This is a pre-print version of the article forthcoming in New Media and Society

The digital skills of Internet-natives

The role of ascriptive differences in the possession of different forms of digital literacy in a random sample of northern Italian high school students

Marco Gui
University of Milano-Bicocca

Gianluca Argentin
University of Milano-Bicocca

ABSTRACT

This article outlines the main results and methodological challenges of a large scale survey on actual digital skills. A test covering three main dimensions of digital literacy (theoretical, operational and evaluation skills) was administered to a random sample of 65 third-year high school classes, producing data on 980 students. Items include knowledge questions, situation-based questions and tasks to be performed online. A Rasch-type model was used to score the results. In agreement with literature, the sample performed better in operational skills, whilst showing a particularly poor performance regarding evaluation skills (although for this dimension the test shows reliability issues). Through a robust regression analysis we investigate if a skills divide based on ascriptive differences, gender and family cultural background, exists among the students. It emerges that cultural background has a significant effect, which is stronger on operational skills,

while gender shows a more definite impact on theoretical knowledge. Methodological problems related to the measurement process are discussed and it is pointed out that a lack of standardised criteria to interpret substantive score differences exists in this field of study.

INTRODUCTION

Research on the digital divide has devoted growing attention to differences in skills, seen as an increasingly important aspect of social inequality in the use of new communication technologies. Theoretical frameworks extending the concept of digital divide beyond physical access have all identified digital skills as a crucial tool for the social inclusion and professional development of individuals (Steyaert, 2000; Castells, 2001; Mossberger et al., 2003, Di Maggio et al., 2004; Liff and Shepherd, 2005; Van Dijk, 2005). Digital skills differences are a primary aspect of what has been called the *second-level digital divide* (Hargittai, 2002).

As figures on internet penetration in developed countries show a clear reduction of the gap between genders, as well as – at a lower pace - between segments of the population with a high and low level of education (Mossberger et al., 2003; Oecd, 2007; Bentivegna, 2009), it has been claimed on the contrary that digital skills differences could manifest themselves in a more stable and even increasingly profound manner (Van Dijk, 2005).

Empirical sociological research has shown relevant 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).

However, a number of problems come about when trying to attribute a meaning to the evidence emerging so far. Firstly, performance tests on digital skills have so far been conducted on small samples, while extensive surveys are usually limited to self-perceived skills and to the knowledge of web-related terms. Secondly, as in some respects we are still undergoing an initial phase in the spread of digital media, we do not know whether the differences that have been found so far constitute a permanent or only a temporary phenomenon. Among today's teenagers in Western

Countries – who were only two or three years old when the Internet appeared – differences in terms of physical access are almost no longer relevant. Nowadays schools are increasingly offering an Internet connection so that access is free and easily available for many high school and college students. In some areas the binary divide between haves and have-nots no longer applies to young people, as Livinsgtone claims in relation to the UK (2007, 676). In the European Union (27 Countries) the percentage of 16 to 24 year-olds who use the internet regularly[1] was 83% in 2008, and in some northern European Countries it was close to 100%. In Italy this percentage is 64% (source: http://ec.europa.eu/eurostat).

Focussing the research regarding digital skills on such a population should limit the risk of detecting temporary disparities. While physical access spreads to the point of reaching the near total number of young people, what exactly happens when it comes to their skills differences? How does this generation perform in the different dimensions of digital literacy?

In this survey, we intend to offer an extensive picture on digital skills possession among young people which have grown up with digital media. On the one hand, we aim to advance the knowledge of digital skills differences among teenagers, thanks both to a larger sample – when compared to existing research – and to a measurement methodology which makes use of performance tasks on different skill levels and a statistically rigorous approach. On the other, we offer schools and educational institutions an insight into what specific aspects of digital skills are mastered to a greater and to a lesser extent by teenagers.

LITERATURE REVIEW

In the sociological research on digital skills, gender and cultural background are the ascriptive social characteristics that have been used the most as independent variables. Education level appears to be the most relevant factor connected to differences in digital skills (Hargittai 2002; Gui, 2007, Van Deursen and Van Dijk, 2009). Gui (2007) finds that, when age is kept constant,

[1] At least once a week on average within the last three months before the survey.

education level and parental education show a relevant impact on the ability to solve complex research tasks on the web.

