Research Institute for the Evaluation of Public Polices

http://irvapp.fbk.eu

FBK-IRVAPP Working Paper series

Research Institute for the Evaluation of Public Policies

The second digital divide in Europe. A cross-national study on students' digital reading and navigation skills

Davide Azzolini

Antonio Schizzerotto

August 2017

FBK-IRVAPP Working Paper No. 2017-02

The second digital divide in Europe. A cross-national study on students' digital reading and navigation skills

Davide Azzolini
FBK-IRVAPP

Antonio Schizzerotto
FBK-IRVAPP, University of Trento

FBK-IRVAPP Working Paper No. 2017-02

August 2017

Research Institute for the Evaluation of Public Policies
Bruno Kessler Foundation
Vicolo dalla Piccola 2, 38122 Trento (Italy)

Phone: (+39) 0461.314209
Fax: (+39) 0461.314240

E-mail: info@irvapp.it
Website: http://irvapp.fbk.eu

The purpose of the IRVAPP Working Papers series is to promote the circulation of working papers prepared within the Institute or presented in IRVAPP seminars by outside researchers with the aim of stimulating comments and suggestions. Updated review of the papers are available in the Reprint Series, if published, or directly at the IRVAPP.

The views expressed in the articles are those of the authors and do not involve the responsibility of the Institute.

The second digital divide in Europe. A cross-national study on students' digital reading and navigation skills*

Davide Azzolini[a†], **Antonio Schizzerotto**[a,b]

[a] *FBK-IRVAPP*
[b] *University of Trento*

Abstract

We investigate gender, migration and social-background disparities in digital reading and navigation skills using PISA 2012 computer-based data from 16 European countries. In comparison to the extant research on information skills disparities, we provide three main improvements. First, we rely on objective measures of navigation skills taken from the log files of the digital reading computer tests. Second, we distinguish a mere quantitative measure of browsing behavior from a more qualitative one which includes individuals' evaluative skills. Third, we compare print and paper-based reading scores, this allowing to assess social disparities in digital competences net of traditional skills and thus providing conservative estimates of the digital gaps. Our results point to the key importance of traditional competences in accounting for the observed digital gaps. Yet, the interplay between digitalization and education inequality is heterogeneous. Boys and girls show markedly different approaches to online navigation: the former show more operational familiarity with technology and 'trial and error' approaches and girls being more prone to 'think and then click' approaches. Concerning the migrant/native gaps, the picture is one in which children of immigrants underperform on the digital reading test, but catch up with natives when traditional reading competences are held equal. Immigrants' children show slightly higher navigation skills that may be employed to compensate their linguistic difficulties. Finally, youths from privileged social backgrounds show higher digital skills that persist even net of traditional competences pointing out the possible existence of a cumulative effect of traditional and digital inequalities.

Keywords: digital divide; digital reading; education inequality; information literacy; log files; navigation skills

* The research benefited from EUN-Schoolnet support. We are also thankful to Francesco Avvisati (OECD) for useful advice provided for the analysis of the PISA computer-based data.

† Corresponding author: azzolini@irvapp.it - *Via Santa Croce 77, I-38122, Trento*

The second digital divide in Europe. A cross-national study on students' digital reading and navigation skills

1. Introduction

In the information age, equal access and use of information and communication technologies (ICTs) have become major policy objectives in most industrialized countries. Following years of considerable digital expansion, the "first digital divide"—i.e., the socioeconomic gap in physical access to ICTs— is nowadays considered as being shrinking almost everywhere or deemed to be closed in the nearest future (Castells 2001, Attewell 2001, Di Maggio *et al.* 2004; Van Dijk and Hacker 2003, Mossberger *et al.* 2003, OECD 2015a). The same holds true for gender disparities (Liff *et al.* 2004) and even more so when the focus is restricted to the younger generations (Livingstone e Helsper 2007), who have benefited from the digital expansion more than the adults.

Major concerns are now posed by the so called "second digital divide"—i.e., the socioeconomic gap in use and ability to take full advantage of ICTs (Attewell 2001). Given the increasing centrality of ICTs and the pervasiveness of digital information, proficiency in using new technologies has become a key asset for a full and active participation in economic, social and civic life (Di Maggio *et al.* 2004). Contrary to a popular belief that increased access to the Internet and related digital technology would lead *per se* to an improved access to information for all, the more mature the digital expansion, the more digital inequality becomes a matter of concern (Di Maggio *et al.* 2004, Hargittai 2010). Higher social classes may implement capital enhancing strategies to maintain or reinforce their advantaged position (Zillien and Hargittai 2009). If a reduction of access inequality is not accompanied by a reduction of inequality on how ICTs are used, digitalization could exacerbate inequality instead of reducing it (Attewell 2001, Di Maggio and Hargittai 2001, Van Deursen and Van Dijk 2014).

This paper aims at shedding light on the intersection between digital competences and educational inequality by focusing on youths' information literacy (Carrettero et al. 2017).[1] More precisely, the paper investigates youths' disparities in digital reading proficiency and online navigation skills. A number of recent studies have pointed out social disparities with respect to the frequency and the modalities in which young people use of the Internet and computers (OECD 2015a, Livingstone and Helsper, 2007, Notten *et al.* 2009, Ono and Zavodny 2007). Not much is known instead on the existence of disparities in young people's competences to fully exploit new technologies and the Internet to retrieve and process digital information (Van Dijk 2005, Warscharuer 2005, Gui 2007). As argued by Gui and Argentin (2011), studies in this field often lack standardized measurements and are based on small, local samples reducing the generalizability of the results. The few studies that attempt to investigate deeper the existence of social disparities on standardized measures of information skills show that having higher educated parents is positively associated with higher skills (Hargittai 2002, Gui and Argentin 2011). Gender differences on digital skills are found to be less pronounced (Hargittai and Schaffer 2006, Gui 2007), while very little is known on migrant/native digital literacy gap. Hargittai (2010) for example finds that students of Hispanic origin and African Americans in the US exhibit lower levels of knowledge of the Internet. One of the biggest shortcomings in existing studies is the impossibility of taking traditional reading skills under control when analyzing digital skills. The former vary

[1] We use competence, skill, proficiency as equivalent terms.

substantially across groups (OECD 2010) hence possibly leading to an overestimation of the digital skill gap, as the competences needed to read and understand texts are intrinsically linked to the competences needed to retrieve and process information online (OECD 2015a).

With this paper, we provide a cross-national analysis of the second digital divide with a specific focus on youths' digital reading and online navigation skills. We use 2012 data from the *Programme for International Student Assessment – Computer Based Assessment* (hereafter also PISA-CBA 2012) relative to 16 European countries. Beyond the wide geographical coverage and the large sample size, this data allows us to contribute to the literature with respect to at least three other important points. First, the data provides objective measures of pupils' online navigation behaviors that were taken while students filled in the digital reading computer tests and that are summarized in two indices of quantity and quality of the online browsing activity (OECD 2015a). Second, PISA-CBA 2012 also provides individual measures of students' standard print reading competences, this allowing an assessment of digital reading and navigation skills gaps keeping 'standard' reading competences under control. Third, we consider multiple forms of inequality. Beyond gender and social origins (as measured by parental education and occupation), we also look at the role played by migration background. The number of children of immigrants in European schools has been growing steadily in the past years, yet not much is known on their digital literacy.

