MEDIA@LSE Electronic Working Papers

Editors: Professor Robin Mansell, Dr. Bart Cammaerts

No. 15

# How deep are different forms of digital skills divide among young people?

## Results from an extensive survey of 1000 northern-Italian high school students

**Marco Gui**
**University of Milano-Bicocca**

**and**

**Gianluca Argentin**
**University of Milano-Bicocca**

**Marco Gui (marco.gui@unimib.it)** is a post-doc fellow at the University of Milano-Bicocca, Italy

**Gianluca Argentin (g.argentin@campus.unimib.it)** is a Ph.D Student at the University of Milano-Bicocca, Italy

Published by Media@LSE, London School of Economics and Political Science ("LSE"), Houghton Street, London WC2A 2AE. The LSE is a School of the University of London. It is a Charity and is incorporated in England as a company limited by guarantee under the Companies Act (Reg number 70527).



ISSN 1474-1938/1946

# How deep are different forms of digital skills divide among young people?

## Results from an extensive survey of 1000 northern-Italian high school students

**Marco Gui and Gianluca Argentin**

**ABSTRACT**

This article outlines the main results of an extensive survey on the digital skills of a random sample of 980 third-year high school students in Italy. The test used in the survey covers three main dimensions of digital skills (theoretical knowledge, operational skills, and evaluation skills) and includes knowledge questions, situation-based questions and tasks performed online. A Rasch-type model was used to score the results and an adjusted regression analysis was undertaken to investigate whether a skills divide based on characteristics such as gender and family education exists among the students. The results indicate that while family education level has a relevant impact with respect to physical access conditions at home, this variable is associated with small differences in students' level of skill. The results for gender, were similar in terms of skills differences but there were no significant differences in terms of access.

These results support the hypothesis that in an environment such as northern Italy (the *Trentino* region), where good quality and relatively equal educational opportunities are in place and where schools provide frequent opportunities to use the web, differences in the characteristics examined in this study do not have a substantial impact on the digital skills level among young people.

## 1. INTRODUCTION

In research on the digital divide, attention has been devoted to differences in skills and usage patterns, seen as an increasingly important aspect of social inequality in the use of new communication technologies. Since 2000, in fact, figures on internet penetration in developed countries have shown a reduction in the gap between high and low income segments of the population as well as between the well-educated and less well-educated segments (NTIA, 2002; Horrigan and Rainie, 2002a,b; Eurobarometer 56-63, 2001-2005, cited in Van Dijk, 2006). However, among the population which is - or soon will be - online, further 'deepening divides' may be emerging depending on the abilities and the opportunities to use these new communication tools for social and personal development (Van Dijk, 2005). This problem has been appropriately labelled the 'second-level digital divide' (Hargittai, 2002).

Theoretical frameworks have been proposed to extend the concept of digital divide beyond physical access (Steyaert, 2000; Mossberger et al., 2003, Di Maggio et al., 2004; Liff and Shepherd, 2005; Van Dijk, 1999, 2003, 2005; Hargittai, 2007). All these frameworks have identified 'skills access' or 'ability to use' as crucial dimensions of 'digital inequality'. Within the skills dimension, there is often a distinction between a technical/operational dimension and an informational one. Van Dijk (2005, 20-22) proposes a very detailed typology: 'operational skills', 'information skills' (further differentiated into 'formal' and 'substantial') and 'strategic skills'.

Empirical sociological research has shown differences in actual digital skills among the population, depending on economic, educational, geographical, and demographical disparities (Hargittai, 2002; De Haan, 2003; Gui, 2007; Van Deursen and Van Dijk, 2008). The most significant differences have been found for the 'information skills' category (Eshet Alkali, 2004; Van Dijk, 2006, 229; Gui, 2007; Rothbauma et al., 2008).

However, problems emerge when we try to attribute meaning to the evidence that has emerged so far. First, performance tests on digital skills have so far been conducted with small and non-random samples, while extensive surveys are usually limited to self-perceived skills and to knowledge of web-related terms. According to Van Dijk (2006, 232), in the study of digital skills there is both a lack of in-depth analysis and of extensive empirical research employing multivariate analysis techniques (Van Dijk, 2006, 232).

Secondly, as the digital divide problem is in a phase of rapid development, we do not know whether the differences that have been found constitute a permanent or a temporary phenomenon. It is possible that the skills divide is displayed very differently within the younger generation which is growing up with new communication technologies in comparison to the adult population. Today's teenage population in western countries, which was only two or three years old when the internet appeared, does not show the same differences in terms of physical access: in some areas, the binary divide between 'haves' and 'have-nots' no longer applies to young people, as Livingstone claims in relation to the UK (2007, 676).

More and more schools are offering an internet connection so that access is free and easy for many high school and college students. In the European Union (27 countries) the percentage of 16 to 24 year-olds having used the internet in the three months before the Eurostat survey was 85%, and in some northern European countries, it was close to 100% (Eurostat, 2007). By focusing research on digital skills on the population as a whole, it should be feasible to limit the risk of detecting temporary disparities. While physical access is spreading almost universally to young people, what is happening to skills differences?

In this survey, we aimed to advance knowledge of digital skills differences among teenagers, based on a large sample – as compared to existing research - and to develop a measurement methodology that uses performance tasks on different skill levels. An in-depth test to investigate the level of digital skills was developed with the intention of addressing the three main dimensions of what the literature defines as 'digital skills': theoretical knowledge, operational abilities and evaluation skills. The test was administered to a random sample of about 1000 Northern Italian high school students in the region of Trentino. This area is technologically and economically advanced when compared with the rest of Italy.

Moreover, in this region the welfare system works well and schools have made noteworthy investments in information and communication technology (ICT). The Trentino region is also among the highest-performing areas in Italy on the Programme for International Student Assessment (PISA) surveys (see OECD, 2000, 2003, 2006). the analysis of these data give us insight into how different aspects of the digital divide - and the skills divide, in particular - can be manifested in a young age group which has grown up with digital media and has enjoyed good and relatively equal educational opportunities.

