did not include any performance assessment, leaving possible gender gaps in students' ICT literacy untouched. Despite the tendency to use self-reported ICT literacy as a proxy for performance and the practice to design school interventions that meet possible gender gaps on the basis of gender differences in self-reports (Meri-Tuulia, Antero,

& Suvi-Sadetta, 2017), we believe that it is time to examine whether the gender differences in motivational constructs and self-perceptions also apply to actual performance. Self-reported skills only provide rough indicators of performance (Honicke & Broadbent, 2016).

Gender differences in students' ICT literacy occurred in many studies; yet, these differences, either significant or insignificant, were inconsistent across studies (Kim, Kil, & Shin, 2014). Many studies reported that female students scored significantly higher than male students (e.g., Aesaert & van Braak, 2015; Baek et al., 2009; Hohlfeld, Ritzhaupt, & Barron, 2013; Kim et al., 2014; MCEETYA, 2007)—however, several studies found the opposite direction (e.g., Hakkarainen et al., 2000; Kuhlemeier & Hemker, 2007; Volman et al., 2005) or reported insignificant gender differences (Siddiq et al., 2017). It should be noted that this diversity in the findings of primary, empirical studies have not yet been explained by study, sample, or publication characteristics. This observation motivates the current meta-analysis to quantitatively synthesize the gender differences in ICT literacy across studies and to investigate the potential factors explaining the inconsistencies.

2.5 The Present Meta-Analysis

In this meta-analysis, we synthesize the existing body of research that examines the effects of gender on students' ICT literacy in K-12 education. To our best knowledge, a synthesis of this kind is still lacking in the existing body of research—this is surprising because such a synthesis could provide more detailed insights into the magnitude and direction of a digital divide that otherwise would neither be available nor obvious based on individual studies (Fan & Chen, 2001). We address the following research questions:

1. Do gender differences in performance measures of ICT literacy, as reported by empirical studies, exist, and to what extent do they vary between studies? (*Overall effect size and between-study variation*)

2. To what extent can study, sample, and publication characteristics explain possible variation in these gender differences between the primary studies in the literature?

(Moderator effects)

3 Methods

3.1 Literature Search and Coding

This meta-analysis is based on Siddiq et al.'s (2016) recent systematic review of performance-based ICT literacy assessments and extends it to a quantitative meta-analysis of gender differences. We updated the literature search that was conducted for studies between November 2014 and August 2017 by applying the same search procedures and selection criteria as they were specified in the systematic review (i.e., the study is concerned with ICT literacy or an equivalent term, reports results from a performance-based assessment conducted on 1-12 education). Five studies were added which were published after November 2014 (Aesaert & van Braak, 2015; Fraillon et al., 2014; Hatlevik, Ottestad, & Throndsen, 2015; Hatlevik, Scherer, & Christophersen, 2017; Siddiq et al., 2017). Relevant information, such as the study, sample, and publication characteristics were extracted and coded. In total, $m = 23$ studies were selected for further meta-analytic inquiries, 22 of which were single studies and one of which was the large-scale International Computer and Information Literacy study (ICILS) that was conducted in 21 countries in 2013. These studies were included because they fit the additional inclusion criteria set for this meta-analysis (i.e., that the study provides sufficient statistical information on gender differences in ICT literacy scores to calculate effect sizes).

Coding of the studies. To further investigate what might explain the inconsistent gender differences across the primary studies, we coded the following sample, study, and publication characteristics: (1) Sample size and number of boys and girls; (2) Country (i.e., which country the data was collected from, grouped as America, Asia, Australia and Europe);

(3) Educational level (i.e., primary or secondary level); (4) Year of publication; (5) Sampling procedure (i.e., convenience sample or randomized and/or stratified sample); (6) Publication type: peer-reviewed journal paper or grey literature (e.g., unpublished dissertation or report); (7) Test characteristics: Reliability of the measure(s), Fairness of the test across gender (e.g., whether invariance testing/differential item functioning was conducted), Test content (i.e., the skills/competences the test aims at measuring, coded as applied skills (i.e., problem solving, communication and technological skills) or other skills (i.e., knowledge-lean and theoretical skills), and Task mode (i.e., whether the tasks in the test were interactive [e.g., simulations] or static [e.g., multiple-choice questions]). Although the competencies the tests measured were originally coded according to the DIGCOMP framework as described in section 2.2 (see also Siddiq et al., 2016), we categorized them into only two groups, namely applied or theoretical skills. This decision was based on the observation from Siddiq et al.'s (2016) systematic review that some of the knowledge and skill areas were scarcely measured.

3.2 Statistical Analyses

Effect size calculations. The primary studies reported gender differences in digital competence in several ways by providing either mean scores and standard deviations or derived statistics, including Cohen's d as an effect size, t -values, and F -values. We transformed all of these statistics into Hedges' g —an effect size that represents a standardized mean difference next to Cohen's d , but provides a more precise effect size measure if sample sizes are small or dissimilar (Lipsey & Wilson, 2001). The extracted effect sizes fulfilled the independence assumption—although some studies reported multiple effect sizes, these effects were obtained from different samples (e.g., of different countries, gender, or age groups).

If the primary studies reported the means and standard deviations of digital performance measures, Hedges' g was calculated from the standardized mean differences ES as follows (Borenstein et al., 2009):

$$ES = \frac{\bar{X}_G - \bar{X}_B}{SD_{Pooled}}$$

where \bar{X}_G and \bar{X}_B represent the mean scores of girls (G) and boys (B), respectively, and SD_{Pooled} the pooled within-groups standard deviation. The latter accounts for possible differences in sample sizes, next to differences in standard deviations:

$$SD_{Pooled} = \sqrt{\frac{(N_G - 1)SD_G^2 + (N_B - 1)SD_B^2}{N_G + N_B - 2}}$$

The standardized mean differences ES represent Cohen's d . Finally, we transformed the resultant effect size ES into Hedges' g , and calculated the corresponding variance v_g and standard error SE_g (with $df = N_G + N_B - 2$; Borenstein et al., 2009):

$$g = \left(1 - \frac{3}{4df - 1}\right) \cdot ES$$

$$v_g = \left(1 - \frac{3}{4df - 1}\right)^2 \cdot \left(\frac{N_G + N_B}{N_G \cdot N_B} + \frac{ES^2}{2(N_G + N_B)}\right)$$

$$SE_g = \sqrt{v_g}$$

For studies that reported t - or F -values, the standardized mean differences ES could be calculated as follows (Lipsey & Wilson, 2001):

$$ES = t \cdot \sqrt{\frac{N_G + N_B}{N_G \cdot N_B}}$$

$$ES = \sqrt{F \cdot \frac{(N_G + N_B)}{N_G \cdot N_B}}$$

Please find all resultant effect sizes along with study, sample, and publication characteristics in the Supplementary Material S1.

Influential cases. We identified influential effect sizes using Viechtbauer's and Cheung's (2010) diagnostics based on the assumption of random effects, that is, significant variation in effect sizes between studies. These diagnostics included student residuals, Cook's

distances, and other leave-one-out deletion measures (Viechtbauer & Cheung, 2010). If indeed several cases are flagged as influential, further sensitivity analyses might be warranted which compare the overall effect size before and after removing these cases.