The main empirical results of our analysis can be summarized as follows. First, we document the existence of inequality in digital competences along the lines of gender, migration background and social origins. Males, children of immigrants and children coming from less privileged family backgrounds exhibit lower digital reading competences. Second, we show that, net of print reading competences, the four individual attributes considered (gender, migration background, parental education and occupation) play a small role on students' digital reading, although in some countries differences due to social background are noticeable. Hence, inequality in digital reading is greatly attributable to students' development of standard reading competences and not to digital ones. Nonetheless, some significant disparities in digital reading persist across groups even net of print reading skills, this pointing to the existence of "digital-specific" gaps. Third, we highlight the possible existence of three distinct mechanisms to account for the relationship between digital skills and educational inequality. Regarding gender, boys show significantly higher navigation activity levels than girls, but when considering the quality of navigation, they fall behind them. This result point to the existence of gendered approaches to ICTs and the Internet, with boys appearing more technically capable to interact with computers but at the same time being more unfocused ("trial and error" approach), while girls appear to be more reflective ("think and then click" approach). These gendered approaches can be rooted in exposure to ICT at home and purposes of use. Boys start using computers at earlier ages, live in technology-richer households and use the Internet for gaming more than girls, who instead use ICTs at home more frequently for school-related tasks. Concerning the migrant/native gaps, the picture is one in which children of immigrants underperform natives in digital reading, but catch up with natives when traditional reading competences are held equal. Also, immigrants' children show slightly higher navigation skills that may be employed to compensate their linguistic difficulties. Finally, when analyzing social-background disparities, we find that these persists even after holding print reading competences fixed. This suggests that, contrary to the previous cases, traditional and new forms of social inequality cumulate, as more educated and affluent families can provide their children with technology-richer environments and a better supervision when using ICTs.

2. Literature review

In nearly all domains, today information is increasingly produced and spread by the means of technological devices rather than printed documents. This technological innovation demands changes in the set of skills needed to adequately retrieving and processing such information. Some authors refer to “informational skills” to define those skills needed to navigate, select, evaluate and reuse digital information (Warschauer 2002, van Dijk 2005, Gui 2007). Carretero and colleagues (2017) identify three sets of competences related to “information and data literacy”: (*i*) browsing, searching and filtering data, information and digital content; (*ii*) evaluating data, information and digital content; and (*iii*) managing data, information and digital content.

Possessing standard reading competences may not be enough to make individuals prepared for mastering digital information. Reading online texts requires new competences that go beyond traditional competences needed for printed documents (Noyes and Garland 2003, Mangen *et al.* 2013, OECD 2015a). First, online texts are characterized by a combination of text, images, animations, videos, etc. Also, hyper-links are often embedded in online texts, this creating non-sequential page structures. This complex configuration requires readers to possess adequate text-processing skills and ability to navigate through different texts (OECD 2015a). Second, web pages also require organisational and spatial skills, such as the ability to construct a mental representation of the structure of a website in order to move confidently across the different pages (OECD 2015a). Third, evaluative skills are also particularly important for online text forms. Individuals have to assess the relevance of information and the credibility of sources by using their prior experience and hints provided by the web pages (e.g., layout, poor grammar and spelling, link names, etc.). Hence, the use of digital information requires individuals to develop a set of specific competences that range from knowing the sources of information, knowing how to search and find the relevant information, being able to critically examining and assessing Internet information sources. In sum, digital reading competences involve both print reading skills and a set of new digital competences that can be labelled as “navigation skills” (OECD 2015a).

Disparities in access to ICT have been documented since the Eighties between genders, immigrant and social backgrounds both in the US and in Europe (Attewell 2001, Van Dijk and Hacker 2003, Di Maggio *at al.* 2004, OECD 2015a). In spite of the evidence that higher-status people use higher-quality digital equipment (Zillien 2006), the social gap in access to ICTs is no longer considered a major concern and it is expected to shrink significantly in the future, if only for the effects of digital saturation. However, social disparities in *how* ICTs are used are found in several contexts. In the US, for example, the patterns of ICTs use vary substantially across social classes, with children coming from less privileged social strata employing more time in unsupervised playing games at home, while affluent children enjoy educationally richer environments (Attewell 2001). Notten and colleagues (2009) using PISA 2003 data from 30 countries find that children of parents with higher education and occupational status use the Internet more often for informational and less often for entertainment purposes than less well educated adolescents. A similar conclusion is reached by Gui (2015) who, relying on a series of local surveys carried out on students in Northern Italy, finds that children of high-educated parents use the Internet more often to search information and are more likely to be supervised by their parents when using new technologies. Moreover, they run lower risk of excessive Internet consumption, which could have disruptive consequences on social activities and learning. Such systematic social disparities in how computers and the Internet are used could reflect ‘capital-enhancing’ strategies implemented by privileged groups to gain more benefits from digitalization (Di Maggio and Hargittai, Zilien and Hargittai 2009). It is well known that the family environment shape young people's resources and habits (Bourdieu 1984, Lareau 2003). High-educated parents are likely more experienced user of new technologies than less-educated and their homes are likely to be more equipped with ICT. Moreover, high-status parents can have

more positive attitudes and be able to better supervise their children in the use of computer and the Internet than low-status parents with poor digital competences (Livingstone and Helsper 2007).

Beyond social disparities, a large number of studies also point out recurrently the existence of gendered patterns in the use of ICT. Boys are found to start using the computer at earlier ages relative to girls (OCED 2015a). The former also show higher use of the Internet and higher likelihood of using it for gaming or informational purposes (Livingstone and Helsper 2007, Notten *et al.* 2009). Concerning the role of migrant or ethnic minority status, instead, there is some evidence that non-native young people may show even higher and more diversified Internet and computer use than natives. Hargittai (2010) studies the web activities of a class of first-year college students and finds that Asian students engage in more activities on the Web regardless of available resources and experience. Gui (2015) finds that immigrant students use the Internet for a wider range of purposes and produce more digital content online than natives in Italy.

When shifting our attention from digital use to digital competences, our knowledge on social disparities is weaker. As pointed out by Gui and Argentin (2011), social disparities shall be more pronounced when it comes to 'informational' skills, as the latter depend more on family's cultural background, than simply operational skills needed to use a computer and software (van Dijk 2005). Empirical studies provide consistent evidence that education level appears to be the most relevant factor connected to differences in digital skills such as skills in solving search tasks online (Hargittai 2002, Gui 2007). Hargittai (2010) finds that higher levels of parental education, being a male, and being white or Asian American are associated with higher levels of Web-use skill, while students of Hispanic origin, and African Americans exhibit lower levels of knowledge of the Internet. Gender differences appear relatively small on online task tests while seem to be larger when it comes to attitudes or self-perceived skills (Hargittai 2002, Hargittai and Shafer 2006).