The main questions we address are the following: how much are sociological characteristics (gender and family socio-cultural background) associated with disparities in access and skills? How big are the differences in students' performance on the three dimensions of digital skills measured in the test?

## 2. DESIGN AND METHODOLOGY

The research, *The digital skills of young students in Trentino* project, was funded by two public institutions of the Trentino Province (ITC-IRST and IPRASE) and is part of a wider Provincial project named 'e-Society'. The aim of the study was to develop an experimental online test to measure the level of digital skills of high school students. This initiative emerged in response to the need of the local school system for both a measurement tool and as a basis for learning initiatives. The research team included sociologists, statisticians, computer science experts and public sector managers[1].

### Test design

The test used in this research was developed by the team through a multi-phase process. First, a review of relevant literature was carried out, examining quantitative and qualitative research on social differences in digital skills and particularly the measuring tools used in those studies. The most common methodology in this kind of research, especially in national surveys, is self-assessment using questionnaires (see for example Bonfadelli, 2002; Hargittai, 2005; Eurostat's annual model surveys on ICT). Given the reliability issues linked to this approach, laboratory tests with real tasks to perform online were implemented (Hargittai, 2002; Gui, 2007; Van Deursen and Van Dijk, 2009).

However, the task-based method has only been applied to small, non-representative samples as it is expensive and presents difficulties in being administered. Moreover, respondents' performance is usually measured on the basis of only a few tasks and without a measuring scale. Another category of measuring tools is represented by tests administered in simulated environments, although they are not developed for social research (they do not collect background information and they are do not involve random samples). Within this approach, the European Computer Driving Licence (ECDL) has so far been the most reliable tool for

[1] See acknowledgements below for details.

measuring digital skills. While it has been limited traditionally to the use of common computer applications (see Van Dijk, 2005, 76-77), an additional interesting instrument has been developed to cover web-related skills needed 'to participate in society as an e-citizen' (www.ecdl.ie). Nonetheless, being based on a fixed and simulated environment, it is still far from a task-based measurement approach which is suggested in the relevant literature, by the experts consulted and the research team conducting this study as the best way to measure the different dimensions of digital skills.

To combine the need for an extensive survey with a measurement procedure based on real task performance, we used a questionnaire approach (suitable for a large sample) enriched by situation-based questions and assignments to perform online (through active links in the question text).

A second phase of the work was focused on the definition of the constructs to be measured and the measurement items of the test. For organisational reasons, the team decided to limit the test to computer-related activities, in particular, those carried out online. Therefore, the test does not measure digital skills linked to the use of different devices or environments (e.g. MP3s, digital television, smart phones).

It was also decided that, apart from the skills which constitute the core of the test, knowledge-based questions could serve as indicators of the broader meaning of "digital skills". Knowledge items were considered which - although not strictly necessary for carrying out online activities - provide students with awareness of the processes underway while they are online. This knowledge can be crucial in problem solving, understanding risks and being creative with technology.

It was therefore decided that the test should include knowledge questions and actual tasks (situation-based questions and tasks to be performed online). We were inspired by Van Dijk's model of digital skills (Van Dijk, 2005) in distinguishing between operational skills (needed to operate the computer and applications to navigate on the internet) and evaluation skills (needed to select, evaluate and re-use information). In summary, we considered three main areas of skills on the basis of which the test was developed[2]:

[2] These categories were inspired by the model proposed by Van Dijk (2005). In a first phase we also planned to develop items concerning the ability to participate in social networks and to be an active contributor to the web (the so-called web 2.0). However, technical and organisational problems emerged and were considered too complex to address in an extensive survey.

1. theoretical knowledge/awareness (33 items in 27 questions);
2. operational skills (27 items in 14 questions);
3. evaluation skills (25 items in 8 questions)[3].

The first area included knowledge-based questions. The second one included Van Dijk's 'operational' and 'formal information skills' together. Research carried out by Gui (2007) has shown how these two dimensions of Van Dijk's model can be unified to examine how they are influenced by differences in socio-educational characteristics. 'Substantial information skills' show different relationships with gender and education capital (ibidem). The 'evaluation skills' section covers this last dimension, simulating results pages in search engines and testing the level of awareness about information evaluation practices. For example, as can be seen in figure 3, a screenshot of the Google results page was presented for the query 'Globalisation'. The question asks the students to visit the links (which remain active during the test) and then to answer the questions about the nature of the site, the organisations behind them and possible biases in the information presented.

In line with the approach of the larger e-society project, we considered digital skills in the light of the perspective of the general framework given by European policies on the 'Information Society' (see European Commission, 2001, 2007). This framework considers ICT skills mainly as tools for achieving full citizenship, for obtaining opportunities in the employment market, and also more generally for improving individuals' quality of life.

In order to validate the first draft of the test and to generate additional suggestions, 13 qualitative interviews were conducted with national and international experts in the field of digital skills. In figure 1, examples of questions in the first area of the test are shown.

[3] The number of questions decreases for the third aspect because the problems presented to the students in that part are complex and need more time to be solved. This is one of the trade-offs in our measurement tool..

**Figure 1 - Examples of items in the theoretical part of the test, translated into English (correct answers are underlined)**

- **What is a website?**

- <u>A collection of web pages organised under the same domain in the World Wide Web</u>
- A collection of files, connected to each other and located in a specific server
- A part of the hard disk where the World Wide Web is kept
- A software which enables a user to access other World Wide Web users' PC
- I don't know

- **The typical path of an email message from a sender to a receiver is:**

1. Senders's PC - any email server - Receiver's PC
2. <u>Sender's PC - sender's email server - Receiver's email server - Receiver's PC</u>
3. Sender's PC - search engine - other PCs - forum - Receiver's PC
4. Sender's PC - chat - Receiver's PC
5. I don't know

- **The Desktop is:**

1. <u>A folder like any other</u>
2. A special folder: it is not contained in the hard disk
3. A special folder: it is contained in the RAM
4. It is not a folder; it is an independent tool
5. I don't know

Examples of questions in the 'operational skills' and 'evaluation skills' parts of the test are listed in Figures 2 and 3.