Gender has shown to have a significant impact on the level of self-perceived skills and knowledge of web related terms (Liff and Shepherd, 2005; Gui, 2007; Hargittai and Hinnant, 2008). Data in the international and Italian context show that young females use the web almost as frequently as their male counterparts but that they tend to use it in a more instrumental way without going into exploration and technical details (Liff and Shepherd, 2005; Gui, 2007). However, research that compared men and women in their ability to solve actual tasks online only found little disparity (Hargittai, 2002; Hargittai and Shafer 2006; Van Deursen and Van Dijk, 2009).

Literature also distinguishes between a range of aspects of digital skills on which social characteristics produce different impacts. A consolidated distinction is the one between a technical/operational aspect of digital skills, needed to operate a computer and a browser, and an informational aspect, linked to the ability to select, evaluate and reuse information (Steyaert, 2000; Mossberger et al., 2003; Van Dijk, 2005). On the basis of the Van Dijk model (Van Dijk, 2005, 20-22), Van Deursen and Van Dijk (2009) propose a very detailed typology which we will use as a scheme of reference in this article: 'operational skills' (needed to operate computers and network hardware and software), 'formal skills' (ability to understand and handle the formal characteristics of computer network and web environments) and 'information skills' (skills to select, evaluate and process information)[2]. Research carried out by Gui (2007) has shown how socio-economic characteristics produce a small impact on operational skills and formal skills, where age is constant. Their possession is connected mainly with frequency of use and access conditions. On the contrary, performances at the 'information skills' level show strong differences depending on education capital (ibid.).

Also, empirical evidence has shown that young people differ from the average population as to their possession of digital skills. It has emerged that while they usually exhibit high operational and formal skills, they experience more problems than adults for what regards information skills (Eshet-Alkali and Hamburger, 2004; Nielsen, 2005; Van Deursen and Van Dijk, 2009). Children and teens show poor research skills (Livingstone, 2003) and a restricted awareness of the risks of their social activities online (Ofcom, 2009). While differences of access and experience with some activities on the web appear widespread among young users, those uses which are more capital-enhancing remain restricted to the segments with a high social-cultural background (Livingstone, 2007).

Notwithstanding these remarkable results, according to Van Dijk (2006, 232), in the study of digital skills there is both a lack of in-depth analyses and of extensive empirical research based on multivariate analysis (Van Dijk, 2006, 232). The economic and organisational costs of testing people's performance in solving actual online tasks have limited the scope of digital skills research. Furthermore, laboratory research on digital skills has so far adopted a basic approach to scoring results and building digital skills indexes, considering the number of tasks completed successfully, the time spent on each task or the number of problems encountered as dependent variables (Hargittai, 2002; Gui, 2007; Van Deursen and van Dijk, 2009). The main problem with these approaches is that questions are not weighted by their actual level of difficulty in the final score composition; moreover, it is not possible to evaluate the reliability of the measurement methodologies and their internal coherence.

HYPOTHESES

On the basis of existing literature we consider the following hypotheses regarding expected disparities in the possession of different dimensions of digital skills and the role of ascriptive characteristics in producing disparities among teenagers.

[2] Van Dijk (2005) also defines 'strategic skills' – the capacity to use all other forms of digital skills – as the means for specific goals and for the general goal of improving one's position in society. In this study we will not discuss this last

Hypothesis 1a: Young people with higher parental education will exhibit higher levels of digital skills.

Hypothesis 1b: Parental education will show a positive impact especially with regards to information skills, showing a smaller effect on operational skills.

Hypothesis 2. Being male will be positively associated to the knowledge of web related terms and the awareness of technological concepts, while gender differences will not be significantly correlated to actual skills differences.

Hypothesis 3. Young people will show a high performance in operational skills, less high with regards to digital media related knowledge and a much lower performance on information skills.

METHOD

An in-depth test to investigate the level of digital skills has been developed with the intent to cover some of the main aspects of what literature defines as 'digital skills'. The test was administered to the students of a random sample of 65 third-year high school classes in the Trentino area, producing data on 980 students[3]. The sample of classes was selected out of the totality of third-year high school classes in the region and stratified by school type and geographical position. The sample is therefore representative of high school students in the Trentino area, but not in the Italian context. Trentino is economically and technologically advanced when compared with the rest of Italy: the unemployment rate in the first quarter of 2009 was 3.9% versus a national value of 7.3%; the percentage of the population living in a household with a personal computer in 2006 was 62.6% in Trentino and 57.1% in the rest of Italy. Finally, Trentino is also among the highest-performing areas in Italy on the PISA surveys (Gentile, 2009). This geographical specificity implies some limitations in the inferences that will be discussed in our conclusions.