Correction for unreliability. Next to possible selection biases in reported effect sizes, measurement bias might influence the report of an overall effect (Schmidt & Hunter, 2015). More precisely, outcome measures—in our case, performance measures of digital competences—are not perfectly reliable, as indicated for instance by internal consistencies below 1. This unreliability might in fact lead to biased effect sizes. Some researchers consequently recommended correcting the effect sizes obtained from primary studies using score reliability measures (Baugh, 2002; Schmidt & Hunter, 2015). However, the current debate surrounding these corrections does not draw a clear picture as to whether such corrections are in fact needed (Cheung, 2015). Moreover, unreliability corrections are oftentimes problematic, because (a) not all primary studies report score reliabilities, (b) reported score reliabilities may differ substantially in their estimation (e.g., Cronbach's α vs. retest-reliability vs. reliabilities based on models of item response theory), (c) a correction for unreliability should include and quantify the between-study variation in reliability scores (Raykov & Marcoulides, 2013). Given these issues, we compare the overall gender differences between uncorrected and corrected effect sizes as a part of the sensitivity analyses.

Meta-analytic models. To aggregate the reported effect sizes across studies and thus estimate an overall effect size that represents gender differences in ICT literacy, we specified two types of univariate meta-analytic models and compared them with each other. First, we specified a fixed-effects model to the data—a model that only provides an estimate of the overall effect sizes without any between-study variation. Second, we allowed for between-study variation in effects by specifying a random-effects model (Borenstein et al., 2009). The between-study variance τ^2 was estimated using the DerSimonian-Laird moment estimator

(DerSimonian & Laird, 1986). Comparing the two meta-analytic models (i.e., fixed- vs. random-effects models) provided evidence for or against a significant variation of effects between studies (Cheung, 2015). If a random-effects model fitted the data better than a fixed-effects model, possible moderator variables were introduced to explain the between-study variance in mixed-effects models. The models were compared using goodness of fit information criterion (i.e., AIC, Akaike, 1974; BIC, Schwarz, 1978) and ratio tests. All models were specified in the R packages ‘metafor’ (Viechtbauer, 2017) and ‘meta’ (Schwarzer, 2017) using the inverse variance weighting method (Borenstein et al., 2009). Please find the R sample code in the Supplementary Material S3.

3.3 Publication Bias and Sensitivity Analysis

We conducted several analyses of publication bias: First, we examined the funnel plot and performed trim-and-fill-analyses (Duval & Tweedie, 2000). Using Egger’s linear regression test, we tested the asymmetry of the resultant funnel plot (Egger et al., 1997). Second, we compared the effect sizes obtained from published studies and grey literature (Schmucker et al., 2017). Third, we estimate fail-safe N ’s on the basis of Rosenthal’s weighted procedure (Borenstein et al., 2009). Fourth, we examined the p -curve that resulted from the statistics underlying the gender differences (Simonsohn, Nelson, & Simmons, 2014). If the primary studies have evidential value, the p -curve should be right-skewed, and left-skewed otherwise. P -curves were obtained from the ‘ P -curve Online App’ (Simonsohn, Nelson, & Simmons, 2017).

Finally, we tested the sensitivity of our findings with respect to two factors: the correction of effect sizes for unreliability (i.e., corrected vs. uncorrected effects) and the treatment of multiple effect sizes obtained from the ICILS 2013 study (i.e., overall ICILS 2013 effect vs. multiple ICILS 2013 effects from the 21 participating countries).

4 Results

4.1 Description of Primary Studies

Table 1 provides a descriptive summary of the $m = 23$ primary studies based on the extracted $k = 46$ effect sizes. Notably, most effect sizes (73.9 %) were based on assessments of ICT literacy that targeted the application of certain skills rather than the reproduction of knowledge. Moreover, for about 80 % of the effect sizes, score reliability coefficients were available, allowing us to examine the differences between uncorrected and corrected effect sizes. Test fairness seemed to play a relevant role in most studies, as almost 70 % of effect sizes were tests for gender invariance. The overall sample of studies comprised more secondary school students (87 %) than primary school students (13 %), which were mostly part of a randomized and/or stratified student sample (84.8 %). This dominance of randomization and stratification was due to the inclusion of national and international large-scale studies assessing digital competence. More than 50 % of the effect sizes were from European samples; effect sizes from Asian, American, and Australian samples were almost equally distributed. Publications were published between 2007 and 2017, and most effect sizes were based on publications from the year 2014 (Table 1).

The overall sample size amounted to $N = 121614$ students (girls: $n = 59489$, boys: $n = 62125$; on average, 2644 students were available per effect size. Sample sizes ranged between 24 and 6237. The average score reliability was 0.86 ($SD = 0.08$, $Mdn = 0.89$, range = 0.61–0.95), providing an acceptable measurement precision.

4.2 Overall Effect Size and Between-Study Variation (RQ 1)

To summarize the effect sizes extracted from the primary studies of gender differences in ICT literacy and their between-study variation (RQ 1), we first specified a fixed-effects model assuming no between-study variation. This model resulted in a small positive but statistically significant effect size in favor of girls, $g = +0.131$, 95% CI = [0.120, 0.142], $k =$

46, $z = 23.2$, $p < .001$. Second, we loosened the variance assumption in the model and specified a random-effects model. The overall effect size was again positive and significant and differed only marginally from that of the fixed-effects model, $g = +0.124$, 95% CI = [0.084, 0.160], $k = 46$, $z = 6.8$, $p < .001$. This model uncovered significant heterogeneity in the effect sizes ($Q[45] = 407.3$, $p < .001$, $I^2 = 89.0\%$) and a positive between-study variance ($\tau^2 = 0.012$, $SE = 0.004$). To test whether this variance is significantly different from zero, we compared the fixed- and random-effects models using the Likelihood-Ratio Test (LRT) and information criteria (Cheung, 2015). The LRT suggested that the random-effects model fitted the data significantly better than the fixed-effects model, $\chi^2(1) = 255.3$, $p < .001$. Moreover, the information criteria were smaller for the random-effects models (AIC = -36.1, BIC = -32.4, cAIC = -35.8) than for the fixed-effects model (AIC = 217.2, BIC = 219.0, cAIC = 217.38). These observations testify to the existence of significant between-study variation on effect sizes of gender differences. We will therefore proceed with the random-effects model as the baseline model for further analyses.

Overall, the positive and significant effect size suggested that girls outperformed boys in performance measures of digital competence. Nevertheless, the overall effect size varied between primary studies. All effect sizes and the results of the fixed- and random-effects modeling are depicted in a forest plot (see Figure 1).

4.3 Moderator Analyses (RQ 2)

Given the significant between-study variance in gender differences, we further examined to what extent study-, sample-, and publication characteristics may explain this variance. Table 2 shows the results of the corresponding moderator analyses.

Neither study characteristics that referred to the interactivity of assessment tasks nor the type of skills assessed (i.e., the application of digital skills in certain contexts vs. more knowledge-lean and theoretical skills) moderated the gender differences. Moreover, whether

researchers made an attempt to examine the fairness of the assessment, often manifested by the invariance of the measures used to indicate ICT literacy across gender, did not affect the overall gender differences.