**Figure 2 - Examples of items in the operational skills part of the test, translated into English (correct answers are underlined)**

- **Take a look at the following images and indicate what they refer to (one answer for each image)[4]:**

1. A blog
2. A P2P software
3. A commercial website
4. A browser

- **You are working on your PC and you find a very useful website you want to visit again over the following days. As the link is very long and complex, how could you record it and easily find it again?**

1. I would use the 'Backup' function
2. <u>I would use the 'Favorites' function</u>
3. I would use the 'Defrag' function

[4] Screenshots were shown in these questions representing different applications and websites; the right definition was to be associated with the right image.

4. I would use the 'Find' function
5. I don't know

- **Surfing on the website www.barilla.it (the link is active) find how many minutes it takes to cook the 'conchiglie rigate' [ribbed shells] pasta variety.**
  The answer is: 12 minutes

**Figure 3 - Examples of items in the evaluation skills part of the test, translated into English and adapted (correct answers are underlined)**

- **Searching for information about 'globalisation' you get the following results on Google. After having visited them (the link is active), choose the right description in the two menus below[5].**

**Globalization - Wikipedia**, the free encyclopedia
Globalization (or **globalisation**) is the term to describe the way countries are becoming more interconnected both economically and culturally. **...**
en.wikipedia.org/wiki/Globalization - 154k - Cached - Similar pages

Which of the following does this site refer to? MENU A
What is the opinion you expect this site to have about globalisation? MENU B

**TUC Globalisation**
Raises key issues and highlights trade union responses to aspects of particular interest to working people. By the Trade Unions Congress, United Kingdom.
www.**tuc**.org.uk/**globalisation**/ - 20k - Cached - Similar pages

Which of the following does this site refer to? MENU A
What is the opinion you expect this site to have about globalisation? MENU B

**Globalisation Management Strategies Conference**
The GMSC **Conference** addresses the issues of managing **globalisation** in the internet age: Executive **Management**, **Strategy** & Planning, Global Marketing **...**
www.**globalisation**.org/ - 19k - Cached - Similar pages

Which of the following does this site refer to? MENU A
What is the opinion you expect this site to have about globalisation? MENU B

MENU A:

1. Open encyclopaedia (correct answer for the first result)
2. Christian volunteering association
3. Lay volunteering association
4. Political party/Trade Union (correct answer for the second result)
5. Governmental
6. University research
7. Professional meeting/association (correct answer for the third result)

[5] The test questions of this type were presented with many simultaneous results, simulating a complete Google results page.

MENU B:

1. Basically negative/critical/sees it as a risk (correct answer for the second result)
2. Basically positive/in favour/sees it as an opportunity (correct answer for the third result)
3. Basically neutral; I expect it doesn't take sides (correct answer for the first result)
4. I don't know

- **You search for the term 'tree' on Google and you get a large number of results. In your opinion, how are these results ordered in the results page?**

1. It would depend on how many times the word 'tree' is written in the pages
2. By date, starting from the most recent
3. Mainly on the basis of how many and what kind of websites link the results with the word 'tree'
4. By the level of reliability of the content
5. In alphabetical order

A pre-test study was carried out in four classes in the Milan area and the results confirmed the reliability of the test (see the *Results* section) and helped to resolve several technical and administration problems. The final test has a total of 85 items.

The test has been implemented as a tool to be administered online with the Mod_Survey (www.modsurvey.org) software. The software was set up to record both answers and completion time (the latter is not discussed here). A detailed questionnaire was added before the test in the administration interface about both social background and the use of ICT at home and at school. The questions, one per page, appear together with a clock counting down from a suggested maximum time for completion. This was done to limit the total duration of the test which appeared to be too long in the first sessions of the pre-test. Although these time limits were not mandatory, they were useful since they gave the subjects an indication of when they should abandon a question they were unable to answer. The non-mandatory time limit was also shown to increase the subjects' concentration. In this way, the questionnaire and the test were engineered to take approximately one and a half hours to be completed. In the pre-test this appeared to be, on average, suitable for the students.

The tool identifies each subject with a unique serial number that is provided by the researchers and inserted at the beginning of the session in a designated box. The internet connection was active throughout the test, as it was not possible to limit its use to the questions that require actual activities online. This might have increased the risks of misuse and copying from other questions, where navigation was not permitted. For this reason, the

tasks to be performed online were concentrated at the end of the test[6]. In this final part a coloured layout appeared in the test interface, making it easy for the researchers who controlled the administration of the test to distinguish between permitted and non-permitted internet surfing[7], by monitoring students' screens.

**Sample and test administration**

A random sample of third-year high school classes in the Trentino area was selected for the research from all the schools in the region; this sample was stratified by school type and geographical position in the region. There were no refusals from schools or classes, probably because the institutions were involved in the project; at student level, all those at school on the test day filled in the questionnaire and the test[8]. A total of about 1,000 students from 65 third-year high school classrooms participated in the study and 980 completed the test, resulting in valid cases for the analysis[9]. This sample is representative of high-school students in the Trentino area.

The demographics of the sample participants were as follows: 50.6% of cases are girls and 49.4% are boys; the birth year distribution varies from 1987 (one case) to 1992 (two cases), with a concentration on the modal birth year - 1991 (72% of cases) - followed by 1990 (21%) and 1989 (6%); 9% of students were not born in Italy. The distribution of the students' cultural background is indicated in the *Sample size* column in table 1. The physical conditions of access of the sample are above the national averages: 66% of the sample has a broadband internet connection at home and 15% have no access at home; 41.9% use the internet every day, while 4.9% never use it.

The test was administered between October and December 2007. In agreement with each school, a two-hour lesson in the computer room was scheduled for each class in order for the test to be administered. The teacher and a researcher were present during the test. The students were required to complete the test and submit data by selecting an appropriate

[6] This also emerged as a useful way to keep students' attention during the test by engaging them in the most stimulating exercises when boredom was more likely to occur.