Sample

dimension, which is very difficult to observe (see van Deursen and Van Dijk, 2009 for the first attempt to do this).

[3] The test was administered to a total number of 1043 students. 63 cases were eliminated due to connection problems during submission, N disability problems of the subjects, N peer-cheating and N random answering as observed by the researchers.

Following are a range of descriptive statistics about the sample demographics: 50.6% of cases are girls and 49.4% are boys; the age distribution varies from 20 years old (one case) to 15 (two cases), with a high concentration on the modal age of 16 (72% of cases) - followed by 17 (21%) and 18 (6%). Among the students, 9% were not born in Italy. The distribution of students' cultural background is as follows: 23.5% have parents at lower secondary education level or less, 55.8% at upper secondary education level, and 20.7% at higher education level. The physical conditions of access of the sample are the following: 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.

Test design

The test used in this research was developed by a team of sociologists, statisticians, computer science experts and ICT public sector managers[4] through a multi-phase process. Firstly, a review of relevant literature was carried out, analysing quantitative and qualitative research on social differences in digital skills and particularly on the measuring tools used in those studies. The most common methodology in this kind of research, especially in national surveys, is self-assessment through questionnaires (see for example Bonfadelli, 2002; Hargittai, 2005; see also Eurostat's annual model surveys on ICT, http://ec.europa.eu/eurostat). Given the reliability issues linked to such an approach, laboratory tests with real tasks to perform online were implemented (Hargittai, 2002; Eshet-Alkali and Hamburger, 2004; Gui, 2007; Van Deursen and Van Dijk, 2009). However, the latter method was only applied to small samples as it is much more expensive and presents difficulties in being administered; moreover, respondents' performance is usually measured on the basis of only a few tasks and without a real measurement scale. Another category of measuring tools is represented by tests administered in simulated environments[5], but they are not engineered

[4] See acknowledgements for details.
[5] See for example the European Computer driving License (ECDL) and its new product "e-citizen" (www.ecdl.ie), the test *iSkills*™, http://www.ets.org, and in the Italian context the Digital Competence Assessment project, www.digitalcompetence.org.

for social research as they do not collect background information and they do not involve random samples.

To combine the large scale nature of the survey with the need to observe real task performances, we used a questionnaire approach (suitable for a large sample) where multiple-choice questions are also applied for testing actual skills online. In these questions, the answers can only be identified by following active links in the question text and completing a task performed online. We are aware that such a research methodology presents limitations as to the scope of information gathered: it only considers the correctness of assignments without looking at the respondents' actions when solving them. Nonetheless, we argue that a multiple choice approach with questions that imply actual navigation is an efficient tool to collect data on actual digital skills of large scale samples and to gain the possibility of validating the reliability of the measuring tool through rigorous statistical methods. However, to vary the kind of items on which the test is built we also inserted some open questions, where respondents must submit a written answer.

A second phase of the work was focussed on the definition of the exact objects to be measured and the items of the test to measure them. First of all, 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, smartphones…).

It was also decided that, apart from actual skills, knowledge based questions could also be used as indicators of "digital skills" in its wider-ranging meaning. Items of knowledge have been considered which – although not strictly necessary for carrying out online activities – provide students with the awareness of the processes underway whilst they are online. This 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). For this last part, we were inspired by Van Dijk's model of digital skills (Van Dijk, 2005) in distinguishing between operational and formal skills (necessary to operate a computer and efficiently navigate on the internet) and information skills (needed to select, evaluate and re-use information).

Hence, at the end of the process, we considered three main areas of skills on the basis of which the test was developed[6]:

1. theoretical knowledge/awareness (33 items) ;
2. operational skills (27 items);
3. evaluation skills (25 items)[7].

While the first area includes knowledge-based questions as mentioned before, the second one refers to Van Dijk's 'operational' and 'formal information skills' together. This area tests the ability to use computer applications but also to recognize specific web environments and to navigate efficiently. Our 'evaluation skills' section instead covers Van Dijk's 'substantial information skills', testing the level of awareness and the actual skills in information evaluation practices.

In order to validate the first draft of the test and to have additional suggestions, 13 qualitative interviews were conducted with national and international experts in the field of digital skills.

In the following figures, examples of questions in the first area of the test are shown.

Figure 1 here

Examples of questions in the 'operational skills' and 'evaluation skills' parts of the test are listed below.