3. Data and methods

3.1 Data

We use PISA 2012 data relative to 16 European countries. The dataset comprises different archives. The most relevant to our purposes are: (*i*) the paper-based assessment (PBA) file from which we derived individuals' scores on the paper-based reading test; (*ii*) the student questionnaire file containing information on pupils' background, attitudes, behaviors, as well as several items on ICT access and use, both at home and at school; (*iii*) the computer-based assessment (CBA) file containing the digital reading test scores (DRA); and (*iv*) the PISA-CBA "Log files" which track students' navigation behaviors when taking the CBA-DRA test.

PISA 2012 was administered differently in the participating countries. While all countries administered the PBA, only a subset of countries (44) administered the CBA, and among the latter only some countries (32) administered the CBA-DRA module. Moreover, within these 32 countries, not all students took all the three CBA tests (i.e., in addition to DRA, the two other computer tests were on mathematics and problem solving), but these were randomly administered to different subsamples of students.[2]

Our analysis includes 15 European Union (EU) countries (Austria, Belgium, Germany, Denmark, Spain, Estonia, France, Hungary, Ireland, Italy, Poland, Portugal, Slovakia, Slovenia, Sweden), and Norway. The remaining European countries are not included in the analysis as they did not administer the digital reading tests. More precisely, seven countries (Luxembourg, Latvia, Lithuania, Romania, Cyprus, Greece, and Malta) did not take part to

[2] See OECD (2014) for further details on the survey design.

the PISA-CBA tests; and six other countries (Bulgaria, Czech Republic, Finland, Great Britain, Croatia, and the Netherlands) administered only the problem-solving module. The analyses presented in this report are based on the "log files" sample - i.e. 24,851 students who took the DRA test and for those the information on navigation skills is available.[3]

3.2 Variables

Our main dependent variables are digital reading scores and navigation skills. For the digital reading score, as well as for the print reading one, the values are standardized across all countries in the first year when the test is administered as a major subject (the pooled OECD mean score is 500, with a standard deviation of 100), making cross-country and cross-test comparisons more meaningful.

Students' online navigation skills are captured by two distinct indices that measure pupils' behaviors when taking the PISA computer-based digital reading test. The CBA system collected behavioral data such as time spent on task, sequence of pages visited, and use of stimulus elements like 'drop down' menus. The CBA log files contain a measure of the "length of navigation sequences" (i.e., number of movements/steps between different pages) as well as qualitative information of each step taken: task-relevant step (from and to a relevant page); misstep (from a relevant to a non-relevant page); correction (from a non-relevant to a relevant page); task-irrelevant step (from and to a non-relevant page). This information was used to construct the two indices used in the analysis: (i) the *index of overall browsing activity*: given by the total number of steps; (ii) the *index of targeted navigation*: given by the total number of task-relevant steps subtracted by missteps and task-irrelevant steps. Since the first index reports a quantitative measure of students' browsing in during the examination, it could not only indicate digital proficiency but also unfocused reading and inefficiency. The second index instead provides a measure of the quality of students' navigation. By taking "right" and "wrong" steps into account, this index captures students' reflective behavior and evaluative skills in answering the computer tests ("think, then click" approach) (OECD, 2015a). The values of the two indices range between 0 and 100—with 0 indicating no (targeted) activity and 100 indicating maximum (or, maximally targeted) activity—and, to account for the different difficulty levels, reflect the percentile score given by the rank of the student among all students who were administered the same digital reading questions.

As for the main independent variables employed in the analyses, they were coded as follows. Gender takes value 1 for males and 0 for females. Migrant background takes value 1 for those students whose both parents were born abroad and 0 otherwise. Parental education takes value 1 for students who have at least one parent with ISCED5 (tertiary) education and 0 otherwise. Parental social position is measured by means of ISEI (International Socio-Economic Index of Occupational Status) scores. ISEI is a standardized measure of the level of social prestige associated to every occupation listed in the 1988 version of the International Standard Classification of Occupation (ISCO88). In the analyses, the ISEI index is recoded as a dummy variable taking value 1 for students with one parent scored in the ISEI index equal or above the 75th percentile and 0 for their counterparts with both parents' ISEI score below the 75th percentile. The percentile distribution is computed country by country to reflect individuals' relative position within her country. A number of other variables concerning pupils' access to and use of ICTs at home and at school are also considered and described in section 5.

[3] Missing values on three of the four main independent variables (i.e., immigrant background, parental education and occupation) range between 4 and 5% of the sample. These observations are included in all analyses as 'missing categories'.

3.3 Methods

The analyses are based on a series of linear regression models which include the five plausible values of the paper and digital reading scores. To account for the survey design, all analyses are estimated using final student weights and the 80 replicate weights provided by PISA to retrieve correct estimates of standard errors. More precisely, we apply the balanced repeated replication (BRR method with a Fay adjustment factor of 0.5. We use the STATA routine *repest* developed by Avvisati and Keslair (2016).

4. Gaps in digital reading

In this section, we investigate how gender, migration background and social origins contribute to disparities in digital reading competences and whether these gaps persist even holding print reading skills constant across groups. Table 2 shows mean differences in paper and digital reading scores across groups. The "gender gap" is given by boys' average score minus the girls' scores, hence the negative values indicate boys underperforming girls. The migrant background gap is obtained subtracting natives' scores from the children of immigrants' ones, with negative values indicating children of immigrants performing worse than natives. The parental education (occupation) gap is the difference between the score of students whose parents are highly educated (employed in socio-economically higher ranked occupations) and those whose parents are lower educated (have lower ranked occupations): positive values on these two gap estimates indicate that students with more privileged family backgrounds perform better. Table 2 shows that gender- and migration-background gaps are smaller in digital vs print reading. This is particularly true for the gender gap, which drops from more than 41 points (two fifths of a standard deviation) to 25 points. When considering the two indicators of social origins instead, it appears that the gaps are substantially and statistically equivalent on the two tests.