[7] Contrary to our expectations, we did not detect misuse of the internet during both the pre-test and the test.

[8] Only one person was not computer literate at all and for this reason could not participate in the survey. Due to absence from school on the test day, we lost 9% of our sample, but we can assume that this selection bias is quite a minor one, considering that absence from school is usually a random event.

[9] The remaining cases were eliminated due to disability problems of the subjects, peer-cheating or random filling in as observed by the researchers.

command when they finished. Their instructions forbade them to skip any of the questions, although the option *I don't know* was almost always featured among the possible answers.

Each student was provided with a random two-digit number and a shared identification number for each class with which they accessed the online tool. Therefore the test was anonymous but the data allow for comparisons between different classes and schools.

## 3. RESULTS

The aim of the analysis was to test whether a skills divide exists among the students according to differences in their various characteristics and to compare evidence of a divide with the physical access divide at home. To do this, we examine the impact of two of the most important variables (gender and family cultural background, assessed through parents' educational level) on the presence of ICT at home and on students' digital skills, measured using the test discussed above. These are the variables that emerge in the literature as being among those that have the strongest influence on access and on skills disparity.

### Differences in access

We begin by focusing on ICT access at home, looking at the presence of incremental levels of ICT equipment. Using the data emerging from the pre- test questionnaire we considered: i) the presence of a personal computer, ii) the presence of an internet connection (of any kind), iii) the availability of a broadband connection, iv) the presence of a personal computer with a broadband connection used only by the student (a sort of ideal situation, where the student has his/her 'personal digital desk'). In table 1, we summarise the results.

**Table 1 - Presence of incremental level of ICT equipment at home by gender and family education (%)**

| | Personal computer | Connection (any kind) | Broadband connection | Broadband + 'Personal digital desk' | Sample size |
|---|---|---|---|---|---|
| | | | | | |
| **Gender** | | | | | |
| Male | 95.0 | 79.1 | 65.3 | 38.0* | 496 |
| Female | 97.2 | 85.7 | 61.7 | 30.0* | 484 |
| | | | | | |
| **Family education** | | | | | |
| Lower secondary | 93.5* | 73.5* | 54.4* | 30.9 | 230 |
| Vocational (2 years) | 93.1* | 80.6* | 58.1* | 28.1 | 160 |
| Secondary (5 years) | 97.4* | 86.1* | 65.1* | 35.9 | 387 |
| Degree | 99.0* | 87.2* | 74.9* | 38.9 | 203 |

**Note:** the symbol '*' in the table means that the *p* value of the Chi square test for that table is less than 0.05

There are small differences between males and females regarding ICT access at home, except for the situation we defined as *personal digital desk*, where males have a small advantage over females. Instead, we note that family education plays an important role in determining this 'first-level digital divide', having an impact on all our indicators of ICT access at home. At this level, statistically significant differences emerge from our data. These differences grow as the levels of ICT equipment improve. Therefore, if we only consider home access to the internet, family background is shown to have an important influence despite the fact that our sample is made up of young students and the survey was conducted in an economically advanced area.

We note that parents' education level has a positive impact on ICT equipment at home, but not on students having their own 'digital desk' (family social class has instead a relevant impact also on this last variable). This mixed effect may be due to the fact that parents with a higher level of education have and use a personal computer themselves and so their children are more likely to share their PC with them.

## Test evaluation

Before focusing on the digital skills divide issue, we discuss the scoring method and the reliability of the test developed for this research.

To score the results of the test we could have adopted an additive measure - such as for example counting the number of right answers. Quantitative research on digital skills has adopted this basic approach of treating the number of tasks completed successfully and the time spent on each task as dependent variables (Hargittai, 2002; Gui, 2007; Van Deursen and van Dijk, 2009). The main problem with this approach is that in the final score questions are not weighted by their level of difficulty. For this reason, we preferred a Rash-type 'partial credit model' (Masters, 1982) to score the results which is the gold standard for competencies measuring (e.g. PISA surveys). With this model it is possible to obtain scores where every item has a specific weight correlated to the respondents' success rate in answering it. Moreover, one is able to test the entire measurement tool and to evaluate whether it is yielding a unique concept; finally, respondents and items are evaluated on the same scale which gives us the possibility to check if the test is correctly set to the respondents' ability level. On this basis, the model provides an ability estimate (score) for each subject and a difficulty estimate for every item[10].

Table 2 reports the means and standard deviations of students' scores calculated with the Rash model on a logit scale: first, the scores of the total test, and then the partial score for each of the three subdimensions discussed above (theoretical, operational, evaluation).