Figure 2 here

Figure 3 here

---

[6] These categories were inspired by the model proposed in 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 for an extensive survey.

[7] It should be noted that in the third part of the test the items were grouped into only eight questions, while in the theoretical and operational part the questions were respectively 27 and 14. This was considered useful because evaluation tasks are more complex and time-consuming. This is one of the trade-offs in our measuring experience.

A pre-test study was carried out in four third-year high school classes in the Milan area[8] whose results showed the reliability of the test (see the *Results* section) and helped to improve several technical and administration problems.

The test was materially implemented as a tool to be administered online with the Mod_Survey (www.modsurvey.org) software. 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 timer counting down from a suggested maximum time for completion. This has been done to limit the total duration of the test which had appeared too long in the first sessions of the pre-test. Although these time frames were not mandatory, they have been shown to be useful since they gave the subjects an indication of when they should abandon a question they were unable to answer. A non-mandatory time limit suggestion also showed to increase the subjects' concentration. In this way, the questionnaire and the test together were engineered to take approximately one and a half hours to be completed. In the pre-test this amount of time appeared to be, on average, suitable for the students.

The internet connection was active during the entire test, as it was not possible to limit its use to the questions that require actual activities online. This, however, could have implied risks of misuse and copying in the other questions, where navigation was not permitted. For this reason the tasks to be performed online were concentrated at the end of the test[9]. In this final part a coloured layout appeared in the test interface, making it easy for the researchers who controlled the administration to distinguish between permitted and non permitted internet surfing[10], by monitoring students' screens.

[8] The choice of an area outside Trentino for the pre-test phase ensured that no classes potentially to be included in the sample were "wasted". The Milan area was also chosen because of its proximity to some of the researchers involved. In is interesting to note that this area presents similar characteristics in comparison with the Trentino area both regarding the economic situation and with respect to ICT access, with penetration rates above the national average.

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

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

Scoring method

As mentioned earlier, additional measures (like counting the number of right answers) present significant problems when used to score results of performance tests. For this reason, we preferred a Rash-type ‘partial credit model’ (Masters, 1982) to score the results, which is the gold standard for competency measuring (e.g. OECD-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 measuring tool and to evaluate whether it is really 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 on the respondents’ ability level. According to the Rasch model, obtaining a mean close to zero is a proof of the good calibration of the test on the population. On this basis, the model provides an ability estimate (score) for each subject and a difficulty estimate for every single item.

Independent variables

Following our hypotheses, we consider two independent variables: gender and family social background. We measure the latter using the higher education credential among student’s mother and father. In this way we distinguish four levels of educational background: lower secondary or less (primary and lower secondary diplomas are considered together because of the limited number of cases); upper secondary diplomas, consistently with Trentino’s educational system, are instead split into *Diploma di Qualifica* (2 years of school) and *Diploma di Maturità* (5 years); finally we consider all higher education qualification together (again because of the limited number of cases). We decided not to consider family social class for two reasons. Firstly, because it would mediate the parental education effects. In this way, we would not have been able to capture the overall impact of cultural background. Secondly, it would have been problematic to consider both variables together in the same model because of collinearity problems, i.e. due to their strong correlation and the available number of cases. Moreover, we observed that at a bivariate level parental education produces stronger differences than family social class, a finding which is consolidated in educational surveys.

Test administration

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 classroom in order for the test to be administered. There were no refusals from schools or classes, probably because of the institutions involved in the project. At student level, all the people at school on the test day completed the assignment[11].

The teacher and a researcher were present during the entire duration of the test to avoid cheating and solve problems. The students were required to complete the test and submit data by selecting an appropriate command when finished. The instructions forbade them to skip any of the questions, although the option *I really 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 comparisons between different classes and schools.

RESULTS

In presenting the results, we will first describe the adequacy of our measurement tool, which will also lead us to compare the students' performance in the three skills dimensions measured by the test. Then we will investigate whether a skills divide exists among the students according to the ascriptive differences chosen as independent variables: parental education and gender.

Test evaluation and students' overall performance

[11] 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.

A preliminary analysis has been carried out on the data to evaluate the main aspects of reliability of the test developed for this research. Firstly, a confirmatory factor analysis supported its unidimensionality (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. Person reliability is 0.81% and the percentage of mis-fitting respondents is 2.1%.

Table 1 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, and of evaluation).