Apart from the educational level the assessments were administered in, neither the sampling procedure nor the geographical region the study was conducted in explained the between-study variation significantly. In primary schools, the gender differences were slightly higher than in secondary schools (primary: $g = +0.201$, secondary: $g = +0.113$; $p < .10$)—this difference explained 7.5 % of the between-study variation.

As mentioned earlier, the status of publication (i.e., published vs. grey literature) did not moderate the gender differences. Overall, these variables moderated the gender differences in digital competence to only a limited extent, leaving the significant between-study variance largely unexplained.

4.4 Publication Bias

As mentioned earlier, we examined the extent to which publication bias occurred using several procedures. First, a graphical inspection of the funnel plot suggested some degree of asymmetry (see Figure 2), and the corresponding trim-and-fill analyses indicated that four studies should be added to adjust for this asymmetry. The resultant overall effect size after adding these studies was $g = +0.142$, 95% CI = [0.104, 0.180], $k = 50$, $z = 7.4$, $p < .001$. This effect was associated with significant heterogeneity ($Q[49] = 490.5$, $p < .001$, $I^2 = 90.0$ %) and a positive between-study variance ($\tau^2 = 0.014$), both of which were close to the original estimates based on $k = 46$ primary studies. Egger's linear regression test of funnel plot asymmetry indicated the symmetry of the plot, $t(44) = -0.79$, $p = .43$, bias = -0.83, $SE = 1.04$.

Second, we compared the overall effect sizes between published and grey literature. Adding publication status as an explanatory variable to the random-effects model did not

provide evidence for significant differences in effect sizes, $Q_M(1) = 1.67, p = .20$. Hence, published ($g = +0.160, 95\% \text{ CI} = [0.095, 0.224], k = 17$) and grey literature ($g = +0.109, 95\% \text{ CI} = [0.066, 0.151], k = 29$) did not differ in the overall effects reported as gender differences.

Third, the fail-safe N analyses based on Rosenthal's method showed that 7898 effect sizes would be needed to turn the observed effects into insignificant effects for the target significance level of 0.05. For a significance level of 0.01, the fail-safe N was 3926, and 2205 for 0.001 respectively. These numbers are large in comparison to the available effect sizes.

Fourth, the p -curve was right-skewed (see Figure 3) and thus suggested that the primary studies had evidential value.

In sum, these analyses did not provide strong evidence for publication bias—in fact, they indicated only a small risk of publication bias in the data.

4.5 Sensitivity Analyses

Effects of correcting for unreliability. After correcting the original effect sizes for unreliability, the overall effect obtained from a fixed-effects model was $g = +0.141, 95\% \text{ CI} = [0.130, 0.152], k = 46, z = 25.0, p < .001$. A random-effects model resulted in $g = +0.135, 95\% \text{ CI} = [0.097, 0.174], k = 46, z = 6.9, p < .001$. The corresponding between-study variance was $\tau^2 = 0.014 (SE = 0.006)$, and the heterogeneity test indicated significant between-study heterogeneity ($Q[45] = 475.5, p < .001, I^2 = 90.5\%$). As for the uncorrected effect sizes, the random-effects model (AIC = -24.1, BIC = -20.4, cAIC = -23.8) was statistically preferred over the fixed-effects model (AIC = 285.4, BIC = 287.2, cAIC = 285.5), as indicated by smaller values of information criteria and the LRT, $\chi^2(1) = 311.4, p < .001$. Hence, the decision for the random-effects model was not affected by the unreliability correction. Considering this and the small differences in both effect sizes and between-study variances between uncorrected and corrected effects, we conclude that our findings on the overall gender differences were not sensitive toward the correction for unreliability. Moreover, the

moderation effects were largely unaffected. Notably, there was a tendency toward higher effect sizes after correcting for unreliability, leading to more pronounced differences in effect sizes between primary and secondary school students. Please see the findings obtained from these sensitivity analyses in the Supplementary Material S2.

Effects of aggregating effect sizes from large-scale studies. In all previous analyses, we include multiple yet independent effect sizes from the international large-scale study ICILS 2013. To test whether these effects may bias the estimation of an overall effect size, we conducted a two-step procedure: In the first step, we aggregated the ICILS 2013 effect sizes using a random-effects model. In the second step, we performed a meta-analysis on the entire sample of effect sizes from all studies and included the overall ICILS 2013 effect size as a single effect size. In other words, we summarized the multiple effects in ICILS as a single ICILS 2013 effect size: For uncorrected effect sizes, the overall ICILS effect was $g = +0.132$ (95% CI = [0.098, 0.165], $k = 21$, $z = 7.6$, $p < .001$) and $g = +0.139$ (95% CI = [0.104, 0.175], $k = 21$, $z = 7.6$, $p < .001$) for corrected effect sizes.

Summarizing the overall effects including the single effect obtained from ICILS 2013 resulted in $g = +0.118$ (95% CI = [0.056, 0.180], $k = 26$, $z = 3.7$, $p < .001$) for uncorrected and $g = +0.133$ (95% CI = [0.066, 0.200], $k = 26$, $z = 3.9$, $p < .001$) for corrected effect sizes under random-effects models. The corresponding between-study variances were $\tau^2 = 0.020$ ($SE = 0.009$) and $\tau^2 = 0.023$ ($SE = 0.010$) respectively. Once again, these values did not differ substantially from those obtained from the random-effects models with multiple independent effect sizes from ICILS 2013.

Given the smaller sample size of studies when an overall ICILS 2013 effect is used, some moderator effects became more pronounced. For instance, the moderation effects for primary and secondary school students were larger, and moderation by publication status became statistically significant ($p < .10$), yet with small effects explaining up to 3.6 % of

between-study variance. Despite this observation, the moderation effects were not sensitive toward the handling of the ICILS 2013 effect sizes. Please find the results obtained from these sensitivity analyses in the Supplementary Material S2.

Power and influential cases. In all conditions, the power to detect meaningful effect sizes, given the overall sample sizes, was 100 % (for a more detailed discussion of power in meta-analyses, please review Valentine, Pigott, & Rothstein, 2010). Moreover, the analysis of influential cases did not flag any effect size to be influential in the sample of effect sizes (see Supplementary Material S2).

5 Discussion

5.1 Summary of Results

In this meta-analysis, we synthesized the existing body of research on gender differences in ICT literacy or, more precisely, in the performance measures of ICT literacy. Comparing the fixed- and random-effects models, we found that the random-effects model fitted that data significantly better. This model revealed a positive and significant effect size, favoring girls ($g = +0.13$) and significant between-study variation of these gender differences. Further moderator analyses showed that only the students' educational level (i.e., primary vs. secondary education) explained this variation significantly; the gender differences were slightly higher in primary schools. Moreover, our analyses indicated little risk of publication bias in the data, and the sensitivity analyses supported the robustness of the overall effect size against the treatment of a large-scale data set (i.e., that of ICILS 2013). Along the same lines, further sensitivity analyses showed that the overall gender differences were not sensitive toward the correction for unreliability, neither did the analysis of influential cases flag any effect sizes.