**Table 2 - Score means, standard deviation, minimum and maximum value: total and three dimensions (n=980)**

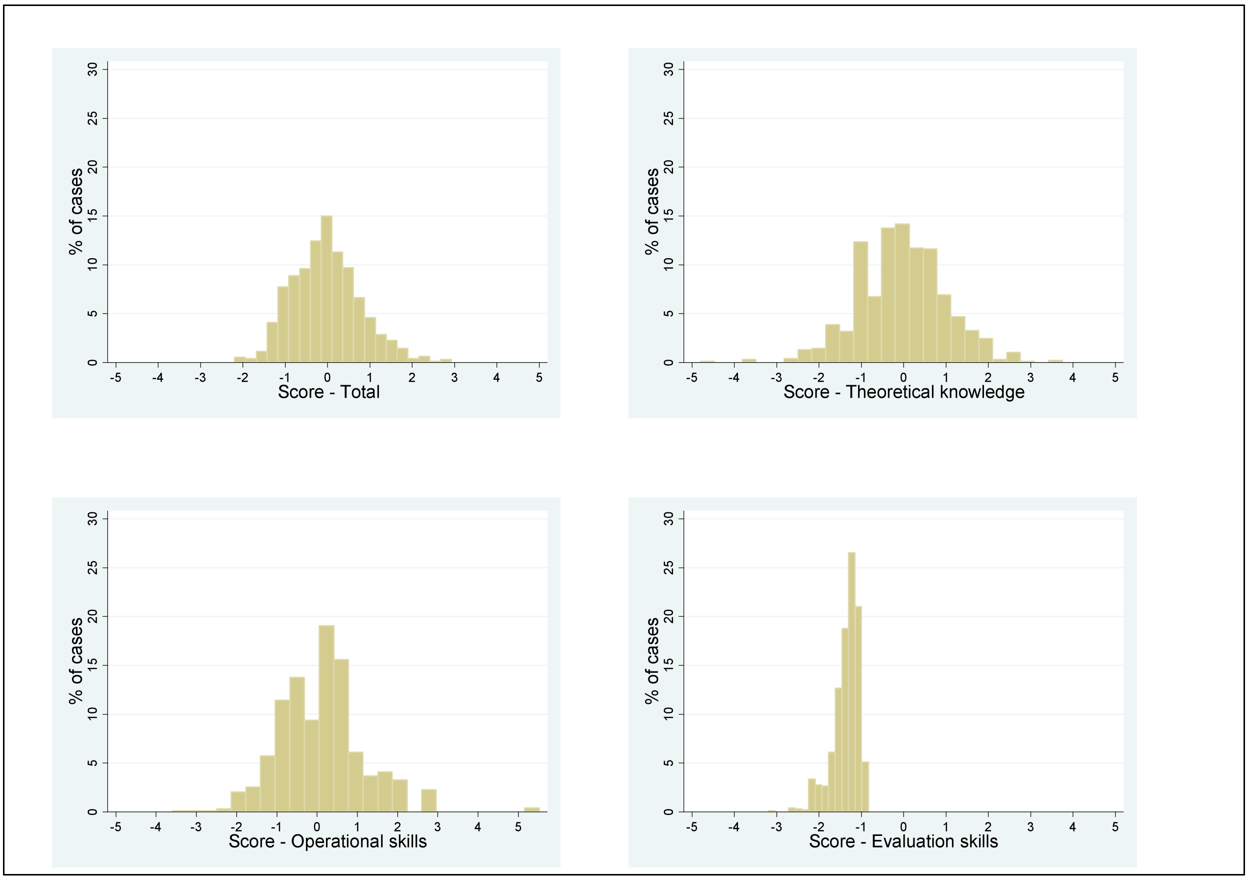
| | **Total score** | **Theoretical knowledge score** | **Operational skills score** | **Evaluation skills score** |
|---|---|---|---|---|
| Mean | - 0.024 | - 0.050 | 0.138 | - 1.351 |
| *Standard deviation* | *0.799* | *1.001* | *1.068* | *0.317* |
| | | | | |
| Minimum | - 2.195 | - 4.800 | - 3.593 | - 3.191 |
| Maximum | 2.924 | 3.423 | 5.519 | - 0.830 |

[10] To validate the test, firstly, a confirmatory factor analysis supported its uni-dimensionality (a prevalent factor explains 64% of the variance). Infit and Outfit values remained within standard levels, confirming the adequacy of questions in relation to the subjects' ability. There were two mis-fitting items, which were eliminated from the overall score and analysed separately. Person reliability is 0.81 and the percentage of mis-fitting respondents is 2.1%.

According to the Rasch model, obtaining a mean equal to zero is evidence of the good calibration of the test to the population. The results show that the test was well-calibrated to the skills level of our students, with the exception of the evaluation skills part which has a very low mean (-1.35).

If we look at the standard deviations and at the range between minimum and maximum values of the total score and the dimension scores, we observe that the evaluation skills part shows a smaller range of variation. We suspect that the measurement of this part of the test could be problematic in the sense that it may be argued that those questions were too difficult for the sample. This is clear in the graphs below: we have very good pseudo-normal distributions for the total score and for the theoretical and the operational dimensions of the test. The same is not the case for the evaluation dimension. We only had eight questions to measure this dimension because the time needed to process and solve the tasks was longer than for the other two dimensions: this could have had an impact on the reliability of this dimension.

**Figure 4 - Distribution of scores: total and the three dimensions (% - Base=980)**



We can also compare the scores obtained by students on the three dimensions. We can conclude that the students attain better results for the operational skills items and significantly worse results for the evaluation items. This finding is in accordance with the existing literature (Eshet-Alkali, 2004; Gui, 2007; Van Deursen and Van Dijk, 2008), but we need to be cautious in interpreting the results.

In conclusion, except for the third dimension, the test appears to be reliable: it generates a pseudo-normal distribution with zero mean and with enough variability in the score assignments.

**Differences in skills**

Turning to the core of our analysis: testing whether characteristics such as gender and family education background affect the students' scores in the test and, hence, whether these characteristics generate a digital divide not only in ICT access at home, but also with respect to digital skills. To test this hypothesis, we compare the score mean differences by independent variable.

To interpret the mean differences more easily the score values have been rescaled. Each score now has a mean equal to 100 and a standard deviation equal to 50; the distribution of cases under the Normal curve suggests that about 95% of the students should reach a score between 0 and 200[11]. Our data show this range of variation, providing additional confirmation that the scores of the subjects show a pseudo-normal distribution.

In table 3, we show the results of this analysis for the total score and also for the scores of the three dimensions, each rescaled as explained. Each cell contains the score mean of the subgroup for the part of the test indicated in the column (total, theoretical knowledge, operational and evaluation skills) and in brackets its standard deviation. We also provide significance tests for the mean differences, comparing each group with the reference category (*ref.* - "male" for gender and "lower secondary" for family education)[12]. For

[11] Obviously, this is not true for the third dimension of the test.

[12] To evaluate the significance of mean differences between groups with normal distributions, it is possible (and usual) to apply the F test. We prefer not to use this statistical instrument because there are two possible violations of its assumptions within our data. The first one is that there are different variances between subgroups (we applied a Levene's test and the result was significant and hence we cannot assume equal variance in the subgroups); the second one is that data were not collected from a sample of individuals, but from a sample of classrooms (hence, it is difficult to assume that there is no correlation between errors). To solve these

example, for the family education rows: we compare vocational with lower secondary, then secondary with lower secondary and finally degree level with lower secondary.