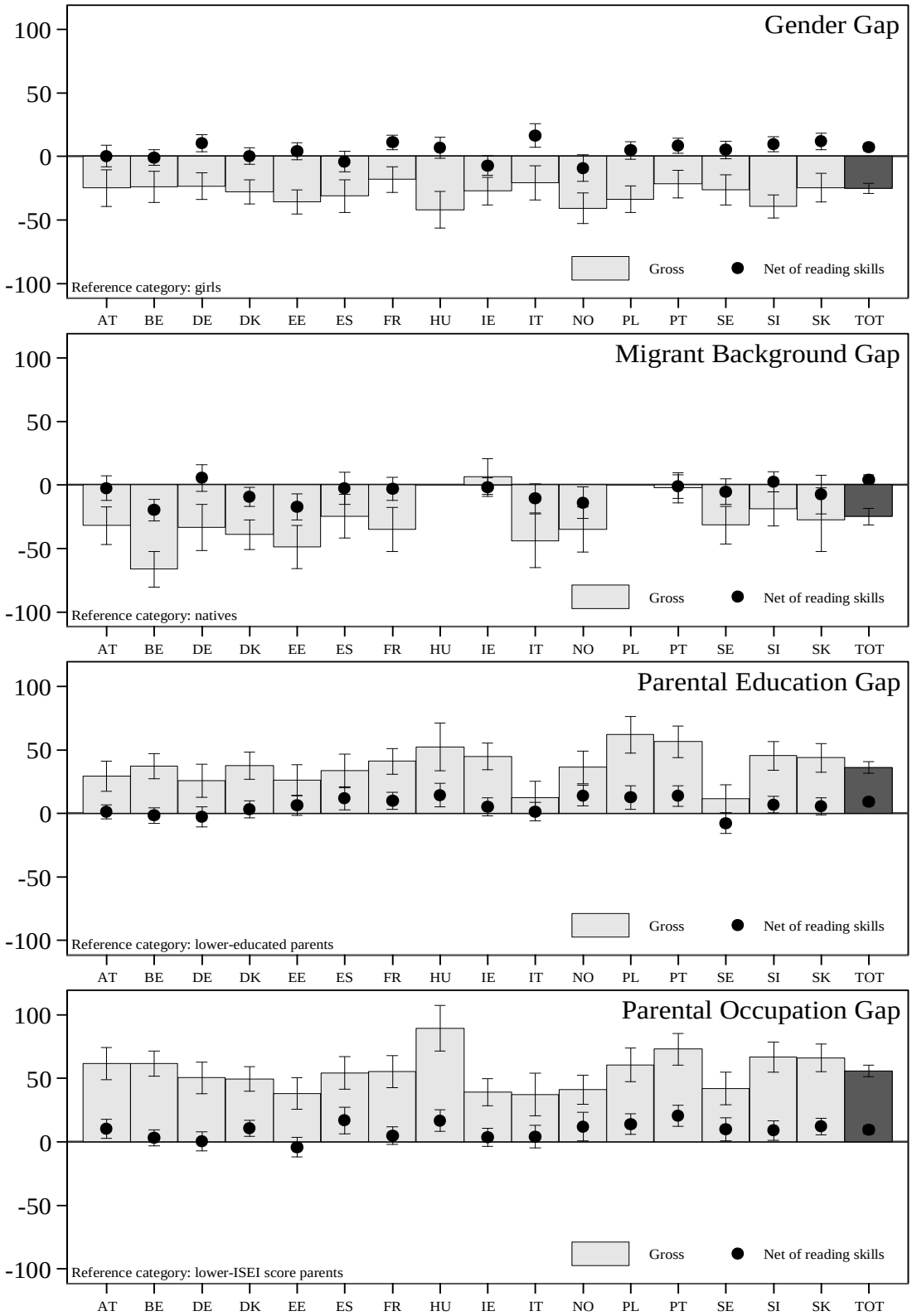
Table 2 Comparison of the gap estimates on print and digital reading tests (N=24,851)

Gap	Print Reading			Digital Reading		
	Diff	95% CI		Diff	95% CI	
		Low	Up		Low	Up
Gender (Boys – Girls)	-41.7	-45.7	-37.7	-25.1	-29.0	-21.3
Migrant background (Non-natives - natives)	-37.1	-44.3	-29.9	-24.9	-31.6	-18.3
Parental education (Higher – lower educated)	34.9	29.5	40.3	36.2	31.7	40.7
Parental occupation (Higher – lower ISEI score)	59.9	54.5	65.4	55.8	51.3	60.4

Source: Own elaborations based on PISA-CBA 2012 data.

Note: Linear regression estimates obtained using the five plausible values, the final student weights and the 80 replicate weights. Observations with missing values on the variables migration background, parental education and ISEI score are included in the analysis as 'missing category' but not shown.

Figure 1 Gross and net (of print reading skills) digital reading gaps (N=24,851)



Source: Own elaborations based on PISA-CBA 2012 data.
Note: Linear regression estimates obtained using the five plausible values, the final student weights and the 80 replicate weights. Lines are 95% confidence intervals. Observations with missing values on the variables migration background, parental education and ISEI score are included in the analysis as 'missing category' but not shown. Migration background gap estimates are not shown for Poland and Hungary because of too small numbers of children of immigrants in the sample.

To investigate deeper into the digital reading gaps, a series of linear regression models is estimated both on the pooled set of countries and within each distinct country. The results are presented in Figure 1, which shows gross mean differences (grey bars) on digital reading and the net mean differences (black circles)—the latter estimated after making “print” reading skills equal across groups. The vertical lines indicate the 95% confidence intervals around the estimates.

The overall picture provided by Figure 1 suggests that net differences are much smaller than gross ones (the circles are always closer to the “zero line”). This means that most of the differences observed in digital reading skills are due to group differences in print reading. However, there are important variations in the results depending on the specific ascriptive factors considered.

When considering gender, males are clearly underperforming girls on the digital reading test in all countries. However, this gap is reversed, with boys outperforming girls, when holding pure reading skills constant across genders. This points to a key role played by digital skills in accounting for gender gaps in online reading. This pattern holds true for nearly all countries. In some countries net gender gaps are virtually zero, while in no country do boys underperform girls once print reading skills are controlled for.