5.2 Gender Differences in ICT Literacy and Their Between-Study Variation (RQ1)

Based on a random-effects model, our research synthesis showed an overall effect size of $g = +0.131$, 95% CI = [0.120, 0.142], representing gender differences in favor of girls. Four key observations accompany this finding:

- (1) *The effect is statistically significant:* The fact that the overall effect size differs significantly from zero testifies to the existence of gender differences or, in terms of equity and diversity, to a gender gap. This finding is by no means trivial, as meta-analyses and large-scale studies in different, yet related domains suggested. For instance, synthesizing the findings from the PISA 2015 study in science, Stoet and Geary (2018) could not find evidence for a significant, overall effect of students' gender on scientific literacy in a large sample of students from 67 countries ($d = -0.01$). Similarly, Lindberg and colleagues (2010), who reviewed 242 studies that reported gender differences in mathematics performance between 1990 and 2007, could not find evidence for a significant gender effect either ($d = 0.05$). Hyde (2005) supported this observation and reported a broad range of effect sizes for different measures of mathematics achievement ($d = -0.14$ – 0.16). Our meta-analysis, however, suggests a significant effect. Considering that the extant literature reports considerable variation of gender effects across domains (Voyer & Voyer, 2014), we suspect that the existence of gender differences may be domain-specific or specific to certain assessment domains. In any case, the gender effect on ICT literacy is evident and deviates from those reported for mathematics in some of the recent research syntheses. Despite this finding for the relatively young domain of ICT literacy, we encourage to monitor gender differences regularly to identify possible changes over time.

Furthermore, we would like to bring to attention our broad conceptualization of ICT literacy as comprised of certain knowledge domains and skills. Our unidimensional view on ICT literacy provided us with only one overall effect size. A further differentiation in further sub-dimensions of the concept may have resulted in deeper insights about which aspects of ICT literacy are more or less prone to gender effects. However, the small number of studies reporting gender differences for these sub-dimensions limited our analyses.

(2) *The effect is positive*: This observation uncovers higher average performance of girls on ICT literacy tasks and aligns with the effects identified in verbal domains. More specifically, in their recent meta-analysis on gender differences in academic achievement, Voyer and Voyer (2014) found effects in favor of girls, especially in language learning ($d = +0.37$), science ($d = +0.15$), and social sciences ($d = +0.17$), whereas those in mathematics were closer to zero ($d = +0.07$). Machin and Pekkarinen (2008) also found extreme differences in gender effects between achievement in reading ($d = +0.30$) and mathematics ($d = -0.10$). Assessments of ICT literacy oftentimes rely on verbal skills, such as communication—this may explain the direction of gender effects in ICT literacy as compared to more verbal domains.

Interestingly, Cai et al. (2017) meta-analyzed the gender differences in ICT attitudes and self-efficacy and found a significant effect favoring boys ($g = -0.17$). Although their meta-analysis examined somehow similar concepts around ICT, it contrasts the finding obtained from our meta-analysis. We believe that this contrast may be explained by a common observation: Boys tend to overestimate their ICT knowledge and skills, while girls tend to underestimate themselves, irrespective of their actual knowledge and skills (Aesaert, Voogt, Kuiper, & van Braak, 2017). It

seems obvious that self-reported performance and actual performance may not necessarily go together (Honicke & Broadbent, 2016). This knowledge is critical because if school interventions developed to correct for the gender gap are based on self-report measures, they might work counterproductively. Examining the gender differences in light of the interplay between ICT self-efficacy and ICT literacy may be subject to future meta-analyses.

- (3) *The effect is small.* In comparison to gender differences in other domains, the effect for ICT literacy is relatively small and comparable to those found in some of the above-mentioned studies in social sciences, science, and mathematics. Moreover, the magnitude of the effect, although in a different direction, is similar to that reported by Cai et al. (2017) for ICT attitudes and self-efficacy ($g = -0.17$). Despite the fact that the overall effect aligns with others, it also indicates that ICT literacy may represent a domain in which equity can be achieved to a certain extent. In contrast to the claims surrounding strong disadvantages of girls in performance in technological domains (King & Winthrop, 2015), our meta-analysis suggests that the gender gap may not be as practically critical as these claims suggest.
- (4) *The effect shows significant variation between studies:* Primary studies abounded in different findings that comprised significantly negative gender effects in favor of boys, insignificant effects, and significantly positive effects in favor of girls. This variation may have been caused by certain design features of the ICT literacy assessments that were not captured by the coding of possible moderators in our meta-analysis (e.g., the time students spent on assessment tasks, their test-taking effort, or prior experience with similar tasks). Moreover, the varying methodologies and models used to report the gender differences may have also

contributed to this variation (e.g., Hohlfeld, Ritzhaupt, & Barron, 2013). Further confounders including motivational constructs and ICT stereotype beliefs could be examined as explanatory variables in future meta-analyses.

5.3 Explaining Between-Study Variation (RQ 2)

Interestingly, our findings showed that only students' educational level moderated gender differences, with higher effects for primary school students. This is quite a surprising result as it seems that the gender gap is higher among younger children. This might be due to the fact that primary school students have less access to and therefore experience with technology at school, and are to a lesser degree taught digital competence (Aasen, Møller, Rye, Ottesen, Prøitz, & Hertzberg, 2012). Moreover, one limitation might be that primary school students' digital competence is less often assessed compared to secondary school students (Siddiq et al., 2016).

Our data did not reveal any further moderation effects of region, sampling procedure, or task mode (i.e., whether the task required interactivity within the test environment). Unexpectedly, test content (i.e., the competences measured in the tests) did not moderate the gender effect. Recent research has shown that, within the ICT literacy framework, boys and girls tend to score differently in different competence areas. For instance, girls were identified to outperform boys on scales related to using learning-related software and tools (e.g., word processing, spreadsheets, presentation software, image processing, and measures related to communication, social networking, and security issues). While boys performed significantly better than girls on scales that required more technical knowledge (e.g., basic operations, information networks, programming, and database operations; Aesaert & van Braak, 2015; Christoph, Goldhammer, Zylka, & Hartig, 2015; Lau & Yuen, 2014; Meri-Tuulia, Antero, & Suvi-Sadetta, 2017). An explanation for this might be that the studies included in this meta-analysis mostly reported ICT literacy as one composite measure and not for specific

knowledge domains or skills—in fact, gender differences for the areas outlined in the DIGCOMP framework were reported to a very limited extent. Because of the recent recognition of ICT literacy as a broad and complex framework, researchers within the field of assessment of ICT literacy have started to consider the different sub-scales or dimensions of the construct (Aesaert & van Braak, 2015; Meri-Tuulia et al., 2017). We encourage researchers to present gender differences for sub-scales or dimensions of ICT literacy. A future update of this meta-analysis could then provide more fine-grained measures of gender differences.

6 Conclusion

The present meta-analysis examined the gender differences in ICT literacy, as measured by performance-based assessments, and the variables that may moderate these differences. The results displayed gender differences favoring girls across K-12 education across all regions included in the primary studies. The effect size was generally small in magnitude, but its consistency suggests that it should not be ignored. We believe that our meta-analysis lays the foundation to establish the existence of a generalized, female advantage in the domain of ICT, next to other domains (Voyer & Voyer, 2014). Moreover, our findings contradict the claims that boys have higher ICT literacy—claims that were merely based on self-efficacy rather than performance measures (Aesaert & van Braak, 2015). Our findings further testify to the fact that the gender gap may not be as strong as expected in the domain of ICT. However, we still see a need for research that determines the factors related to gender differences in ICT literacy and their underlying causes.