Table 1 here

The results show that the test was well calibrated on the skills level of our students, with the exception of the evaluation skills part which has a very low mean value (-1.35).

If we look at the standard deviations and at the ranges between minimum and maximum values of the total score and the dimension scores, we observe that the evaluation skills part also shows a smaller range of variation. We suspect that the measurement of this part of the test could have been affected by some problems and, more specifically, we feel that those questions were too difficult for the sample. This argument becomes clear when looking at Figure 4: 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 true for the evaluation dimension. More specifically, we are missing the right part of the normal distribution curve for evaluation skills and this confirms that the questions were too difficult also for the best performing students.

Figure 4 here

If we assume that also this third dimension was calibrated on a adequate skills standard for 15-16 year-olds, we can use this result to compare the scores obtained by students in the three dimensions. Their best performance is in the operational skills score. They perform a little worse at the theoretical level and significantly worse in the evaluation skills part. This finding is in accordance with existing literature (Eshet-Alkali and Hamburger, 2004; Gui, 2007; Van Deursen and Van Dijk, 2008). However, as we said before, there are good reasons to be cautious about the fitness of our test at the evaluation skills level.

In conclusion, except for the third dimension, from these analyses the test appears to be reliable: we generate a pseudo-normal distribution with zero mean and with enough variability in the scores assignment.

Differences in skills due to ascriptive characteristics

Finally, we can focus on the core of our analysis: testing whether ascriptive characteristics (gender and family education background) affect the students' scores in the test and hence, assessing whether these characteristics generate a digital divide with respect to digital skills. To test this hypothesis, we have run four regression models, one for every score considered as dependent variable (total, theoretical, operational and evaluation skills). To easily interpret the mean differences the score values have been standardised and every score now has a mean equal to 0 and standard deviation equal to 1. In Table 2 we show the results of this analysis, reporting the four regression models containing parental education and gender as independent variables[12].

Looking at the results of significance tests, the table shows that there are mean differences that are not random and prove the existence of a digital skills divide; hence, gender and family background generate inequality, according to literature. This is not entirely true for all the four scores because

[12] We used an adjusted linear regression model containing all the dummies for the two independent variables, accounting for heteroskedasticity and sample clusterisation in the estimation of standard errors (we used the options "Robust" and "Cluster", employing Stata9). We also observed that, our results would even be reinforced using a regression model for each variable or improperly using the F test (there are two possible violations of its assumptions within our data: there are different variances between subgroups; data were not collected from a sample of individuals, but from a sample of classrooms, hence there is correlation between errors). We also tested the differences showed in

not all differences are significant, but we note that for many comparisons (seven out of 20) we observed *p values* lower than 0.05.

More precisely, looking at the total score, we find significant differences in the students' performance both by gender and social origins. The cumulative effect of the two ascriptive variables is 0.62 standard deviations, which separates males with high educated parents from females with lower secondary or less educated parents. It should be considered that the effect of the two variables on the overall score resides in different parts of the test. For what concerns gender, it is interesting to note that the differences are mainly due to a better performance of males in theoretical skills. Instead, social origins play a higher role in operational skills and, secondly, in theoretical ones, but only when we consider the highest groups against the lowest. Gender and cultural background, instead, do not produce significant differences in evaluation skills, but we should remember the possible weakness of our test on this point.

We also ran a model where there was an interaction between the two independent variables (gender, *parental education[13]), using the total score as dependent variable. It is interesting to observe that the cultural background gap in performance is higher among males. Male students from high cultural backgrounds performed 0.39 standard deviations better than male students coming from family with a lower cultural background, while among females this difference is only 0.07 standard deviations.

Table 2 here

DISCUSSION AND CONCLUSIONS

the table, using the confidence intervals approach instead of significativity tests. Our conclusions remain the same also with this kind of analysis.

In this paper we have presented the results of a study measuring the digital skills of 980 teenagers attending the third year of high school. For this, an *ad hoc* test was implemented, containing both survey questions and performance tasks, and covering three dimensions of digital skills.

With these data we are endeavouring to supply the three hypotheses presented in the introduction with an answer. Firstly, we consider the impact of parental education on disparities in skills. We expected (Hypothesis 1) that parental education could produce significant differences in the possession of digital skills. In agreement with literature (see van Dijk, 2006), our data support to this hypothesis. According to Hypothesis 1a we expected that the impact of parental education was the strongest for what regards information skills but that it produced only small differences in operational skills. The data falsify this hypothesis, as the biggest differences found on the basis of parental education are at the operational skills level, then at the theoretical level. Differences at the evaluation skills level are not significant.