**Table 3 - Rescaled score means and standard deviation (total and three dimensions) by gender and family education**

| | **Total Score** | **Theoretical knowledge score** | **Operational skills score** | **Evaluation skills score** | **Size of sample** |
|---|---|---|---|---|---|
| | | | | | |
| **Gender** | | | | | |
| Male *(ref.)* | 107.6 *(57.3)* | 110.0 *(55.6)* | 104.6 *(56.7)* | 100.7 *(52.9)* | 496 |
| Female | 92.5* *(40.4)* | 90.3* *(41.7)* | 95.5 *(42.0)* | 99.3 *(47.0)* | 484 |
| | | | | | |
| **Family education** | | | | | |
| Lower secondary *(ref.)* | 93.2 *(49.6)* | 95.3 *(47.2)* | 91.7 *(48.2)* | 97.6 *(48.9)* | 230 |
| Vocational (2 years) | 93.2 *(50.4)* | 94.7 *(53.5)* | 94.1 *(46.6)* | 93.8 *(53.1)* | 160 |
| Secondary (5 years) | 102.0* *(46.0)* | 101.0 *(47.0)* | 102.5* *(49.8)* | 102.1* *(48.8)* | 387 |
| Degree | 109.4* *(55.6)* | 107.7* *(54.9)* | 109.3* *(53.2)* | 103.6* *(50.7)* | 203 |

**Note:** the symbol '*' in the table near a category signifies that the mean of that group is significantly different from the mean of the reference category *(ref.) :* the *p* value is less than the usual threshold of 0.05. For details about the significance test see note 12.

The table shows that there are mean differences that are not random and confirm the existence of a digital skills divide; hence, gender and family background differences are associated with inequality in line with the literature. This is not entirely so for all the four scores because not all differences are significant, but we observe that for many comparisons, the *p values* are lower than 0.05.

What is more interesting is that all the differences are very small if we compare them to the score range of variability. This means that the impact of gender and family education are

problems we used an adjusted linear regression model for every single independent variable, accounting for heteroskedasticity and sample clusterisation in the estimation of standard errors (we used the options "Robust" and "Cluster", employing Stata9) and, in this, we obtained correct significance test (we do not report all regression coefficients and standard errors, but they are available from the authors on request). Anyway, we observe that, using a multiple regression model or improperly using the F test, our results would even be reinforced. We also tested the differences showed in the table, using the confidence intervals approach; our conclusions remain the same also with this analysis.

significant but not very strong[13]. For example, let us consider the biggest mean difference in the table, that between male and female on theoretical knowledge. This difference is 19.7 points (110.0-90.3) and this is not a wide gap, considering that the range of the score is about 200 points and that the standard variation is 50 points. When we look at the confidence interval of the estimation, this difference varies from 8.3 to 31.1 points (95% level of confidence); and accepting the highest value of the estimate, the means difference is not great, compared to the range variation and to the standard deviation. This is so for all our independent variables and for all our dependent scores, especially considering that many of the other differences are clearly smaller than the one in this example.

For gender, it is also interesting to note that the difference between male and female is mainly due to a different distribution of the two groups within the score range: males show a greater presence among high-performance respondents. This is also reflected in the standard deviations which are higher among males.

## 4. DISCUSSION

These results assist us in answering the two research questions in the introduction to this paper.. First, we consider the impact of gender and family cultural background on disparities in access and skills. Gender does not have a relevant impact on ICT access while it does have has a statistically significant impact on skills (in the total score and in the theoretical dimensions). Family education, in contrast, produces big and significant differences on ICT access, except in the case of "personal digital desk". Family education produces some significant differences with respect to digital skills only when we consider the highest segments as compared to the lowest.

We observed that the width of skills differences for both gender and family education are not very large: the most relevant difference is that between males and females on theoretical skills (19 points on a 0-200 scale). It should be noted, however, that males have scores with higher variability than females and show a different distribution with a higher concentration in the topmost performance groups.

[13] We also used a model with the two independent variables together, obtaining similar results. We presented the significance test of the three separated models in the table because the aim of our use of regression is not to estimate the net effects of every single variable but to correct the possible bias that we would meet using the normal standard errors or the F test.

It is not possible to directly compare data on ICT access with skills differences observed with the test because the measurement units as well as the variable characteristics are very different. Nevertheless, we can suggest on the basis of these data that the impact of family background is stronger on ICT access than on the digital skills level.

With respect to the second question, the sample performed better on operational skills than on theoretical skills and significantly lower on evaluation skills. The impact of the gender and education level variables was not different between the three dimensions of skills. The much lower performance of the sample on the third dimension (evaluation skills) was present irrespective of the students' social backgrounds.

To interpret these findings, reflections are needed at both the methodological and theoretical levels. First, as for the representativeness of these results, the size of the sample and the fact that it was selected randomly offers more scope for generalisation in comparison with past studies in the field. At the same time, our findings are not generalisable to the digital divide situation in different socio-economic contexts and geographical areas. More data from performance tests using our tool in different countries is needed to validate our conclusions.

Second, although the test overall was shown to be reliable (the scores display a normal distribution), we need an improved test for the evaluation skills section. It is possible that the test was not able to measure the meaningful differences on the third dimension (evaluation skills) due to the excessive difficulty of the items. This may have hidden a specific skills gap in evaluation skills at a lower level of difficulty.

Finally, there is the more complicated issue regarding the validity of this test. It cannot be considered as an exhaustive measure of digital skills. The concept is too broad to be covered uni-dimensionally since it includes dimensions as diverse as technical operation and socio-emotional abilities (Eshet-Alkali, 2004). Nonetheless, this test can be taken as a good measure of those skills that are needed for information seeking and social participation on the web and these skills are important prerequisites for educational activities, political awareness and civic participation. The results of this study do appear to tell us something about how digital divide problems can develop in situations where: i) people have grown up with digital media, ii) there is a good socio-economic situation and iii) there are relatively equal schooling opportunities.