Tables

Table 1

Description of primary studies on gender differences in measures of ICT literacy

Characteristics	<i>k</i>	Proportion of effect sizes
Study characteristics		
Type of outcome measure		
Interactive tasks	25	54.3 %
Static tasks	21	45.7 %
Skills assessed		
Applied skills	34	73.9 %
Theoretical skills	12	26.1 %
Test fairness		
Fairness not examined	14	30.4 %
Fairness examined	32	69.6 %
Score reliability		
Reliability not reported	9	19.6 %
Reliability reported	37	80.4 %
Sample characteristics		
Educational level		
Primary school	6	13.0 %
Secondary school	40	87.0 %
Sampling		
Convenience sample	7	15.2 %
Randomized and/or stratified sample	39	84.8 %
Region of the study sample		
America	7	15.2 %
Asia	8	17.4 %
Australia	7	15.2 %
Europe	24	52.2 %
Publication characteristics		
Publication status		
Published	17	37.0 %
Grey literature	29	63.0 %
Publication year		
2007	1	2.3 %
2008	2	4.3 %
2010	6	13.0 %
2012	5	10.9 %
2013	5	10.9 %

2014	23	50.0 %
2015	2	4.3 %
2017	2	4.3 %

ACCEPTED MANUSCRIPT

Table 2

Analyses of moderation effects by study, sample, and publication characteristics

<i>Moderator variables</i>	<i>k</i>	<i>g</i>	95 % CI	$Q_M(df)$	<i>p</i>	$Q_E(df)$	<i>p</i>	R^2
Study characteristics								
Type of outcome measure								
Interactive	25	+0.141	[0.092, 0.190]	1.02 (1)	.31	407.20 (44)	< .001	0.0 %
Static	21	+0.103	[0.049, 0.158]					
Skills assessed								
Applied skills	34	+0.135	[0.095, 0.176]	1.11 (1)	.29	374.98 (44)	< .001	4.8 %
Theoretical skills	12	+0.092	[0.022, 0.162]					
Test fairness								
Fairness not examined	14	+0.099	[0.023, 0.174]	0.58 (1)	.45	395.00 (44)	< .001	0.0 %
Fairness examined	32	+0.132	[0.091, 0.173]					
Sample characteristics								
Educational level								
Primary school	6	+0.201	[0.105, 0.297]	2.82 (1)	.09	358.18 (44)	< .001	7.5 %
Secondary school	40	+0.113	[0.075, 0.150]					
Sampling								
Convenience sample	9	+0.151	[0.050, 0.251]	0.31 (1)	.58	401.30 (44)	< .001	0.0 %
Randomized and/or stratified sample	37	+0.120	[0.081, 0.159]					
Regions								
America	7	+0.185	[0.098, 0.272]	3.07 (3)	.38	321.55 (42)	< .001	12.4 %
Asia	8	+0.145	[0.059, 0.231]					
Australia	7	+0.091	[0.012, 0.170]					
Europe	24	+0.112	[0.064, 0.159]					
Publication status								
Published literature	17	+0.160	[0.095, 0.224]	1.67 (1)	.20	382.05 (44)	< .001	1.6 %
Grey literature	29	+0.109	[0.066, 0.151]					

Figures

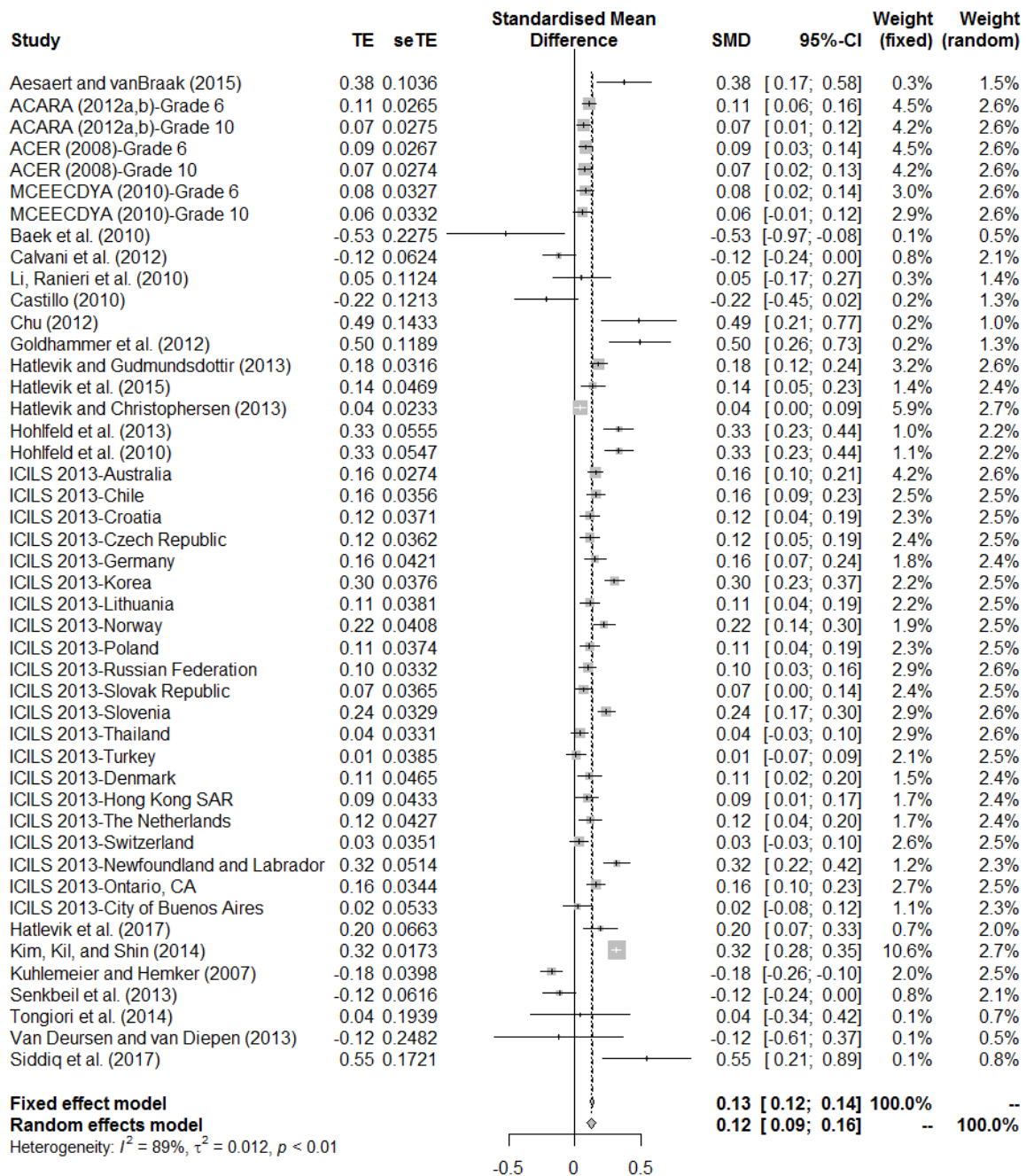


Figure 1. Forest plot of the uncorrected effect sizes, including the overall effects derived from the fixed- and random-effects modeling.