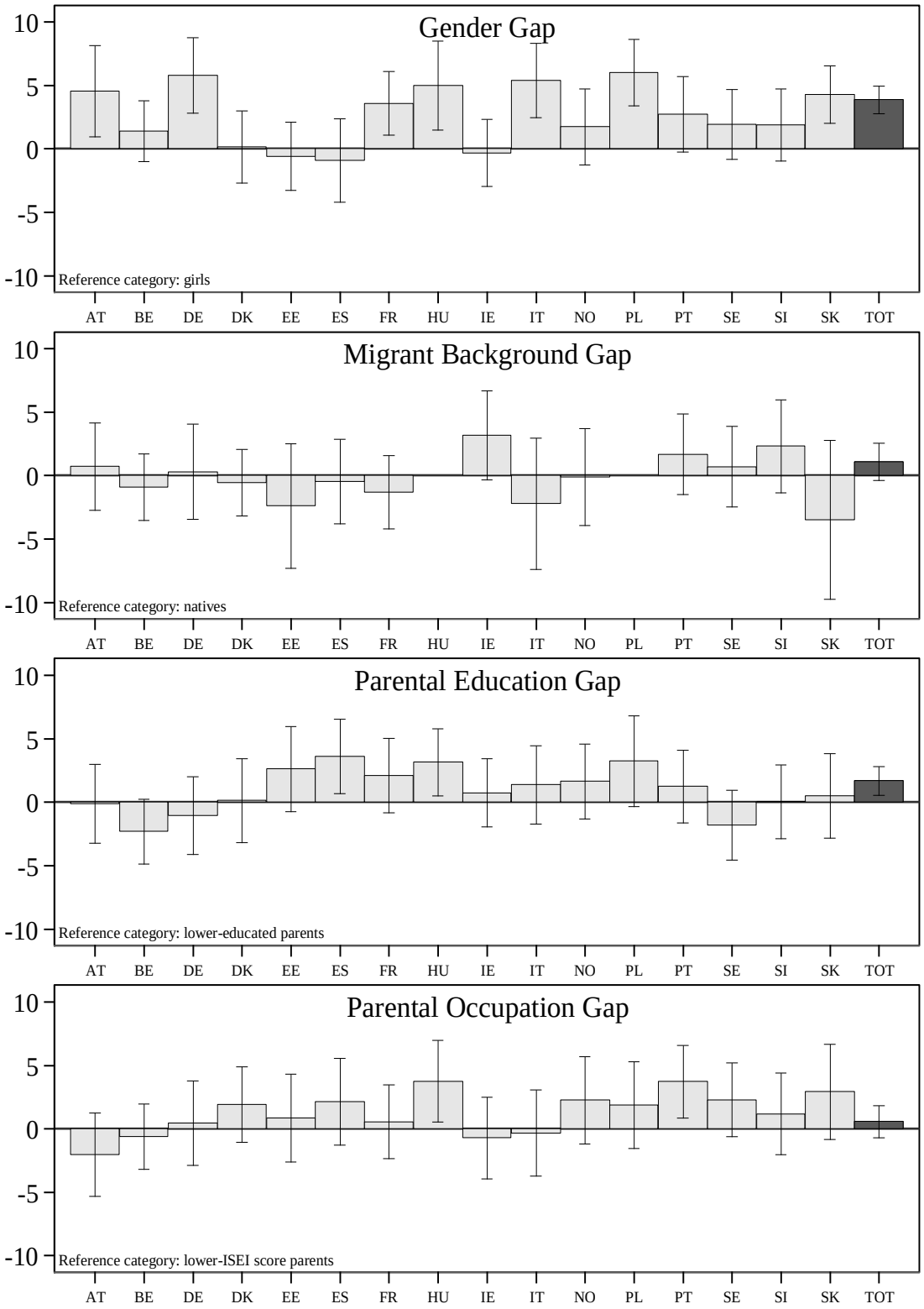
Turning now to migration background, children of immigrants underperform natives on the digital reading test, but once pure reading skills are held fixed, the gap vanishes. This result holds true for almost all countries. In some countries, however, a migrant ‘penalty’ persists (Belgium, Denmark, Estonia, and Norway), suggesting that immigrants’ children face a double disadvantage: not only in host-country language acquisition but also in digital skills’ development.

Finally, in all countries children of high-educated parents and parents with high occupational status show higher performance on the digital reading test relative to their counterparts with lower educated and less socio-economically advantaged parents. This pronounced social-background gap in digital reading is almost entirely explained by print reading skills. Overall, digital reading gaps between social groups are reduced by 80-90% when print reading skills are modelled, suggesting that social inequality is more pronounced on basic traditional competences than on new digital competences. However, even if small, residual significant gaps persist in many countries (Austria, Denmark, France, Hungary, Norway, Poland, Portugal, Spain, Sweden, Slovenia, Slovakia), suggesting that the “advantage” of children of more privileged social backgrounds is not always limited to standard competences but also extends to digital ones.

5. Gaps in navigation skills

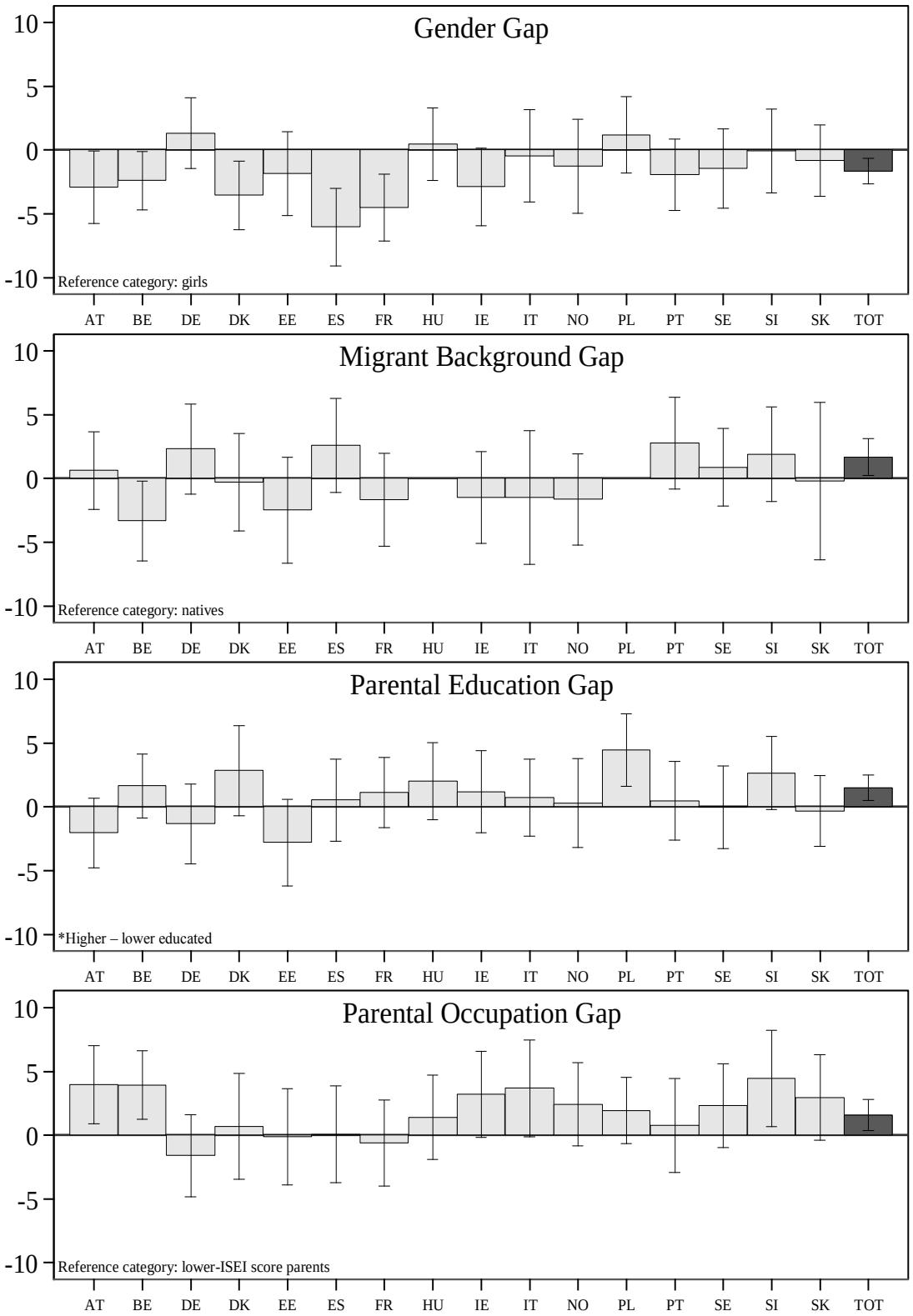
After establishing the extent to which gender, migration background and social origins affect students' digital reading proficiency, this section investigates the role played by the same characteristics with respect to students’ ability in navigating the web to answer the digital reading test questions. Figures 2 and 3 show the results obtained with a series of linear regressions aimed at testing the existence of these gaps on the two indices (overall browsing activity index and targeted navigation index) net of print reading skills. Results show that boys clearly outperform girls on the overall browsing activity index, while they underperform them on the targeted navigation index. This result suggests that boys are more active and quick in taking action when answering digital reading tasks than girls (“trial and error approach”) while girls are more reflective and read more carefully before taking action (“think before clicking approach”). Results are fairly consistent across countries: in a number of countries, gender differences are not statistically significant, but "opposite" effects are never detected. In no country do girls show higher browsing activity or lower targeted navigation than boys.