## 5. CONCLUSIONS

In this paper we have presented and discussed the results of a study measuring the digital skills of about 1,000 teenagers attending the third year of high school. For this, a test was developed and implemented, containing both survey questions and performance tasks, and covering several dimensions of digital skills. This study reports the first extensive survey on digital skills based on tasks to be performed online. Its extensive coverage of three dimensions and the measurement approach provides a basis for generalisation to a greater extent than past studies in the field.

The results show that, consistently with the literature, there is a high degree of variability in the scores obtained by the subjects. In line with earlier work, (Eshet Alkali, 2004, Gui, 2007, Van Deursen and Van Dijk, 2008), the sample performed better on operational skills than on theoretical skills and was especially poor on evaluation skills, although the validity of this latter finding is unclear.

In line with Van Dijk (2006) and others, we found that family background has a strong impact on ICT access conditions at home (presence of a computer and a connection at home, availability of broadband) and access disparities are significantly associated with family education (and socio-economic status). However, surprisingly, these differences do not automatically translate into a stratification of skills levels. Family education level does not seem to have a strong impact on performance: we found some statistically significant effects but the width of these differences was small. This result is in contrast with most of the existing literature. For gender we found some statistically significant effects only on the total score and on the theoretical dimension of the test.

Although we cannot directly compare the results on the access divide and those on the skills divide it is plausible to argue that we found smaller differences on skills level between groups by family education as compared to differences on domestic ICT equipment. In the case of gender, the opposite is the case.

Acknowledging the limitations of this study, how do these data match with theoretical expectations and past results of digital divide research?

While the theoretical literature has suggested the hypothesis of a 'second level digital divide' which includes skills and usage type as "deepening" dimensions of the digital divide (Van Dijk, 2005), based on our results for our sample the skills divide seems to be smaller than the access divide at home. In environments such as Trentino, it is not the case that familiar social and cultural differences in society are reflected in evidence of variations in digital skills. It seems that some factors are leading to the homogenisation of the skills of the subjects, irrespective of their origins. This may be due to the equalising effects of both the schools and the students' peers. We intend to test this hypothesis in the future.

The general claim of digital divide research - that the differences between social categories, which were already un-equal in terms of 'old' resources and capital, are amplified by the use of digital media - seems to be challenged by these results at least with respect to digital skills differences in younger age groups. These findings suggest a picture of a possible future trend in the normalisation/stratification problem of access to digital media (Norris, 2000; Van Dijk, 2005). It seems that while stratification still exists with respect to access conditions at home, normalisation is instead to be expected with respect to the digital skills of the younger generation.

It is possible, alternatively, that socio-economic differences translate more directly into different types of usage (this will be investigated with the same data in future analysis). While the skills gap is about the differential ability to use the media, the usage gap is a 'broader thesis about a differential use of whole applications in daily practices' (Van Dijk, 2006). This field of stratification might be appearing more clearly (see Livingstone, 2006; Cotten and Jelenewicz, 2006) and may emerge as the crux of social differentiation in the use of digital media in the long-run.

This first extensive survey of actual digital skills and its unexpected results indicate the urgent need for more standardised and shared measurement tools in this field. We encourage the use of statistically rigorous measurement techniques, such as those used for international research on education performance in the PISA surveys in order to obtain reliable and comparable results in this field of digital competence. If we want to have a clear picture of what the skills divide is like among young people, and generally among the entire population, we also need measurement tools which: i) use a large range of items, ii) address different dimensions of digital skills (operational, formal and substantial), and iii) possibly take into account different communication practices (information seeking, communication, e-

commerce). We would also encourage large scale surveys based on random samples as only in this way will it be possible to generalise research results and analyse differences between sub-populations, avoiding potential biases due to local factors (as in this research) or variations in administration context conditions (as in most research in this field).

Finally we suggest that future research should pay particular attention to what we have called "evaluation skills", or - in Van Dijk's words - "substantial information skills". On the one hand, our results show that this is the most difficult dimension to measure accurately. On the other hand, the lower performance of the sample on this part of the test seems to confirm the findings of the existing literature, suggesting that this dimension of digital competence is the one most in need of new media literacy interventions among young people.

**Acknowledgements**

This research is part of the 'e-society' programme, funded by the 'Provincia Autonoma di Trento' (www.provincia.tn.it), FBK-IRST (www.fbk.eu) and IPRASE (www.iprase.tn.it), IARD Institute (www.istitutoiard.it ).

As well as the authors of this work, the research team included: Carlo Buzzi (University of Trento), Bruno Caprile (FBK-IRST), Vincenzo D'Andrea (University of Trento), Andrea Dipace (IARD Institute), Gianni Lazzari (FBK-IRST), Romano Nesler (IPRASE), Roberto Ricci (IRRER), Chiara Tamanini (IPRASE). Special thanks go to the panel of experts who were interviewed during the initial stage of the project and to the researchers who helped us during the test administration.

## REFERENCES:


Cotton, S.R. and Jelenewicz, S.M. (2006) 'A Disappearing Digital Divide Among College Students? Peeling away the Layers of the Digital Divide', *Social Science Computer Review* 24 (2): 497-506.

De Haan, J. (2003) 'IT and Social Inequality in the Netherlands', *IT & Society* 1(4): 27-45.

DiMaggio, P., Hargittai E., Celeste C. and Shafer S. (2004) 'From Unequal Access to Differentiated Use: A Literature Review and Agenda for Research on Digital Inequality', in K. Neckerman (ed.) Social Inequality, York: Russell Sage Foundation.

Eshet-Alkali, Y. (2004) 'Digital literacy: a conceptual framework for survival skills in the digital era', *Journal of Educational Multimedia and Hypermedia* 13(1): 93-106.

European Commission (2001), e-Inclusion: The Information Society's potential for social inclusion in Europe, See URL: http://ec.europa.eu/employment_social/soc-dial/info_soc/esdis/eincl_en.pdf [Last Consulted 22 May 2009].

European Commission (2007), To be part of the information society, See URL: http://ec.europa.eu/information_society/activities/einclusion/docs/i2010_initiative/comm_native_com_2007_0694_f_en_acte.doc [Last Consulted 22 May 2009].