We find full evidence for Hypothesis 2 as gender is a significant factor in producing differences only at the level of theoretical knowledge and it is not relevant in differentiating the level of skills in the operational and evaluational dimensions. A significant effect is also visible in the total score but this is due for the greatest part to the differences found at the theoretical level. In terms of inequality this might mean that females are usually as skilled as males in routine activities online but that they might experience more problems when facing technical problems or unexpected outcomes.

We have also checked the strength of the interaction effects of the two variables finding that parental education is a more important factor of differentiation among males than among females. A possible interpretation of this finding relates to specific different access conditions between males and females in our sample, as pointed out in another analysis on the same data set (Gui and Argentin, 2009), which could confer a relative advantage to males with a high social-cultural background. This hypothesis however needs further analysis.

[13] In this model, with the aim of being more parsimonious, we distinguish only between parents with tertiary or upper secondary (5 years) diploma and the others, with lower credentials. The coefficient table is not shown here, but available

Given the existence of statistically significant differences in our research results, it is much more complicated to interpret their substantive relevance. If we calculate score differences in terms of standard deviation units and use the conventional categories of small effects (more than 0.2 standard deviation) medium effects (more than 0.5) and large effects (more than 0.8 – see Cohen, 1992), one can argue that we only found small effects by both gender (0.40) and parental education (0.30). Only when we consider their cumulative impact (0.62), the numbers would go beyond the threshold of medium effects. This uncertainty in the interpretation of the findings leads us to consider how research on digital skills urgently needs standardised criteria to substantively comment on differences between individuals and groups when considered with statistical measuring tools.

Finally, the data supports Hypothesis 3. In accordance with existing literature (Eshet Alkali and Hamburger, 2004; Gui, 2007; Van Deursen and Van Dijk, 2008), the sample performed better in operational skills than in theoretical skills and was particularly poor on evaluation skills. The impact of ascriptive variables does not differ between the three dimensions of skills. The much lower performance of the sample in the third dimension (evaluation skills) also happens irrespective of students' social backgrounds.

To attribute greater significance to these findings, some reflections are needed at both the methodological and theoretical levels. Firstly, as for the representativeness of these results, the size of the sample in this research and the fact that it has been selected randomly allows for wider generalisation in comparison with past studies in the field. At the same time, we cannot argue that our findings can be representative of the digital divide situation in different social-economic context and geographical areas. We need more data from performance tests in different countries to validate our conclusions.

Secondly, as mentioned before, the test overall has shown to be reliable but we need an improved test for the evaluation skills section. It seems that this part of the test does not have the same discriminatory strength of the first two parts. Either the sample has very low and homogenised

from the authors on request.

evaluation skills or the items used to measure this dimension were too difficult even for the most skilled among the students. It is possible that our questions were influenced by the fact that they were originally inspired by previous studies using adults and university students as samples (Wang et al., 2000; Ford et al., 2001; Hargittai, 2002; Gui, 2007).

Finally, a wider discussion regards the validity of this test. Obviously, it cannot be considered an exhaustive measure of digital skills. The concept is too broad to be covered unidimensionally, since it includes as diverse dimensions as technical operationality and social-emotional abilities (Eshet-Alkali and Hamburger, 2004). We considered digital skills not only in terms of actual know-how but also as a measure of the awareness of the technical and logical structures beneath digital environments. Some of these resources (especially when the theoretical part of the test is considered) are not of direct use for ordinary activity online but they are nonetheless important for a critical participation in digital environments, in finding creative solutions, and in being aware of the sources of possible problems.

Despite all these limitations in validity, we underline however that thanks to the chosen measurement approach it was possible – for the first time in this context – to statistically validate the reliability of the measuring tool.

We would therefore encourage the use of statistically rigorous measuring techniques, such as those applied in international research on education performances in OECD-PISA surveys, in order to obtain reliable and comparable results also in the field of digital competence. If we wish to have a clear picture of how the skills divide is configured among young people, and generally among the entire population, we also need measuring tools which i) use a large spectrum 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; 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 happens in this research) or administration context conditions (as in traditional research in this field).

Finally we believe that future research should pay particular attention to what we have called “evaluations 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 be measured accurately. On the other hand, the lower performance of the sample in this part of the test seems to confirm the findings of existing literature and suggests that this dimension of digital competence is one the most in need of new media literacy interventions among young people.