Note. TE = Primary ‘treatment effect’ size Hedges’ g , seTE = Standard error of Hedges’ g , SMD = Standardized mean difference (equal to Hedges’ g). The 95% Wald confidence intervals are shown.

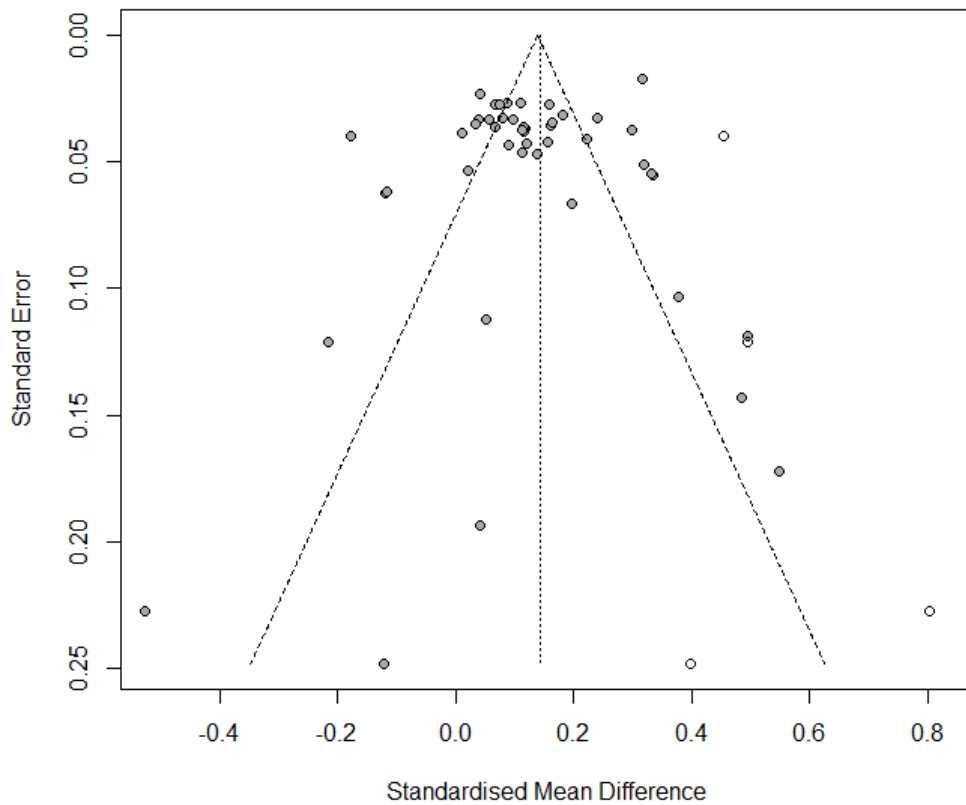
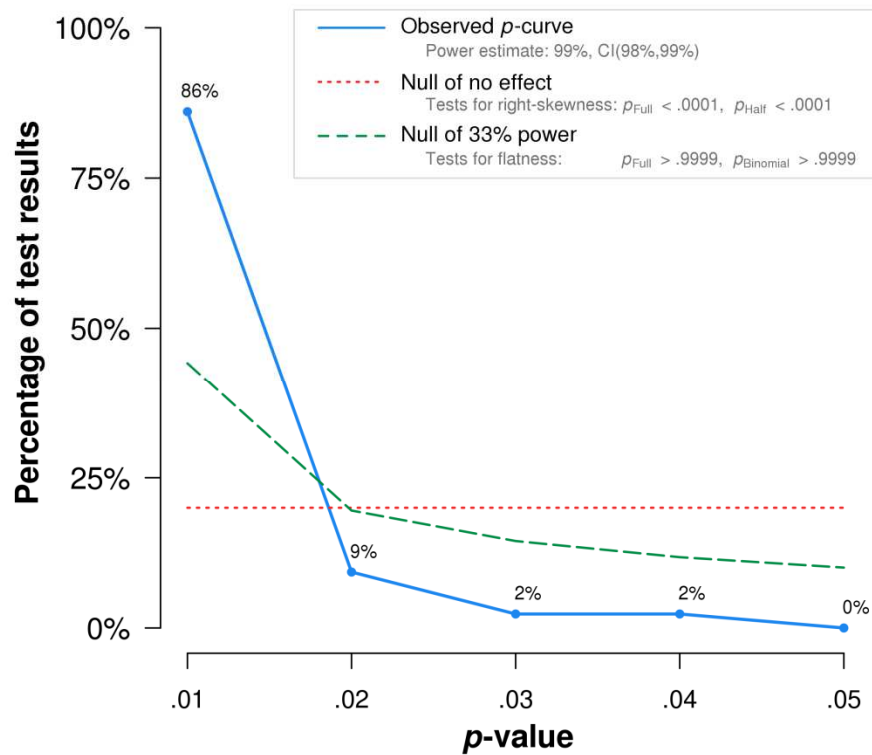


Figure 2. Funnel plot of the uncorrected effect sizes (grey circles), including the effects suggested by the trim-and-fill-analyses (white circles).



Note: The observed p -curve includes 43 statistically significant ($p < .05$) results, of which 42 are $p < .025$. There were 25 additional results entered but excluded from p -curve because they were $p > .05$.

Figure 3. P -curve based on uncorrected effect sizes of gender differences ($k = 46$).

References

Studies included in the meta-analysis are marked with an asterisk ().*

Aasen, P., Møller, J., Rye, E., Ottesen, E., Prøitz, T. S., & Hertzberg, F. (2012).

Kunnskapsløftet som styringsreform – et løft eller et løfte? Forvaltningsnivåenes og institusjonenes rolle i implementeringen av reformen [eng. The ‘Knowledge promotion’ reform – The role of administration and institutions in the implementation of the reform]. Report 20/2012. Oslo: NIFU and the Department of Educational Sciences, UiO.

*ACARA. (2012). National assessment program - ICT literacy years 6 & 10 report 2011.

Sidney: Australian Curriculum, Assessment and Reporting Authority. Retrieved from http://www.nap.edu.au/verve/_resources/nap_ictl_2011_public_report_final.pdf.

Aesaert, K., van Nijlen, D., Vanderlinde, R., & van Braak, J. (2014). Direct measures of digital information processing and communication skills in primary education: Using item response theory for the development and validation of an ICT competence scale. *Computers & Education*, 76, 168-181. doi:10.1016/j.compedu.2014.03.013

*Aesaert, K., & van Braak, J. (2015). Gender and socioeconomic related differences in performance based ICT competences. *Computers & Education*, 84, 8-25.

Aesaert, K., Voogt, J., Kuiper, E., & van Braak, J. (2017). Accuracy and bias of ICT self-efficacy: An empirical study into students’ over- and underestimation of their ICT competences. *Computers in Human Behavior*, 75, 92-102.

doi:10.1016/j.chb.2017.05.010

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.