Figure 2 Overall Browsing Activity Gaps, net of print reading skills (N=24,851)



Source: Own elaborations based on PISA-CBA 2012 data and CBA log files.
Note: Linear regression estimates obtained using the five plausible values, the final student weights and the 80 replicate weights. Lines are 95% confidence intervals. All models control for students' print reading skills. Observations with missing values on the variables migration background, parental education and ISEI score are included in the analysis as 'missing category' but not shown. Migration background gap estimates are not shown for Poland and Hungary because of too small numbers of children of immigrants in the sample.

Figure 3 Targeted Navigation Gaps, net of print reading skills (N=24,851)



Source: Own elaborations based on PISA-CBA 2012 data and CBA log files.
Note: Linear regression estimates obtained using the five plausible values, the final student weights and the 80 replicate weights. Lines are 95% confidence intervals. All models control for students' print reading skills. Observations with missing values on the variables migration background, parental education and ISEI score are included in the analysis as 'missing category' but not shown. Migration background gap estimates are not shown for Poland and Hungary because of too small numbers of children of immigrants in the sample.

In regard to migrant/native differences, evidence is weaker, partly as a consequence of the small number of observations for students with an immigrant background. If any, a small advantage of children of immigrants is detected with respect to the targeted navigation index. But, overall, migrant/native differences in navigation behaviors are close to zero in nearly all countries.

With respect to the role played by social background, Figures 2 and 3 report substantially small positive effects of parental education and parental occupation. In many countries, differences between groups are not significantly different from zero, but when significant, they always point to a positive association between parental education and parental occupation on both indices. In none of the considered countries do students from lower social backgrounds outperform their counterparts from more privileged backgrounds on the digital skills indices considered. Hence, students from poorer social backgrounds show lower navigation skills even net of print reading skills.

The digital skills gaps presented in Figures 2 and 3 could be partially related to group differences with respect to students' familiarity with ICTs, which could arise as a consequence of different ages of ICT first exposure, variation in home availability of ICTs, as well as different patterns and intensity of use of ICTs and the Internet. Table 3 shows a comparison of several indicators across genders, migration and social backgrounds.

In support of the interpretation of the results presented so far, boys happen to have been exposed earlier than girls to ICTs: 80% of boys used the computer or the internet before the age of 9, while only 73% of girls do so. Also, males state that they have more ICT resources at home and use ICTs for doing homework less frequently than girls while using the Internet more often for searching information and gaming. All together, these figures support the hypothesis that boys have more operational skills to use digital resources and this could account for their higher browsing activity when taking the PISA digital reading tests. Interestingly, boys also display more negative attitudes towards the use of ICTs in their education, and this could be reflected in lower performance on the reading test.

Overall weaker differences exist between immigrants' children and native students, although the former are exposed later to ICTs; use less the Internet for gaming and have more negative attitudes with respect to the role of ICTs in education. When looking at social background, a clear and systematic divide comes up, with students from parents with higher education and occupational status having been exposed earlier to ICTs and living in homes that are better technologically equipped. Also, they report using ICTs and the Internet more for doing homework and keeping up to date with news and less for gaming. Furthermore, they show more positive attitudes towards the use of ICT in education. In sum, it seems that children of more privileged backgrounds enjoy a technology-richer environment as well as a family context in which children use less often ICTs in a completely unsupervised way.

Table 3 Access and use of ICT and the internet and attitudes toward ICTs, by gender, migration background and social origins

	Used ICTs before the age of 9 [b] (%)	ICT availability at home [a, c, d]	ICT Use at Home for School-related tasks [a]	Internet use for gaming [a, d]	Internet use for news [a, d]	Internet use for social networks [a, d]	Limitations of the Computer as a Tool for School Learning [a]
Gender							
Females	**.73**	**-.04**	**.02**	**-.54**	**-.16**	-.04	**-.13**
Males	**.80**	**.18**	**-.03**	**.41**	**.02**	-.01	**<-.01**
Migration background							
Natives	**.78**	.06	<.01	**-.05**	-.06	-.02	**-.08**
Non natives	**.73**	.11	-.02	**-.12**	-.13	-.01	**.02**
Parental education							
Up to secondary	**.73**	**-.05**	**-.04**	-.05	**-.10**	-.02	**-.02**
Tertiary	**.82**	**.23**	**.07**	-.08	**-.02**	-.02	**-.12**
Parental occupation							
Low ISEI score	**.76**	**.04**	**-.02**	**-.04**	**-.09**	**<-.01**	**-.03**
High ISEI score	**.83**	**.21**	**.08**	**-.13**	**.04**	**-.09**	**-.18**

Source: Own elaborations based on PISA-CBA 2012 data.

Note: Estimates are obtained using the five plausible values, the final student weights and the 80 replicate weights. Statistically significant differences (p<.05) in bold. Observations with missing values on the variables migration background, parental education and ISEI score are included in the analysis as 'missing category' but not shown. [a] Average value of an Index with mean 0 and standard deviation 1 (at the OECD level). [b] The age of nine is chosen for it is a critical period in childhood, when the majority of children make the transition from “learning to read” to “reading to learn”. [c] The index measures households’ possession of different objects such as a desktop or portable computer, an internet connection, a printer, a cell phone, etc. [d] In France, this information was not collected.

Finally, individual and school-level determinants of students’ development of navigation skills are jointly investigated by the means of a series of multiple linear regression models estimated on the pooled set of countries (Table 4). The first model (M1) includes the four students’ characteristics (i.e., immigrant background, sex, parental education and occupation). Model 2 (M2) includes ICT access and use indicators examined in Table 3. Model 3 (M3) includes indicators of the school quality (such as student-teacher ratio and school average print reading score) and the school’s availability and use of ICTs (such as the PC-student ratio in the school and an index measuring the intensity to which internet is used in schoolwork, as reported by school principals). Model 4 (M4) adds individual score in print reading and the grade attended relative to the country modal grade attended by same-aged students. To account for factors that vary across countries, all models include country fixed effects.

The results of Model 1 inform about the fact that boys, migrants and young people from lower social status have lower navigation skills. This result holds true for both indices, although the gender gap, in line with what discussed above, is more pronounced on the second one (targeted navigation skills).