Acknowledgements

This research is part of the project *The digital skills of young students in Trentino*, funded by the ‘Provincia Autonoma di Trento’ (www.provincia.tn.it), FBK-IRST (www.fbk.eu), IPRASE (www.iprase.tn.it), IARD Institute (www.istitutoiard.it), and is part of a wider Provincial project named ‘e-Society’. 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.

TABLES

Table 1 – Students performance on the total score and on the three specific 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

Table 2 – Linear regression models on the rescaled score means (total and three dimensions) (base=980)

	Total score			**Theoretical knowledge score**			**Operational skill score**			**Evaluation skill score**		
	Coef	**Robust SE**	**P value**	**Coef**	**Robust SE**	**p value**	**Coef**	**Robust SE**	**p value**	**Coef**	**Robust SE**	**p value**
Male	0.30*	0.11	0.01	0.40*	0.11	0.00	0.18	0.11	0.09	0.03	0.08	0.49
Female	*ref*	- -	- -	*ref*	- -	- -	*ref*	- -	- -	*ref*	- -	- -
Parental education: Higher education	0.32*	0.11	0.01	0.25*	0.11	0.03	0.35*	0.10	0.00	0.12	0.10	0.16
Parental education: Upper secondary (5 years)	0.18*	0.07	0.01	0.12	0.07	0.09	0.22*	0.07	0.00	0.09	0.08	0.18
Parental education: Upper secondary (2 years)	0.00	0.09	0.69	-0.01	0.10	0.62	0.05	0.07	0.36	-0.08	0.09	0.29
Parental education: Lower secondary or less	*ref*	- -	- -	*ref*	- -	- -	*ref*	- -	- -	*ref*	- -	- -
Constant	-0.29*	0.08	0.00	-0.29*	0.08	0.00	-0.26	0.08	0.00	-0.06	0.08	0.30
R-squared		0.038			0.048			0.025			0.005	

*p<0.05.

FIGURES

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

- **What is a website?**

1. <u>A collection of web pages organised under the same domain in the World Wide Web</u>
2. A collection of files, connected to each other and located in a specific server
3. A part of the hard disk where the World Wide Web is kept
4. A software which enables a user to access other World Wide Web users' PC
5. 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

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)[14]:**

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. 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

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

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 links are active), choose the right description in the two menus below[15].**

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)

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

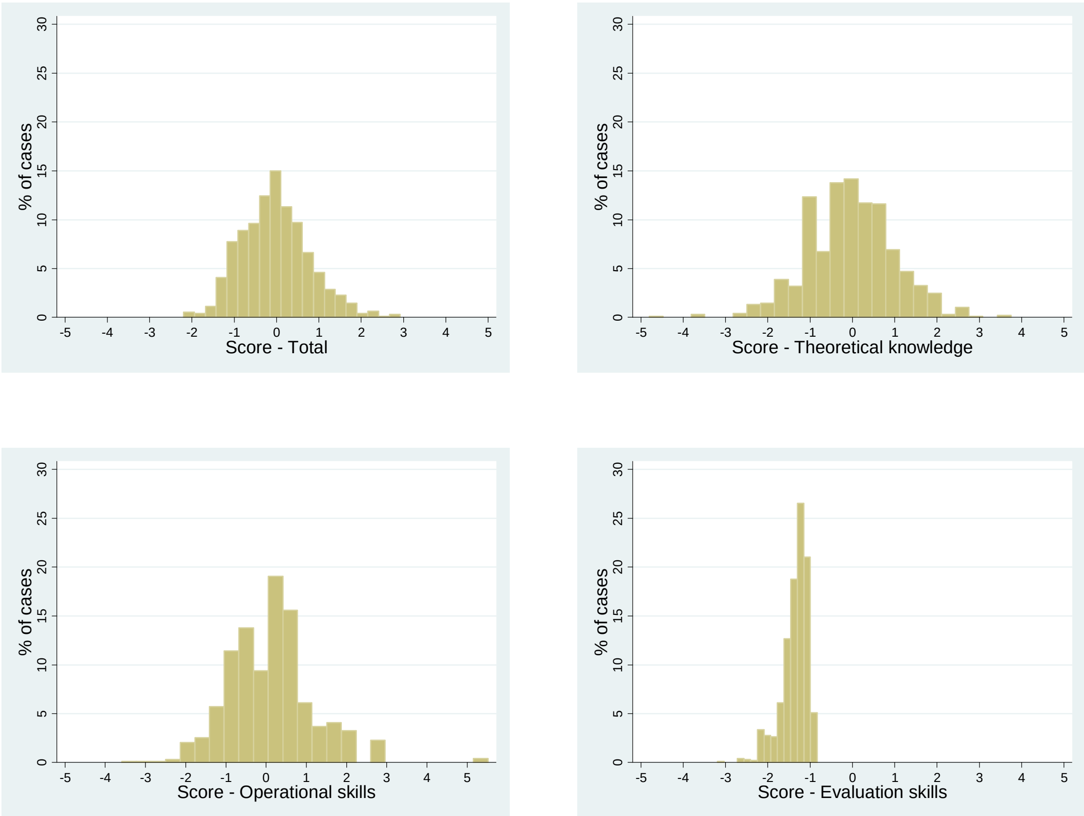
- **You search for the term 'tree' on Google and you get a large number of results. In your opinion, how are these results ordered in the results page?**