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), See URL: http://www.firstmonday.org/

Hargittai, E. (2002) 'Second-Level Digital Divide: Differences in People's Online Skills', *First Monday* 7(4), See URL: http://www.firstmonday.org/

Hargittai, E. (2005) 'Survey Measures of Web-Oriented Digital Literacy', *Social Science Computer Review.* 23(3): 371-379.

Hargittai, E. (2007). 'A Framework for Studying Differences in People's Digital Media Uses', in Kutscher N. and H. Otto (eds.), *Cyberworld Unlimited*, pp. 121-137, Verlag für Sozialwissenschaften/GWV Fachverlage GmbH.

Horrigan, J. and Rainie L. (2002a), 'Getting serious online: as Americans gain experience, they pursue more serious activities', *Washington DC: Pew internet and American Life Project*, See URL: http://www.online-publishers.org/media/202_W_PIP_Getting_Serious_Online3ng.pdf [Last Consulted 22 May 2009].

Horrigan, J. and Rainie, L. (2002b), 'The broadband difference: how online behavior changes with high-speed internet connections. *Washington DC: Pew internet and American Life Project*, See URL: http://www.pewinternet.org/~/media//Files/Reports/2002/PIP_Broadband_Report.pdf.pdf [Last Consulted 22 May 2009].

Liff, S. and A. Shepherd(2004), 'An Evolving Gender Digital Divide?' *Oxford internet Institute, internet Issue Brief,* No. 2, see URL: http://www.oii.ox.ac.uk/resources/publications/IB2all.pdf [Last Consulted 22 May 2009].

Livingstone S. and Helsper E. (2007) 'Gradations in digital inclusion: children, young people and the digital divide', *New Media and Society* 9(4): 671-696.

Martin S.P. (2003) 'Is the digital divide really closing? A critique of inequality measurement in 'A nation online', *IT & Society* 4(1): 1-13.

Masters, G.N. (1982) 'A Rasch model for partial credit scoring', *Psychometrika* 47: 149-174.

Mossberger, K., Tolbert C.J. and Stansbury M. (2003) *Virtual Inequality: Beyond the Digital Divide,* Washington, DC: Georgetown University Press.

Norris, P. (2001) *Digital Divide, Civic Engagement Information Poverty and the internet Worldwide*, Cambridge: Cambridge University Press.

NTIA (National Telecommunications and Information Administration) (2002), 'A nation online: how Americans are expanding their use of the internet', See URL: http://www.ntia.doc.gov/ntiahome/dn/anationonline2.pdf/ [Last Consulted September 2008].

Organisation for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2000

Organisation for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2003

Organisation for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006

Rothbaum F., Martland N. and Beswick Jannsen J. (2008) 'Parents' reliance on the Web to find information about children and families: Socio-economic differences in use, skills and satisfaction', *Journal of Applied Developmental Psychology* 29(2): 118-128.

Van Deursen, A. and Van Dijk J. (2008) 'Measuring Digital Skills. Performance tests of operational, formal, information and strategic internet skills among the Dutch population', paper presented at the 57th Annual conference of the International Communication Association (ICA), Montreal, 22-26/08.

Van Deursen, A.. and Van Dijk J. (2009). 'Improving digital skills for the use of online public information and services'. *Government Information Quarterly*

Van Dijk, J. (1999) *The Network Society: Social Aspects of New Media*, London: Sage Publications.

Van Dijk, J. (2003) 'A framework for digital divide research', *Electronic Journal of Communication/Revue de Communication Electronique* 12(1), see URL: http://www.cios.org/EJCPUBLIC/012/1/01211.html [Last Consulted 22 May 2009].

Van Dijk, J. (2005) *The deepening divide. Inequality in the information society*, London: Sage Publications.

Van Dijk, J. (2006) 'Digital divide research, achievements and shortcomings', *Poetics* 34: 221-35.

Van Dijk, J. and Hacker K. (2003) 'The Digital Divide as a Complex and Dynamic Phenomenon', *The Information Society* 19: 315-27.

Steyaert, J. (2000) 'Digitale vaardigheden: Geletterdheid in de informatiesamenleving', Den Haag: Rathenau Instituut.

# Electronic Working Papers

Media@lse Electronic Working Papers will:

- Present high quality research and writing (including research in-progress) to a wide audience of academics, policy-makers and commercial/media organisations.
- Set the agenda in the broad field of media and communication studies.
- Stimulate and inform debate and policy. All papers will be published electronically as pdf files, subject to review and approval by the Editors and will be given an ISSN.

An advantage of the series is a quick turnaround between submission and publication. Authors retain copyright, and publication here does not preclude the subsequent development of the paper for publication elsewhere.



**Notes for contributors:**

Contributors are encouraged to submit papers that address the social, political, economic and cultural context of the media and communication, including their forms, institutions, audiences and experiences, and their global, national, regional and local development. Papers addressing any of the following themes are welcome:

- Communication and Difference
- Globalisation and Comparative Studies
- Innovation, Governance and Policy
- Democracy, Politics and Journalism Ethics
- Media and Identity
- Media and New Media Literacies
- The Cultural Economy

Contributions are welcomed from academics and PhD students. In the Michaelmas Term each year we will invited selected Master's students from the preceding year to submit their dissertations which will be hosted in a separate part of this site as 'dissertations' rather than as Working Papers. Contributors should bear in mind when they are preparing their paper that it will be read online.

Papers should conform to the following format:

- 6,000-8,000 words (excluding bibliography, including footnotes)
- 150-200 word abstract
- Headings and sub-headings are encouraged
- The Harvard system of referencing should be used!
- Papers should be prepared as a Word file.
- Graphs, pictures and tables should be included as appropriate in the same file as the paper.
- The paper should be sent by email to:

Dr. Bart Cammaerts, Deputy Editor, Media@lse EWP-Series (b.cammaerts@lse.ac.uk)
Department of Media and Communications
Houghton Street
London
WC2A 2AE

ISSN: 1474-1938/1946