Table 4 Individual and school level determinants of student navigation skills (N=18,247)

	Overall browsing activity								Targeted navigation							
	M1		M2		M3		M4		M1		M2		M3		M4	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Non native (ref. native)	**-2.85**	**1.06**	**-2.76**	**1.04**	-0.76	0.99	1.55	1.00	**-3.03**	**0.85**	**-2.97**	**0.85**	-1.07	0.80	1.17	0.81
Male (ref. female)	**-1.93**	**0.69**	**-2.82**	**0.70**	**-1.96**	**0.67**	**1.62**	**0.65**	**-5.76**	**0.71**	**-6.11**	**0.74**	**-5.30**	**0.67**	**-1.66**	**0.69**
High-educated parents (ref. low/medium)	**2.85**	**0.81**	**2.89**	**0.83**	1.00	0.82	0.36	0.80	**2.09**	**0.69**	**2.08**	**0.70**	0.31	0.72	-0.32	0.75
High-ISEI parents (ref. Low)	**6.06**	**0.79**	**5.81**	**0.79**	**2.02**	**0.75**	-0.01	0.78	**7.08**	**0.81**	**6.85**	**0.80**	**3.25**	**0.80**	1.23	0.84
Used ICTs before the age of 9			**3.02**	**0.93**	**1.82**	**0.90**	0.48	0.89			**3.30**	**0.81**	**2.15**	**0.80**	0.79	0.78
ICT availability at home			**-1.55**	**0.43**	**-1.46**	**0.41**	-0.69	0.40			**-0.98**	**0.41**	**-0.89**	**0.36**	-0.09	0.35
ICT Use at Home for School-related tasks			-0.81	0.47	**-1.21**	**0.41**	-0.31	0.37			**-0.98**	**0.40**	**-1.32**	**0.35**	-0.39	0.32
Internet use for gaming			0.62	0.37	**1.28**	**0.38**	**1.02**	**0.36**			0.02	0.38	0.64	0.37	0.38	0.36
Internet use for news			**1.91**	**0.41**	**1.29**	**0.39**	0.43	0.39			**1.26**	**0.44**	0.65	0.42	-0.22	0.41
Internet use for social networks			-0.27	0.37	-0.05	0.34	0.04	0.32			0.11	0.34	0.34	0.34	0.43	0.32
Student-teacher ratio					0.04	0.11	0.05	0.12					-0.08	0.10	-0.06	0.10
PC-student ratio					0.72	0.48	0.65	0.50					1.02	0.59	0.96	0.61
Index of internet use in school					-0.10	0.47	0.02	0.45					-0.67	0.39	-0.56	0.38
School average print reading score					**0.14**	**0.01**	**0.04**	**0.01**					**0.14**	**0.01**	**0.04**	**0.01**
Print reading score							**0.11**	**0.01**							**0.11**	**0.00**
Grade (relative to country mode)							**2.24**	**0.65**							**1.43**	**0.56**
R-squared	0.07		0.08		0.16		0.24		0.05		0.06		0.13		0.21	

Source: Own elaborations based on PISA-CBA 2012 data. Note: Linear regression estimates obtained using the five plausible values, the final student weights and the 80 replicate weights. All models control for country fixed effects. Statistically significant ($p < .05$) coefficients in bold. Missing information on the use of ICT and the Internet are dropped. France is entirely dropped from the analysis because of missing information on ICT and internet use questions. Observations with missing values on the variables migration background, parental education and ISEI score are included in the analysis as 'missing category' but not shown.

M2 adds individual information on the degree to which pupils have access to (and use) digital technologies. The inclusion of these variables hardly changes the coefficients estimated with M1. If any, a slight—yet statistically insignificant—increase in the gender gap is detected. This may be explained by the fact that boys have earlier access to ICTs than girls and use the Internet more often, especially for retrieving information and for gaming-related purposes (as shown in Table 3). Early access to ICT is indeed strongly and positively correlated with both indices. Interestingly, the index of home digital equipment is negatively correlated with both navigation scores, suggesting that once family background is controlled for, the mere possession of technological devices at home is not conductive of higher digital competences.[4]

M3 adds school-level controls. PC-student ratio and Internet use at school do not seem to be relevant for students' development of navigation skills, net of individual characteristics. However, the "quality" of the school attended (measured with the school's average print reading score) is positively and significantly correlated with both indices. Hence, students attending good quality schools are also more likely to exhibit better digital skills. This could explain the reduction in the social origin and immigrant background gaps and point to the salience of unequal sorting of students into schools based on ascriptive characteristics or residential segregation.

M4 adds the individual standard reading score and the relative grade attended. The two variables, paper reading especially, are strongly and positively associated with the two indices and the regressions' goodness of fit (R-squared statistics) improves substantially, confirming the importance of basic standard skills for developing information skills. The inclusion of these variables also change the gap estimates of models M1-3. Boys perform better on the first index and worse on the targeted navigation one and children from more privileged social backgrounds show no advantage over their counterparts from lower social strata. Concerning the role of the different ICT uses, the only still significant indicator is the one related to gaming when considering overall browsing activity. This suggesting that gaming could increase youths' skills in operating the computer over and beyond standard competences.

In sum, net of all individual and school factors, only the gender gap persists significant, confirming boys' higher navigation activity and girls' more focused navigation behavior. In turn, immigrant- and social background gaps are accounted for by individual and school characteristics.

6. Discussion and conclusions

Access to information and the possession of the adequate information competence have become key assets in economic and social life and may give rise to a new source of stratification in the society. The "knowledge gap" is far from being a new concept (Tichenor, Donohue, and Olien 1970), but it may achieve increased importance in the digital era (Bonfadelli 2002). The pervasiveness of digital technology and ICTs has reached unprecedented levels in human life history. Information is more and more accessible on the Internet and digital devices. According to most studies, the so called first digital divide (i.e., the socioeconomic gap in access to technology) is decreasing and deemed to disappear alongside the maturation of the digitalization process. Yet, despite the popular assumption that digitalization and the Internet would lead to an overall and equal increase of knowledge in the population, recent empirical evidence points to substantial social disparities in how

[4] Additional analyses including a quadratic term of the ICT possession index to capture possible nonlinearities in the relationship with navigation skills partly confirm this interpretation, as the negative association is only detected at high values of the index. Too much technology around in the home environment does not translate into better use and higher skills.

people use new information technologies (the so called second digital divide). Searching, processing and evaluating information in a digital world requires a new set of skills (i.e., information skills). Higher social classes may employ their higher resources and cultural capital to take more advantage from new technologies and hence achieve faster and better access to information.

This paper provides a comparative analysis of digital reading and navigation skill gaps across gender, migration background and social origins in 16 European countries. In addition to most studies on the topic, it uses objective measures of navigation skills, which allow investigating both a quantitative measure of browsing behavior and a qualitative index, which includes an evaluative component in browsing activity. Also, it provide estimates of digital reading and navigation skills gaps, net of print standard reading competences, which has rarely been done in the literature, and that provides digital gaps estimates which are less affected by omitted variable bias.