- Ala-Mutka, K. (2011). Mapping digital Competence: Towards a conceptual understanding. Seville.: European Commission, Joint Research Centre. Institute for Prospective Technological Studies. Retrieved from <http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id!44699>.
- *Baek, S. G., Lim, C. I., Kim, H. S., Yu, Y. L., Kim, M. R., & Lee, S. R. (2010). Analysis of influential variables on performance-based ICT literacy level for Korean students. In J. Sanchez, & K. Zhang (Eds.), Proceedings of world conference on e-learning in corporate, government, healthcare, and higher education 2010 (pp. 1560-1565). Chesapeake, VA: AACE.
- Balanskat, A. (2009). Study of the impact of technology in primary schools—STEPS. Synthesis report. European Commission, Directorate General Education and Culture. Retrieved from http://erte.dge.mec.pt/sites/default/files/Recursos/Estudos/synthesis_report_steps_en.pdf
- Baugh, F. (2002). Correcting Effect Sizes for Score Reliability: A Reminder that Measurement and Substantive Issues are Linked Inextricably. *Educational and Psychological Measurement*, 62(2), 254-263. doi:10.1177/0013164402062002004
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, West Sussex: John Wiley & Sons, Ltd.
- Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, 105, 1-13. <http://dx.doi.org/10.1016/j.compedu.2016.11.003>
- *Calvani, A., Fini, A., Ranieri, M., & Picci, P. (2012). Are young generations in secondary school digitally competent? A study on Italian teenagers. *Computers & Education*, 58, 797-807. <http://dx.doi.org/10.1016/j.compedu.2011.10.004>.

- *Castillo, M. (2010). Technological literacy: Design and testing an instrument to measure eight-grade achievement in the technology education (Paper presented at the American Society for Engineering Education, Louisville, KY).
- Cheung, M. W.-L. (2015). *Meta-Analysis: A Structural Equation Modeling Approach*. Chichester, West Sussex: John Wiley & Sons, Ltd.
- Christoph, G., Goldhammer, F., Zylka, J. & Hartig, J. (2015). Adolescents' computer performance: The role of self-concept and motivational aspects. *Computers & Education*, 81, 1–12. DOI: 10.1016/j.compedu.2014.09.004
- *Chu, S. K. W. (2012). Assessing information literacy: A case study of primary 5 students in Hong Kong (Vol. 15). American Association of School Librarians. Retrieved from <http://www.ala.org/aasl/slr/volume15/chu>.
- Claro, M., Preiss, D. D., San Martín, E., Jara, I., Hinostroza, J. E., Valenzuela, S., et al. (2012). Assessment of 21st century ICT skills in Chile: Test design and results from high school level students. *Computers & Education*, 59, 1042-1053.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177-188. doi:10.1016/0197-2456(86)90046-2
- 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(2), 455-463. doi:10.1111/j.0006-341X.2000.00455.x
- Educational Testing Service. (2007). *Digital transformation: A framework for ICT literacy. A report of the International ICT Literacy Panel*. Princeton, NJ: Center for Global Assessment. Retrieved from [https://www.ets.org/Media/Tests/Information and Communication Technology Literacy/ictreport.pdf](https://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf)

- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629-634. doi:10.1136/bmj.315.7109.629
- Fan, X., Chen, M. (2001). Parental involvement and students' academic achievement: A Meta-Analysis. *Educational Psychology Review*, *13*(1), 1-22.
- Ferrari, A. (2013). *DIGCOMP: A framework for developing and understanding digital competence in Europe*. In Y. Punie & B. N. Brečko (Eds.), *JRC scientific and policy reports*. Luxembourg: Publications Office of the European Union. doi:10.2788/52966
- *Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). Preparing for life in a digital age. The IEA international computer and information literacy study (ICILS), international report. IEA: Springer Open.
- Gallardo-Echenique, E. E., de Oliveira, M. J., Marques-Molias, L., & Esteve-Mon, F. (2015). Digital competence in the knowledge society. *MERLOT Journal of Online Learning and Teaching*, *11*(1). Retrieved from jolt.merlot.org/vol11no1/Gallardo-Echenique_0315.pdf.
- *Goldhammer, F., Naumann, J., & Keßel, Y. (2012). Assessing individual differences in basic computer Skills: Psychometric characteristics of an interactive performance measure. *European Journal of Psychological Assessment*, *29*, 263-275. <http://dx.doi.org/10.1027/1015-5759/a000153>.
- Hakkarainen, K., Ilomaki, L., Lipponen, L., Muukkonen, H., Rahikainen, M., Tuominen, T., et al. (2000). Students' skills and practices of using ICT: Results of a national assessment in Finland. *Computers & Education*, *34*, 103-117.
- *Hatlevik, O. E., & Christophersen, K. (2013). Digital competence at the beginning of upper secondary school: Identifying factors explaining digital inclusion. *Computers & Education*, *63*, 240-247. Doi:10.1016/j.compedu.2012.11.015.

- *Hatlevik, O. E., & Gudmundsdottir, G. (2013). An emerging digital divide in urban school children's information literacy: Challenging equity in the Norwegian school system. *First Monday*, 18(4). doi:10.5210/fm.v18i4.4232.
- *Hatlevik, O. E., Ottestad, G., & Throndsen, I. (2015). Predictors of digital competence in 7th grade: A multilevel analysis. *Journal of Computer Assisted Learning*, 31, 220-231. <http://dx.doi.org/10.1111/jcal.12065>.
- *Hatlevik, O. E., Scherer, R., & Christophersen, K. (2017). Moving beyond the study of gender differences: An analysis of measurement invariance and differential item functioning of an ICT literacy scale. *Computers & Education*, 113, 280-293. doi:10.1016/j.compedu.2017.06.003
- *Hohlfeld, T. N., Ritzhaupt, A. D., & Barron, A. E. (2010). Development and validation of the student tool for technology literacy (ST2L). *Journal of Research on Technology in Education*, 42, 361-389.
- *Hohlfeld, T. N., Ritzhaupt, A. D., & Barron, A. E. (2013). Are gender differences in perceived and demonstrated technology literacy significant? It depends on the model. *Education Technology Research and Development*, 61, 639-663. Doi:10.1007/s11423-013-9304-7
- Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, 17, 63-84. doi:10.1016/j.edurev.2015.11.002
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60(6), 581-592. doi:10.1037/0003-066X.60.6.581
- Jackson, L. A., Zhao, Y., KolenicIII, A., Fitzgerald, H. E., Harold, R., & Von Eye, A. (2008). Race, Gender, and Information Technology Use: The New Digital Divide. *CyberPsychology & Behavior*, 11(4). <http://doi.org/10.1089/cpb.2007.0157>