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

15 The test questions of this type were presented with many simultaneous results, simulating a complete *Google Results* page.

5. In alphabetical order

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



REFERENCES

Bentivegna, S. (2009), *Disuguaglianze digitali. Le nuove forme di esclusione nella società dell'informazione,* Roma-Bari, Laterza.

Bonfadelli, H. (2002), 'The Internet and knowledge gaps: A theoretical and empirical investigation', *European Journal of Communication*, 17, 1, pagg. 65-84.

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

Cohen, J. (1992) 'A power primer', *Psychological Bulletin*, 112, 155.

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-Alkalai, Y. and Amichai-Hamburger, Y. (2004) 'Experiments in digital literacy'. *Cyberpsychology and Behavior* 7, 421–429.

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].

Ford, N., Miller, D., and Moss, N. (2001), 'The Role of Individual Differences in Internet Searching: An Empirical Study', Journal of the American Society for Information Science and technology, 52(12), 1049-1066.

Gentile M. (ed.) (2009), *Rapporto provinciale PISA 2006, I dati OCSE-PISA 2006 per il Trentino*, IPRASE del Trentino, Trento, See URL: http://www.iprase.tn.it/attivit%C3%A0/studio_e_ricerca/indagini_internazionali/download/IPRASE/PISA_2006.pdf [Last consulted 6 October 2009].

Gui M. and Argentin G. (2009), 'How deep are different forms of digital skills divide among young people? Results from an extensive survey of 1000 northern-Italian high school students', Media@LSE electronic working papers, 15, London School of Economics and Political Science, London, UK.

Gui M. (2007a), 'Formal and substantial Internet information skills: The role of sociodemographic differences on the possession of different components of digital literacy', First Monday, 1 (1), 2007.

Hargittai, E. (2002) 'Second-Level Digital Divide: Differences in People's Online Skills', *First Monday* 7(4).

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

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. (2007). ‘A Framework for Studying Differences in People’s Digital Media Uses' in Kutscher N. and H. Otto (eds.), *Cyberworld Unlimited*, 121-137.

Hargittai, E. and Hinnant, A. (2008). ‘Digital Inequality: Differences in Young Adults’ Use of the Internet’, *Communication Research*, 35(5):602-621.

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. (2003) ‘Children’s use of the Internet: Reflections on the emerging research agenda’, *NewMedia & Society*, 5(2), 147-167.

Livingstone S. and Helsper E. (2007) ‘Gradations in digital inclusion: children, young people and the digital divide’, *New Media & 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.

Nielsen, J. (2005) *Usability of Websites for teenagers*. Fremont, CA: Nielsen Norman Group.

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

Ofcom (2009) ‘Digital Lifestyles. Young adults aged 16-24’. See URL: http://www.ofcom.org.uk/advice/media_literacy/medlitpub/medlitpubrss/digital_young/young_digital_lifestyles.pdf [Last Consulted October 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

Organisation for Economic Cooperation and Development (OECD – 2007), Working Party on the Information Economy. Broadband and ICT access and use by households and individuals, See URL: http://www.oecd.org/dataoecd/44/11/39869349.pdf [Last Consulted October 2009].

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

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

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', *Interacting with Computers.*

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.

Wang, P., Hawk, W. B., and Tenopir, C. (2000) 'Users' interaction with the World Wide Web resources: an exploratory study using a holistic approach', *Information Processing and Management*, 36(2), 229–251.