The novel contribution of this study lies in the finding that the largest part of these digital reading gaps is due to differences in print reading skills. Hence, inequality in digital reading competences is strongly attributable to students' development of standard competences and not to digital ones. Yet, some significant gaps in digital reading persist across groups even net of print reading skills and highlight the possible existence of three distinct mechanisms underlying digital inequality. The case of gender differences is remarkable. Boys underperform girls on the digital reading test but, when print reading skills are equalized across genders, their "disadvantage" turns into an "advantage", suggesting that they are more proficient than girls with computers and ICTs. Boys do indeed show significantly higher navigation activity levels than girls, but when considering the quality of navigation, they fall behind them. The latter result points to the possible existence of gendered approaches to ICTs and the Internet, with boys being, on average, more "technically" capable to interact with computers but at the same time being also less focused when navigating; whereas girls appear to be more reflective. In other words, boys seem to be more prone to a "trial and error" approach while girls are more keen to a "think, then click" one. These gaps may arise from the existence of gender differences in ICT familiarity, which this report has documented on several factors like age of first exposure to ICT, frequency and patterns of use of computers and the internet, as well as attitudes towards the use of ICT in education.

Digital reading gaps between natives and children of immigrants are close to zero. However, the fact that migrant/native gaps are smaller when reading tests are digitally administered rather than on paper leads to the hypothesis that children of immigrants deploy their navigation skills to compensate, at least partially, for their lower mastery of the host country language. More studies are needed to shed further light on the factors that produce or alleviate the reported digital reading gaps across genders, migration- and social backgrounds.

When considering social background gaps in digital reading, these are strongly attributable to print reading competences. However, in some countries there is evidence of a small advantage of children of more privileged social backgrounds that is not limited to standard competences but that also includes digital ones. Moreover, children of more affluent or more educated families also show higher navigation skills. Hence, digital skills might represent a further area of concern for educational inequality, which might cumulate on traditional aspects.

All together, these results add to the debate on the link between digitalization and educational inequality. While there is an abundant literature on social inequality with respect to traditional measures of educational performance (e.g., reading and mathematics skills above all), our knowledge on the existence of social disparities on digital and information skills is still scarce. The broader question is whether the widespread use of technology in teaching and the increased number of competence domains that require an intense use of ICTs amplify or reduce educational inequalities. Answering to such a question is beyond the

scope of this paper, but the findings presented here seem to suggest that the answer depends on the specific source of inequality taken into consideration. When it comes to gender and migration background gaps, disparities tend to be larger in traditional competences and less relevant in digital ones. However, when looking at social background inequality, it seems that the digitalization does not reduce disparities and, if any, it increase them, pointing to a cumulative effect of traditional and new forms of educational inequality. More research is needed to investigate further this point as well as to shed light on the mechanisms underlying the transmission of these skills.

References

Attewell P (2001) Comment: The first and second digital divides. *Sociology of education* 74(3): 252-259.

Avvisati F and Keslair F (2016) REPEST: Stata module to run estimations with weighted replicate samples and plausible values. *Statistical Software Components*.

Biagi F and Loi M (2013) Measuring ICT use and learning outcomes: Evidence from recent econometric studies. *European Journal of Education* 48(1):28-42.

Bonfadelli H (2002) The internet and knowledge gaps. A theoretical and empirical investigation. *European Journal of Communication* 17: 65–84

Bucy EP (2000) Social access to the internet. *Harvard International Journal of Press/Politics* 5: 50–61.

Castells M (2001) *The Internet Galaxy: Reflections on the Internet, Business, and Society*. New York: Oxford University Press.

Carretero S, Vuorikari R. and Punie Y (2017) *DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and examples of use* (No. JRC106281). Joint Research Centre (Seville site).

DiMaggio P and Hargittai E (2001) From the 'digital divide' to 'digital inequality': Studying Internet use as penetration increases. *Princeton: Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University 4*(1): 4-2.

DiMaggio P, Hargittai E, Coral C and Shafer S (2004) *Digital Inequality: From Unequal Access to Differentiated Use*. in K. Necherman (eds.) *Social Inequality*. New York: Russell Sage Foundation: 355–400.

Gui M (2007) Formal and substantial internet information skills: The role of sociodemographic differences on the possession of different components of digital literacy. *First Monday* 12(9). Available at:http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2009/(retrieved in May 2017).

Gui M and Gianluca A (2011) Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. *New media & society* 13(6): 963-980.

Hargittai E (2002) Second-level digital divide: Differences in people's online skills. *First Monday* 7(4). Available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/942/864/ (retrieved in May 2017).

Hargittai E and Shafer S (2006). Differences in actual and perceived online skills: The role of gender. *Social Science Quarterly* 87(2): 432-448.

Hargittai E and Hinnant A (2008) Digital inequality differences in young adults' use of the Internet. *Communication Research* XXXV(5): 602-621.

Hargittai E (2010) Digital na(t)ives? Variation in internet skills and uses among members of the "net generation". *Sociological inquiry* 80(1): 92-113.

Liff S, Shepherd A, Wajcman J, Rice R and Hargittai E (2004) An evolving gender digital divide? *OII Internet Issue Brief* 2.

Livingstone S and Helsper E (2007) Gradations in digital inclusion: Children, young people and the digital divide. *New media & society* 9(4): 671-696.

Mangen A, Walgermo BR and Brønnick K (2013) Reading linear texts on paper versus computer screen: Effects on reading comprehension, *International Journal of Educational Research* 58: 61-68.

Mossberger K, Mary KMCJT, Tolbert CJ and Stansbury M (2003) *Virtual inequality: Beyond the digital divide*, Georgetown, Georgetown University Press

Notten N, Peter J, Kraaykamp G and Valkenburg PM (2009) Research note: Digital divide across borders—A cross-national study of adolescents' use of digital technologies. *European Sociological Review* 25(5): 551-560.

Noyes JM and Garland KJ (2003) VDT versus paper-based text: Reply to Mayes, Sims and Koonce, *International Journal of Industrial Ergonomics* 31(6): 411-423.

OECD (2015a) *Students, Computers and Learning: Making the Connection*, OECD Publishing.

OECD (2010b) *PISA 2009 Results: Overcoming social background. Equity in learning opportunities and outcomes (Vol. II)*. Paris: OECD Publishing.

OECD (2015b) *Education at a glance 2015. OECD Indicators*, OECD Publishing.

OECD (2010) *PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes (Volume II)*, OECD Publishing.

OECD (2014) *PISA 2012 Technical Report*, OECD Publishing.

Ono H and Zavodny M (2007) Digital inequality: a five country comparison using microdata. *Social Science Research* 36: 1135–1155

Tichenor PJ, Donohue GA and Olien CN (1970) Mass media flow and differential growth in knowledge. *Public opinion quarterly* 34(2): 159-170.

Van Deursen AJ and Van Dijk JA (2014) The digital divide shifts to differences in usage. *New media & society* 16(3): 507-526.

Van Dijk JA and Hacker K (2003) The digital divide as a complex and dynamic phenomenon. *The information society 19*(4): 315-326.

Van Dijk JA (2005) *The Deepening Divide: Inequality in the Information Society*, SAGE Publications.

Warschauer M (2002) "Reconceptualizing the digital divide," *First Monday*, volume 7, number 7 (July), at http://www.firstmonday.org/issues/issue7_7/warschauer/, accessed May 2017

Zillien N and Hargittai E (2009) Digital distinction: Status-specific types of Internet usage, *Social Science Quarterly* XC(2): 274-291.