- *Kim, H.-S., Kil, H.-J., & Shin, A. (2014). An analysis of variables affecting the ICT literacy level of Korean elementary school students. *Computers & Education*, 77, 29-38.
doi:10.1016/j.compedu.2014.04.009
- King, E. M., & Winthrop, R. (2015). Today's challenges for girls' education. *Global Economy & Development*, Working Paper No. 90. Washington, DC: The Brookings Institution. Retrieved from <https://www.brookings.edu/wp-content/uploads/2016/07/Todays-Challenges-Girls-Educationv6.pdf>
- *Kuhlemeier, H., & Hemker, B. (2007). The impact of computer use at home on students' Internet skills. *Computers & Education*, 49, 460-480.
- Lankshear, C., & Knobel, M. (2008). *Digital literacies. Concepts, policies, and practices*. New York: Peter Lang Publishing Inc.
- Lau, W. W. F. & Yuen, A. H. K. (2014). Developing and validating of a perceived ICT literacy scale for junior secondary school students: Pedagogical and educational contributions. *Computers & Education*, 78, 1–9. DOI:10.1016/j.compedu.2014.04.016
- Law, N., Lee, Y., & Yuen, H. K. (2009). The impact of ICT in education policies on teacher practices and student outcomes in Hong Kong. In F. Scheuermann, & F. Pedro (Eds.), *Assessing the effects of ICT in education - Indicators, criteria and benchmarks for international comparisons* (pp. 143-164). France: OECD.
- Lennon, M., Kirsch, I., von Davier, M., Wagner, M., & Yamamoto, K. (2003). *Feasibility study for the PISA ICT literacy assessment, Report to Network A*. Paris: OECD.
- *Li, Y., & Ranieri, M. (2010). Are “digital natives” really digitally competent? - A study on Chinese teenagers. *British Journal of Educational Technology*, 41, 1029-1042.
- Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies in Educational Evaluation*, 32, 317–344.
doi:10.1016/j.stueduc.2006.10.002

- Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*(6), 1123-1135. doi:10.1037/a0021276
- Lipsey, M. W., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Machin, S., & Pekkarinen, T. (2008). Global Sex Differences in Test Score Variability. *Science*, *322*, 1331-1332. doi:10.1126/science.1162573
- Markauskaite, L. (2006). Toward an integrated analytical framework of information and communications technology literacy: From intended to implemented and achieved dimensions. *Information Research*, *11*(3).
- *MCEECDYA. (2010). National assessment program e ICT literacy years 6 & 10, report 2008. Carlton South: Ministerial Council for Education, Early Childhood Development and Youth Affairs. Retrieved from http://www.nap.edu.au/verve/_resources/2008_nap_ictl_public_report.pdf.
- *MCEETYA. (2007). National assessment program - ICT literacy years 6 & 10, report 2005. Carlton South: Ministerial Council on Education, Employment, Training and Youth Affairs. Retrieved 20.02.2015 from http://www.nap.edu.au/verve/_resources/2005_ICTL_Public_Report_file_main.pdf.
- Meri-Tuulia, K., Antero, K., & Suvi-Sadetta, K. (2017). Differences between the genders in ICT skills for Finnish upper comprehensive school students: Does gender matter?. *Seminar.Net*, *13*(2). Retrieved from <https://journals.hioa.no/index.php/seminar/article/view/2304>
- Pagani, L., Argentin, G., Gui, M. & Stanca, L. (2016). The impact of digital skills on educational outcomes: evidence from performance tests. *Educational studies*, *42*, 137-162. DOI: 10.1080/03055698.2016.1148588

- Pamuk, S., & Peker, D. (2009). Turkish pre-service science and mathematics teachers' computer related self-efficacies, attitudes, and the relationship between these variables. *Computers & Education*, 53, 454-461. doi:10.1016/j.compedu.2009.03.004
- Pellegrino, J.W., & Hilton, M.L. (Eds.) (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. National Research Council. Washington, D.C., The National Academies Press.
- Raykov, T., & Marcoulides, G. A. (2013). Meta-Analysis of Scale Reliability Using Latent Variable Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 20(2), 338-353. doi:10.1080/10705511.2013.769396
- Sáinz, M., & Eccles, J. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. *Journal of Vocational Behavior*, 80(2), 486-499.
- Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3 ed.). Thousand Oaks, CA: Sage.
- Schmucker, C. M., Blümle, A., Schell, L. K., Schwarzer, G., Oeller, P., Cabrera, L., . . . on behalf of the, O. c. (2017). Systematic review finds that study data not published in full text articles have unclear impact on meta-analyses results in medical research. *PLoS One*, 12(4), 1-16. doi:10.1371/journal.pone.0176210
- Schwarzer, G. (2017). *meta: General Package for Meta-Analysis*. R package version 4.8-4.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- *Senkbeil, M., Ihme, J. M., & Wittwer, J. (2013). The test of technological and information literacy (TILT) in the national educational panel Study: Development, empirical testing, and evidence for validity. *Journal for Educational Research Online*, 5, 139-161.
- Siddiq, F., Hatlevik, O. E., Olsen, R. V., Throndsen, I., & Scherer, R. (2016). Taking a future

perspective by learning from the past – A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educational Research Review*, 19, 58-84. <http://dx.doi.org/10.1016/j.edurev.2016.05.002>

*Siddiq, F., Gochyyev, P., & Wilson, M. (2017). Learning in Digital Networks–ICT literacy: A novel assessment of students' 21st century skills. *Computers & Education*, 109, 11-37.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-Curve: A Key to the File-Drawer. *Journal of Experimental Psychology: General*, 143(2), 534-547.
doi:10.1037/a0033242

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2017). *P-curve Online App Version 4.06*.
<http://www.p-curve.com/app4/>

Søby, M. (2013). Learning to Be: Developing and understanding digital competence. *Nordic Journal of Digital Literacy*, 8, 134-138. Retrieved from
<https://www.idunn.no/file/pdf/62414156/#page=3>.

Stoet, G., & Geary, D. C. (2018). The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science*, 29(4), 581-593.
doi:10.1177/09567976177417

*Tongori, A., & Pluhar, Z. (2014). An instrument to assess the basic technological actions of ICT literacy. *TEE SZEMLE*, IV, 7-20.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How Many Studies Do You Need? A Primer on Statistical Power for Meta-Analysis. *Journal of Educational and Behavioral Statistics*, 35(2), 215-247. doi:10.3102/1076998609346961

*Van Deursen, A. J. A. M., & Van Diepen, S. (2013). Information and strategic internet skills of secondary students: A performance test. *Computers & Education*, 63, 218-226.

- Vekiri, I., & Chronaki, A. (2008). Gender issues in technology use: perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. *Computers & Education*, 51(3), 1392–1404.
- Viechtbauer, W. (2017). *metafor: Meta-Analysis Package for R*. R package version 2.0-0.
- Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112-125. doi:10.1002/jrsm.11
- Volman, M., van Eck, E., Heemskerk, I., & Kuiper, E. (2005). New technologies, new differences: Gender and ethnic differences in pupils' use of ICT in primary and secondary education. *Computers and Education*, 45, 35-55.
- Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of Curriculum Studies*, 44, 299-332.
- Voyer, D., & Voyer, S. D. (2014). Gender Differences in Scholastic Achievement: A Meta-Analysis. *Psychological Bulletin*, 140 (4), 1174–1204. doi:10.1037/a0036620.
- Zhong, Z. J. (2011). From access to usage: the divide of self-reported digital skills among adolescents. *Computers & Education*, 56, 736–746.
DOI:10.1016/j.compedu.2010.10.016

Highlights

- Girls perform better than boys on performance-based ICT literacy assessments.
- Gender differences are larger in primary schools than in secondary schools.
- The overall effect size is robust across several analysis conditions.
- No evidence of publication bias could be found.
- Overall, the gender differences in ICT literacy are significant but small.